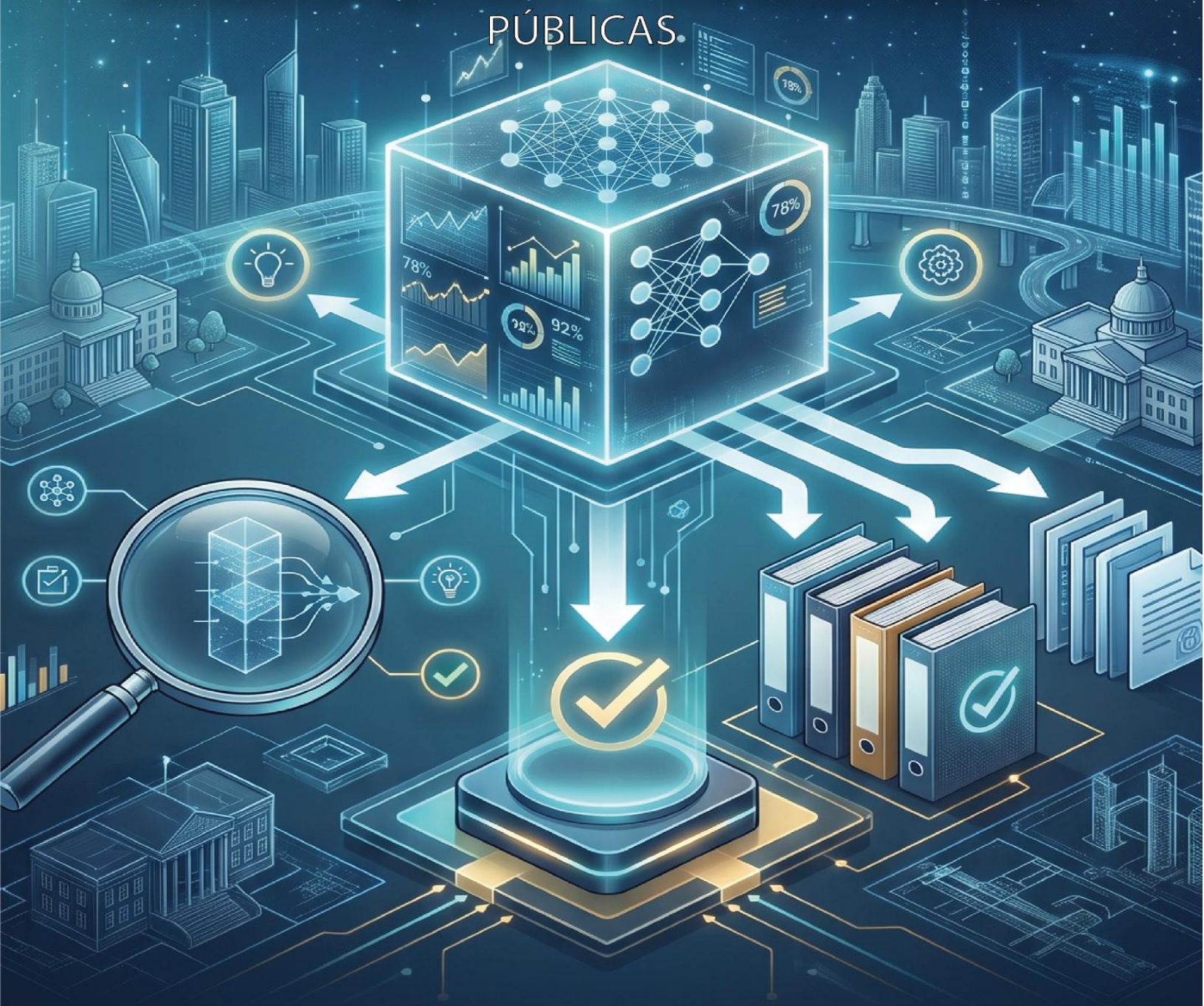
INTELIGENCIA ARTIFICIAL EXPLICABLE Y MODELOS PREDICTIVOS

PARA LA TOMA DE DECISIONES EN SISTEMAS DE RECOMENDACIÓN DE LICITACIONES PÚBLICAS



Miguel Molina Villacís | María Molina Miranda | Angel Cuenca Ortega

Luis Espín Pazmiño | Renzo Padilla Gómez

INTELIGENCIA ARTIFICIAL EXPLICABLE Y MODELOS PREDICTIVOS PARA LA TOMA DE DECISIONES EN SISTEMAS DE RECOMENDACIÓN DE LICITACIONES PÚBLICAS

AUTORES:
Miguel Giovanny Molina Villacís
Universidad de Guayaquil
https://orcid.org/0000-0002-7080-2354
miguel.molinav@ug.edu.ec

María Fernanda Molina Miranda
Universidad de Guayaquil
https://orcid.org/0000-0002-4237-4364
maria.molinam@ug.edu.ec

Ángel Eduardo Cuenca Ortega
Universidad de Guayaquil
https://orcid.org/0000-0001-7798-611X
angel.cuencao@ug.edu.ec

Luis Arturo Espín Pazmiño
Universidad de Guayaquil
https://orcid.org/0000-0002-1663-2489
luis.espinp@ug.edu.ec

Renzo Rogelio Padilla Gómez
Universidad de Guayaquil
https://orcid.org/0000-0003-4301-1335
renzo.padillag@ug.edu.ec

Detalles de publicación	
Título:	Inteligencia artificial explicable y modelos predictivos para la toma de decisiones en sistemas de recomendación de licitaciones públicas
Autores:	Miguel Giovanny Molina Villacís, María Fernanda Molina Miranda, Ángel Eduardo Cuenca Ortega, Luis Arturo Espín Pazmiño y Renzo Rogelio Padilla Gómez
Editorial:	Editorial de Educación, Investigación y Cultura Académica
Diseño de portada:	Editorial de Educación, Investigación y Cultura Académica
Formato:	PDF
Páginas:	246
Tamaño:	A4 (21 × 29.7 cm)
Acceso:	En línea
Fecha de publicación:	17- 07- 2026
ISBN:	978-9907-829-07-5
DOI:	10.5281/zenodo.21362368

Primera edición, 2026.

Publicado por Editorial de Educación, Investigación y Cultura Académica.

Esta obra ha sido sometida a un proceso de revisión por pares ciegos, cumpliendo con estándares académicos y editoriales de calidad bajo la supervisión de la editorial. La editorial garantiza la integridad de dicho proceso; sin embargo, el contenido, la veracidad y la precisión de los datos presentados son responsabilidad exclusiva del autor. Se permite la descarga, reproducción y distribución del libro en cualquier medio, siempre que se reconozca adecuadamente la autoría, no se realicen modificaciones y no se utilice con fines comerciales. Su uso está destinado a fines educativos y de divulgación académica.

® *Inteligencia artificial explicable y modelos predictivos para la toma de decisiones en sistemas de recomendación de licitaciones públicas*

© *2026* Miguel Giovanny Molina Villacís, María Fernanda Molina Miranda, Ángel Eduardo Cuenca Ortega, Luis Arturo Espín Pazmiño y Renzo Rogelio Padilla Gómez

---

Licencia y derechos de uso

La obra *Inteligencia artificial explicable y modelos predictivos para la toma de decisiones en sistemas de recomendación de licitaciones públicas* está licenciada bajo la Licencia Creative Commons Atribución-NoComercial-SinDerivadas 4.0 Internacional (CC BY-NC-ND 4.0).

Se permite la reproducción y distribución del material en cualquier medio o formato, siempre que se reconozca adecuadamente la autoría, no se utilice con fines comerciales y no se realicen obras derivadas.

Editorial de Educación, Investigación y Cultura Académica

Primera edición, 2026

DEDICATORIA

El presente libro se lo dedica a la comunidad académica que busca siempre trabajar en proyectos innovadores utilizando inteligencia artificial y conocimientos de Tecnologías de la Información

AGRADECIMIENTO

Agradecemos a Dios y a nuestras familias por su apoyo incondicional en el desarrollo académico basado en la Investigación y a todas las personas que han aportado con su conocimiento para la realización de este libro, un agradecimiento a Angel Josue Jimenez Villao y Darla Lisbeth Luna Chiriboga por su valiosa colaboración en las pruebas y simulaciones desarrolladas para la elaboración de este libro. Su apoyo técnico, participación en la validación de escenarios y contribución en la verificación de los modelos predictivos fueron fundamentales para fortalecer la calidad y confiabilidad de los resultados presentados.

PRÓLOGO

La incorporación de la inteligencia artificial en los procesos de análisis y toma de decisiones representa una de las transformaciones tecnológicas más significativas de la época contemporánea. Su capacidad para procesar grandes volúmenes de información, identificar relaciones entre variables y generar estimaciones ha ampliado las posibilidades de innovación en organizaciones públicas y privadas. Sin embargo, este avance también ha puesto en evidencia una tensión que no puede ser ignorada: mientras los modelos adquieren mayor capacidad predictiva, sus resultados pueden resultar cada vez más difíciles de comprender para las personas que deben interpretarlos, revisarlos y asumir la responsabilidad de las decisiones derivadas de su uso.

Esta tensión adquiere especial relevancia en la contratación pública. En este ámbito convergen necesidades institucionales, normas jurídicas, especificaciones técnicas, restricciones presupuestarias, condiciones operativas y criterios de evaluación que deben ser analizados de manera articulada. Las decisiones adoptadas no solo producen efectos administrativos, sino que repercuten en la utilización de recursos públicos, la continuidad de los servicios, la ejecución de políticas y la satisfacción de necesidades colectivas. Por ello, cualquier herramienta tecnológica que intervenga en estos procesos debe responder a exigencias de transparencia, trazabilidad, responsabilidad y supervisión humana.

La obra Inteligencia artificial explicable y modelos predictivos para la toma de decisiones en sistemas de recomendación de licitaciones públicas se inscribe dentro de este escenario. Su propósito no consiste en presentar la inteligencia artificial como una solución automática ni en atribuir a los algoritmos una autoridad superior al criterio profesional. Por el contrario, propone una aproximación crítica y metodológicamente delimitada sobre la manera en que los modelos predictivos y los mecanismos de explicabilidad pueden integrarse en una arquitectura experimental de apoyo analítico.

Uno de los méritos de la obra se encuentra en la diferenciación entre predicción, puntuación, recomendación y decisión. Estos conceptos suelen emplearse como si representaran etapas equivalentes; sin embargo, poseen funciones y niveles de responsabilidad diferentes. Un modelo puede estimar una categoría o una probabilidad; un sistema puede ordenar alternativas de acuerdo con una puntuación; y una interfaz puede presentar una opción como resultado de ese procesamiento. La decisión, en cambio, exige interpretar la información, contrastarla con otras evidencias, aplicar criterios técnicos y jurídicos y asumir las consecuencias correspondientes. Esta última responsabilidad permanece en las personas y autoridades competentes.

La inteligencia artificial explicable ocupa un lugar central en esta discusión. La explicación no debe concebirse como un elemento decorativo añadido después de generar una predicción. Tampoco puede reducirse a una gráfica atractiva o a una enumeración general de variables. Una explicación adecuada debe mantener correspondencia con el comportamiento efectivo del modelo, mostrar los factores que influyeron en el resultado, comunicar la dirección de sus contribuciones y advertir sobre las limitaciones de la estimación.

La obra examina esta problemática desde una perspectiva conceptual, tecnológica, metodológica y ética. Sus capítulos presentan los fundamentos de la inteligencia artificial, el aprendizaje automático, los modelos de clasificación, los sistemas de recomendación y las principales técnicas de explicabilidad. Esta revisión proporciona al lector las herramientas necesarias para comprender que el rendimiento estadístico representa solo una dimensión de la calidad de un sistema. También deben considerarse la pertinencia de las variables, la calidad de los datos, la estabilidad, la posibilidad de auditoría, la comprensión del usuario y los riesgos asociados con la automatización.

Otro aporte significativo se encuentra en el tratamiento de la gobernanza de datos. Los algoritmos no operan sobre una realidad neutral y completamente disponible. Procesan representaciones construidas a partir de decisiones previas sobre qué información recolectar, cómo definir las variables, qué registros conservar, qué transformaciones aplicar y qué resultado intentar predecir.

Por esta razón, la calidad y documentación del conjunto de datos resultan tan importantes como el algoritmo seleccionado.

El prototipo presentado utiliza información sintética y posee una finalidad académica y demostrativa. Esta delimitación constituye una fortaleza metodológica, porque permite diferenciar la construcción de una prueba de concepto de la validación de una herramienta institucional. Los datos generados programáticamente permiten comprobar el funcionamiento del código, examinar el flujo de procesamiento y representar mecanismos de explicación. No permiten, sin embargo, evaluar empresas reales, anticipar adjudicaciones ni demostrar efectos sobre la eficiencia presupuestaria.

La transparencia con la que se reconocen estas restricciones fortalece la integridad del trabajo. En un campo caracterizado con frecuencia por afirmaciones excesivas sobre las capacidades de la inteligencia artificial, resulta necesario distinguir entre lo que una herramienta demuestra, lo que potencialmente podría realizar y lo que todavía requiere investigación. Esta obra adopta esa distinción y propone una ruta progresiva para la consolidación futura del prototipo.

El lector encontrará una discusión que trasciende la programación. El diseño de un sistema explicable implica cuestiones relacionadas con responsabilidad institucional, supervisión humana, sesgos, privacidad, trazabilidad, calidad de la información y comunicación de la incertidumbre. La tecnología adquiere sentido cuando se integra dentro de procedimientos que permiten comprender, revisar y cuestionar sus resultados.

La obra se dirige a investigadores, docentes, estudiantes, profesionales de tecnologías de la información, especialistas en contratación pública, responsables institucionales y lectores interesados en la aplicación responsable de la inteligencia artificial. Su contenido puede utilizarse como referencia conceptual, como recurso para la enseñanza y como punto de partida para nuevas investigaciones orientadas a conectar la innovación tecnológica con las exigencias de la gestión pública.

Más que ofrecer una respuesta definitiva, este libro abre un espacio de análisis sobre las condiciones necesarias para desarrollar herramientas explicables, reproducibles y sometidas a control humano. Su aporte se encuentra en demostrar que la innovación no debe medirse únicamente por la complejidad del algoritmo, sino por la capacidad para documentar sus decisiones de diseño, reconocer sus límites y mantener la responsabilidad institucional fuera de la automatización.

Esta perspectiva convierte la obra en una contribución pertinente para el debate contemporáneo sobre inteligencia artificial responsable. Su lectura invita a comprender que explicar un modelo no significa justificarlo automáticamente; que una puntuación no constituye una decisión; y que la eficiencia tecnológica solo resulta legítima cuando se encuentra acompañada por transparencia, gobernanza, revisión crítica y compromiso con el interés público.

RESUMEN

La presente obra analiza la convergencia entre inteligencia artificial explicable, modelos predictivos, sistemas de recomendación y toma de decisiones en el contexto de la contratación pública. Su propósito es fundamentar y documentar el diseño de un prototipo experimental de apoyo analítico capaz de procesar variables asociadas con perfiles sintéticos de proveedores tecnológicos, generar puntuaciones mediante modelos de clasificación y presentar información explicativa que facilite la revisión humana de los resultados. El libro se estructura en cuatro capítulos. El primero examina la contratación pública como un entorno complejo de decisión y desarrolla aspectos relacionados con transparencia, trazabilidad, calidad de la información y riesgos de la opacidad algorítmica. El segundo presenta los fundamentos de la inteligencia artificial, el aprendizaje automático, los modelos predictivos, los sistemas de recomendación y las principales técnicas de explicabilidad. El tercero describe el diseño metodológico, la gobernanza del conjunto sintético, la operacionalización de variables, el procesamiento de datos, la evaluación del modelo y la arquitectura funcional del prototipo. El cuarto discute sus aportes, limitaciones, riesgos y posibilidades de desarrollo futuro. La propuesta utiliza datos generados programáticamente y se plantea como una prueba de concepto académica, no como una plataforma institucional validada. Las puntuaciones obtenidas no representan probabilidades reales de adjudicación ni criterios oficiales para seleccionar proveedores. El principal aporte de la obra consiste en integrar procesamiento predictivo, documentación de datos, explicación global y local, visualización, trazabilidad y supervisión humana dentro de una arquitectura reproducible. Se concluye que la explicabilidad puede favorecer la comprensión y revisión de los modelos, pero no sustituye la calidad de los datos, la validación técnica, el análisis jurídico ni la responsabilidad institucional. La futura aplicación del prototipo requerirá datos reales documentados, evaluación externa, análisis de sesgos, pruebas con usuarios y mecanismos permanentes de auditoría y control.

Palabras clave: inteligencia artificial explicable; modelos predictivos; sistemas de recomendación; contratación pública; gobernanza de datos; supervisión humana.

ÍNDICE GENERAL

ÍNDICE TABLAS

ÍNDICE GRÁFICOS

INTRODUCCIÓN

La transformación digital de las instituciones ha incrementado la disponibilidad de datos, documentos y herramientas capaces de apoyar el análisis de problemas complejos. Dentro de este proceso, la inteligencia artificial ha adquirido una presencia creciente debido a su capacidad para identificar patrones, clasificar observaciones, estimar resultados y organizar grandes volúmenes de información. Estas posibilidades han despertado interés en sectores donde las decisiones dependen de múltiples variables y requieren procesar evidencias procedentes de fuentes diversas.

La contratación pública constituye uno de esos sectores. Los procedimientos mediante los cuales las instituciones adquieren bienes, contratan servicios y ejecutan obras requieren articular condiciones jurídicas, técnicas, económicas, operativas y documentales. La evaluación de alternativas no puede reducirse a la observación aislada del precio, la experiencia o cualquier otra característica. Debe considerar la correspondencia entre la necesidad institucional, las condiciones del procedimiento, las capacidades demostradas por los participantes y las responsabilidades de las personas que intervienen en la decisión.

La digitalización permite almacenar y consultar una cantidad creciente de información sobre proveedores, procedimientos, ofertas y contratos. Sin embargo, la disponibilidad de datos no garantiza automáticamente decisiones más transparentes o adecuadas. La utilidad de la información depende de su integridad, consistencia, actualización, pertinencia y trazabilidad. Del mismo modo, los resultados generados por un modelo dependen de la manera en que el problema fue formulado, de las variables incorporadas, de la calidad de los registros y de los criterios utilizados para evaluar el sistema.

Los modelos predictivos pueden contribuir a organizar información e identificar relaciones que resultarían difíciles de observar mediante procedimientos exclusivamente manuales. Un clasificador puede estimar la pertenencia de un caso a una categoría; un sistema de puntuación puede ordenar alternativas; y un recomendador puede presentar aquellas que considera pertinentes

según determinada función de utilidad. No obstante, ninguna de estas operaciones constituye por sí misma una decisión administrativa.

Esta diferencia representa uno de los principios orientadores de la presente obra. La predicción expresa una estimación condicionada por los datos y el modelo. La puntuación representa el valor asignado a una alternativa dentro de una estructura previamente definida. La recomendación organiza o selecciona opciones para facilitar su revisión. La decisión implica interpretar esa información, contrastarla con otros elementos, aplicar disposiciones jurídicas y técnicas y asumir la responsabilidad correspondiente.

La creciente complejidad de algunos algoritmos ha generado un problema adicional. Un modelo puede alcanzar un desempeño estadístico aparentemente adecuado y, al mismo tiempo, proporcionar poca información sobre las razones que sustentan sus resultados. Esta situación suele describirse mediante la noción de caja negra y representa una dificultad en ámbitos donde las decisiones deben ser justificadas, revisadas y auditadas.

La inteligencia artificial explicable surge como un campo orientado a reducir la distancia entre el funcionamiento técnico de un sistema y la comprensión que los usuarios pueden alcanzar sobre sus resultados. Su finalidad consiste en proporcionar razones, evidencias o representaciones que permitan identificar los factores relacionados con una predicción, examinar su comportamiento y reconocer sus limitaciones.

La explicación no garantiza que un modelo sea correcto, equitativo o jurídicamente procedente. Un sistema puede explicar con claridad una predicción producida a partir de variables inadecuadas o datos deficientes. Por esta razón, la explicabilidad debe acompañarse de procedimientos de validación, documentación, evaluación de riesgos y supervisión humana. Explicar el modelo permite examinarlo, pero no sustituye la revisión de sus fundamentos.

La aplicación de estas tecnologías en contratación pública plantea oportunidades y riesgos. Entre las posibilidades se encuentran la organización de información, la identificación de patrones, la priorización de casos y la presentación de comparaciones. Entre los riesgos se encuentran la reproducción de decisiones históricas, la exclusión de nuevos participantes, la utilización de variables sustitutas, la confianza excesiva en resultados numéricos y la delegación indebida de responsabilidades.

La obra se desarrolla a partir de la necesidad de explorar cómo pueden integrarse modelos predictivos, mecanismos explicativos y criterios de gobernanza dentro de una herramienta experimental de apoyo. Su pregunta orientadora se formula en los siguientes términos:

¿Cómo puede diseñarse un prototipo de apoyo analítico que integre un modelo predictivo y mecanismos de explicabilidad para facilitar la comprensión de estimaciones relacionadas con perfiles de proveedores en escenarios simulados de contratación pública?

Esta pregunta se plantea desde una perspectiva de diseño y no como una demostración definitiva de impacto. La obra no pretende probar que la inteligencia artificial reduce el gasto público, elimina errores o garantiza decisiones equitativas. Estos efectos requerirían datos reales, una aplicación institucional, indicadores definidos y procedimientos de evaluación capaces de comparar resultados antes y después de la implementación.

El propósito general del libro consiste en analizar, diseñar y documentar un prototipo experimental que articule un modelo de clasificación, un sistema de puntuación y mecanismos de explicabilidad para apoyar la comprensión de resultados obtenidos a partir de perfiles sintéticos de proveedores tecnológicos.

Este propósito se desarrolla mediante varias dimensiones complementarias. En primer lugar, se examina la contratación pública como un entorno de decisión que exige transparencia, calidad de la información y trazabilidad. En segundo lugar, se presentan los fundamentos de la inteligencia

artificial, el aprendizaje automático, los sistemas de recomendación y la explicabilidad. En tercer lugar, se documentan el origen de los datos, las variables, las transformaciones, el entrenamiento, la evaluación y la arquitectura funcional del prototipo. En cuarto lugar, se analizan sus aportes, riesgos, limitaciones y condiciones de desarrollo futuro.

La propuesta posee un alcance académico, tecnológico y demostrativo. El prototipo utiliza un conjunto de datos sintético generado mediante programación. Cada registro representa un perfil hipotético y no una empresa, oferta o adjudicación real. Las variables relacionadas con experiencia, proyectos, costos, tiempos, satisfacción, certificaciones, personal, cobertura, soporte, especialización y cumplimiento fueron seleccionadas para construir un escenario tabular que permitiera ejecutar el flujo computacional.

Los valores, rangos y pesos utilizados no constituyen criterios oficiales de contratación. Tampoco proceden de expedientes administrativos o de una evaluación institucional. Su finalidad consiste en hacer explícitas las relaciones incorporadas durante la simulación y permitir que el modelo sea evaluado dentro de un entorno controlado.

Esta delimitación responde a una consideración metodológica fundamental. Un modelo solo puede aprender relaciones presentes en los datos y en la variable objetivo. Cuando la etiqueta se genera sin relación con los predictores, el algoritmo carece de un patrón sustantivo que recuperar. La obra revisa esta situación y propone construir una clase de compatibilidad sintética mediante una regla transparente, reproducible y expresamente identificada como demostrativa.

El modelo principal corresponde a un bosque aleatorio, seleccionado por su capacidad para trabajar con variables tabulares, representar relaciones no lineales y combinar múltiples árboles. Su comportamiento debe compararse con una estrategia ingenua y con una regresión logística, de modo que la complejidad adicional pueda evaluarse frente a alternativas más sencillas.

La evaluación no se limita a la exactitud global. Se consideran la matriz de confusión, la exactitud balanceada, la precisión, la sensibilidad, la medida F1 y otros indicadores que permiten examinar el desempeño frente a las diferentes clases. Esta decisión resulta especialmente importante cuando la distribución de las categorías no es uniforme, porque un porcentaje elevado de aciertos podría ocultar la incapacidad para reconocer la clase menos frecuente.

La explicabilidad se incorpora mediante dos niveles. La explicación global busca identificar las variables de las que depende el comportamiento general del modelo. La explicación local se concentra en una observación específica y muestra cómo sus características contribuyeron a la puntuación obtenida. Para estas finalidades se consideran la importancia por permutación y los valores SHAP.

Estas técnicas describen el funcionamiento del modelo, pero no permiten establecer relaciones causales. Una variable puede influir en una puntuación porque el algoritmo aprendió una asociación dentro del conjunto, sin que ello signifique que modificarla produzca necesariamente un resultado equivalente en la realidad. La explicación algorítmica debe interpretarse como un recurso para la inspección y no como una demostración de causalidad.

La gobernanza de datos constituye otra dimensión central. El conjunto sintético se acompaña de una ficha que documenta su nombre, versión, finalidad, procedencia, composición, usos permitidos, restricciones y limitaciones. También se establece un registro de las transformaciones, los parámetros y las versiones del entorno computacional.

La trazabilidad permite reconstruir cómo se produjo cada resultado. Este proceso incluye identificar la versión del conjunto, las variables procesadas, las transformaciones aplicadas, el modelo utilizado, la puntuación obtenida y la explicación presentada. En una futura implementación, debería incluir además la persona que revisó el resultado y la decisión adoptada.

La obra reconoce que la confianza en un sistema no depende únicamente de su precisión. Requiere documentación, estabilidad, seguridad, posibilidad de revisión y comunicación de incertidumbre. Una interfaz moderna o una representación gráfica no convierten automáticamente una predicción en confiable. La confianza debe ser proporcional a la evidencia disponible y permitir que el usuario cuestione el resultado.

La supervisión humana ocupa, por tanto, una posición transversal. Su finalidad no consiste en confirmar automáticamente la alternativa priorizada por el modelo. Implica revisar los datos, interpretar las explicaciones, identificar información omitida, detectar casos fuera del alcance y rechazar la salida cuando existan razones justificadas.

El libro se organiza en cuatro capítulos. El capítulo I, denominado Contratación pública, transparencia y toma de decisiones, presenta el contexto institucional, la complejidad de los procedimientos, la importancia de las especificaciones técnicas y los riesgos de utilizar modelos opacos. También delimita el problema, el propósito y el alcance de la obra.

El capítulo II, titulado Inteligencia artificial explicable y sistemas de recomendación, desarrolla el marco conceptual. Examina la evolución de la inteligencia artificial, los paradigmas del aprendizaje automático, los modelos de clasificación, las métricas, la interpretabilidad, los métodos de explicación y los principales enfoques de los sistemas de recomendación.

El capítulo III, denominado Diseño metodológico y gobernanza de datos, documenta la naturaleza del prototipo, su estrategia de desarrollo, las condiciones de factibilidad, la procedencia sintética de los datos, la operacionalización de las variables, la construcción de la clase, el procesamiento, el entrenamiento, la evaluación y la arquitectura funcional explicable.

El capítulo IV, titulado Discusión, conclusiones y proyección del prototipo, examina los aportes de la propuesta, sus limitaciones, los riesgos de automatización, la reproducción de patrones

históricos, las variables sustitutas y las explicaciones persuasivas. Asimismo, formula conclusiones ajustadas a la evidencia y presenta una ruta progresiva para la maduración del prototipo.

La obra se dirige a investigadores, docentes, estudiantes, especialistas en tecnologías de la información, profesionales de contratación pública y responsables interesados en la aplicación responsable de la inteligencia artificial. Su lectura no requiere asumir que la innovación tecnológica debe reemplazar el juicio profesional. Por el contrario, invita a analizar cómo los sistemas pueden fortalecer la comprensión, la documentación y la revisión sin desplazar la responsabilidad humana.

El aporte del libro se encuentra en integrar dimensiones que con frecuencia se presentan por separado. El modelo predictivo se relaciona con la calidad de los datos; la explicación se vincula con la fidelidad y la comprensión; la recomendación se diferencia de la decisión; y la innovación tecnológica se conecta con la gobernanza, la ética y la rendición de cuentas.

La propuesta debe entenderse como un punto de partida. Una futura implementación requeriría fuentes reales documentadas, validación de variables con especialistas, comparación de modelos, análisis de sesgos, evaluación con usuarios, revisión jurídica, infraestructura segura y mecanismos permanentes de auditoría. Hasta que estas condiciones se cumplan, el prototipo debe mantenerse dentro de su finalidad académica y demostrativa.

Desde esta perspectiva, la inteligencia artificial explicable no se presenta como una respuesta definitiva, sino como un medio para hacer visibles las relaciones que intervienen en una estimación y favorecer su examen crítico. La innovación adquiere valor cuando permite comprender mejor la información, reconocer la incertidumbre y mantener la decisión bajo responsabilidad de las personas. Ese principio orienta el desarrollo de los capítulos que integran la presente obra.

CAPÍTULO I

CONTRATACIÓN PÚBLICA, TRANSPARENCIA Y TOMA DE DECISIONES

CHAPTER I: PUBLIC PROCUREMENT, TRANSPARENCY, AND DECISION-MAKING

CAPÍTULO I: CONTRATACIÓN PÚBLICA, TRANSPARENCIA Y TOMA DE DECISIONES

El Sistema Nacional de Contratación Pública y el rol del SERCOP

La contratación pública constituye uno de los principales mecanismos mediante los cuales las instituciones del Estado adquieren bienes, ejecutan obras y contratan servicios necesarios para atender las demandas de la sociedad. Su importancia no se limita a la realización de transacciones administrativas, puesto que las decisiones adoptadas dentro de estos procedimientos influyen en el uso de los recursos públicos, la continuidad de los servicios institucionales, la ejecución de políticas públicas y la satisfacción de necesidades colectivas. En consecuencia, la contratación estatal requiere procedimientos que articulen eficiencia, legalidad, transparencia, competencia y responsabilidad en cada una de sus etapas.

En Ecuador, el Sistema Nacional de Contratación Pública integra principios, normas, actores, procedimientos y herramientas orientados a planificar, administrar, controlar y ejecutar las contrataciones realizadas por las entidades sujetas a la legislación correspondiente. Dentro de este sistema, el Servicio Nacional de Contratación Pública ejerce la rectoría institucional y desarrolla funciones de regulación técnica, administración de herramientas, generación de información y acompañamiento a los actores vinculados con los procedimientos de contratación. La normativa vigente caracteriza al SERCOP como una entidad de derecho público, técnico-regulatoria, con personalidad jurídica propia y autonomía administrativa, técnica, operativa, financiera y presupuestaria.

Las atribuciones del SERCOP comprenden, entre otras, asegurar el cumplimiento de los objetivos del Sistema Nacional de Contratación Pública, administrar el Registro Único de Proveedores, desarrollar el Portal de Contratación Pública, generar estadísticas y facilitar el acceso a la información relacionada con los procedimientos de contratación. La legislación vigente dispone, además, que la información del portal sea gestionada bajo el concepto de datos abiertos y contempla

el apoyo de herramientas de inteligencia artificial para facilitar el acceso, la consulta y el análisis de documentos, datos y reportes.

Esta incorporación normativa de los datos abiertos y de herramientas de inteligencia artificial representa una oportunidad para fortalecer el análisis de la información pública, pero también introduce exigencias adicionales. La disponibilidad de grandes volúmenes de datos no garantiza, por sí sola, que las decisiones sean correctas, transparentes o equitativas. Para que la información resulte útil, debe reunir condiciones de integridad, consistencia, pertinencia, actualización y trazabilidad. De igual manera, cualquier modelo algorítmico utilizado para procesarla debe permitir que sus resultados sean examinados críticamente por las personas responsables de tomar decisiones.

La contratación pública constituye un entorno particularmente complejo porque cada procedimiento puede incorporar requisitos jurídicos, condiciones técnicas, restricciones presupuestarias, criterios de calidad, plazos, garantías, experiencia previa y capacidades operativas. Estos elementos no siempre poseen la misma relevancia ni pueden ser evaluados mediante una única medida. En consecuencia, la decisión no consiste simplemente en identificar la oferta de menor precio, sino en analizar la correspondencia entre la necesidad institucional, las condiciones establecidas en los documentos del procedimiento y la capacidad demostrada por los participantes.

En este contexto, las herramientas tecnológicas pueden contribuir a organizar la información, reconocer relaciones entre variables y presentar resultados que apoyen el análisis. Sin embargo, su incorporación debe realizarse con precaución. Cuando un sistema genera una clasificación, una predicción o una recomendación sin mostrar los elementos que influyeron en el resultado, se dificulta la revisión de sus criterios, la identificación de errores y la justificación de la decisión. Por esta razón, la transparencia tecnológica no debe entenderse únicamente como acceso a una plataforma o publicación de información, sino también como la posibilidad de comprender, evaluar y cuestionar la manera en que los datos son procesados.

Desde esta perspectiva, la inteligencia artificial explicable adquiere relevancia como un enfoque orientado a proporcionar información comprensible sobre los factores que intervienen en una predicción. Su aplicación en contratación pública no debe concebirse como un mecanismo de adjudicación automática, sino como una alternativa de apoyo para fortalecer la trazabilidad del análisis y facilitar la supervisión humana. El valor de una herramienta de esta naturaleza dependerá, en última instancia, de la calidad de los datos, de la pertinencia del modelo, de la claridad de sus explicaciones y de su integración responsable dentro de los procedimientos institucionales.

La contratación pública como entorno complejo de toma de decisiones

La contratación pública constituye un proceso de decisión complejo porque en ella convergen necesidades institucionales, disposiciones jurídicas, requerimientos técnicos, restricciones presupuestarias y condiciones operativas. Cada procedimiento requiere que la entidad contratante defina con claridad el objeto de la contratación, determine las características del bien o servicio requerido y establezca criterios que permitan verificar si las ofertas recibidas responden efectivamente a la necesidad identificada. Por consiguiente, la decisión no se reduce a escoger entre varias propuestas, sino que supone analizar evidencias diversas, comparar alternativas y justificar la opción adoptada.

La complejidad aumenta cuando los criterios de evaluación poseen naturalezas diferentes. Algunos pueden expresarse mediante valores cuantitativos, como el precio, el tiempo de entrega, los años de experiencia o la capacidad operativa. Otros requieren apreciaciones cualitativas, como la pertinencia de una metodología, la correspondencia de una solución con los requerimientos institucionales o la calidad de los mecanismos de soporte ofrecidos. Esta diversidad obliga a establecer relaciones entre variables que no siempre pueden compararse directamente y cuya importancia puede variar de acuerdo con las características de cada contratación.

En determinadas circunstancias, una oferta puede presentar un precio reducido, pero requerir mayores tiempos de ejecución o contar con una capacidad técnica limitada. Otra puede ofrecer

mejores condiciones operativas, aunque suponga una inversión económica superior. También pueden existir propuestas que cumplen formalmente los requisitos documentales, pero presentan dificultades para demostrar la experiencia o los recursos necesarios para ejecutar el contrato. Estas posibles combinaciones evidencian que la toma de decisiones debe apoyarse en criterios previamente definidos y no únicamente en la observación aislada de una variable.

La objetividad en este campo no implica eliminar toda intervención humana. Una decisión objetiva requiere que los criterios sean conocidos, que las evidencias puedan verificarse, que las reglas se apliquen de forma consistente y que las razones de la elección queden registradas. El juicio profesional continúa siendo necesario, pero debe encontrarse vinculado con información comprobable y con procedimientos que permitan revisar el proceso seguido.

Desde esta perspectiva, la trazabilidad se convierte en una condición esencial. No basta con conocer el resultado final de una evaluación; también debe ser posible reconstruir qué información se utilizó, qué criterios se aplicaron, qué comparaciones se realizaron y qué personas intervinieron. La ausencia de esta trazabilidad puede dificultar la revisión posterior de la decisión, especialmente cuando participan numerosos proveedores, documentos y variables.

Las plataformas digitales permiten almacenar, consultar y organizar grandes volúmenes de información. No obstante, la digitalización no elimina automáticamente las dificultades del proceso. Un sistema puede acelerar la consulta de documentos y facilitar determinadas comparaciones, pero sus resultados seguirán dependiendo de la calidad de la información registrada, de la correcta definición de los criterios y de la forma en que estos hayan sido traducidos a reglas o variables procesables.

Esta distinción es relevante porque una herramienta tecnológica no interpreta por sí misma la necesidad pública que originó una contratación. Tampoco determina de manera autónoma si una condición es jurídicamente procedente o si un criterio técnico refleja adecuadamente la realidad institucional. El sistema procesa la información que recibe conforme a una estructura previamente

diseñada. Por ello, cualquier error en la definición del problema, en la selección de las variables o en la preparación de los datos puede trasladarse al resultado generado.

En este contexto, los modelos predictivos y los sistemas de recomendación pueden contribuir a identificar patrones, organizar alternativas y mostrar relaciones que resultarían difíciles de observar mediante una revisión exclusivamente manual. Sin embargo, su utilización debe mantenerse dentro de límites claramente establecidos. Una predicción expresa una estimación construida a partir de determinadas variables, mientras que una recomendación organiza o prioriza opciones según criterios definidos. Ninguna de las dos constituye, por sí sola, una decisión administrativa.

La diferencia entre predicción, recomendación y decisión debe mantenerse en toda la obra. El modelo desarrollado no debe presentarse como una tecnología capaz de adjudicar contratos ni como un mecanismo que reemplaza a los responsables institucionales. Su función debe describirse como apoyo analítico: procesa información, genera estimaciones y ofrece elementos explicativos que pueden ser examinados por las personas encargadas de valorar las propuestas.

Esta delimitación permite conservar el propósito innovador del proyecto sin atribuirle capacidades que no han sido demostradas. La utilidad de la herramienta no dependerá de sustituir la intervención humana, sino de contribuir a que el análisis sea más organizado, comprensible y susceptible de revisión.

Calidad de la información y especificaciones técnicas

La calidad de la información constituye uno de los principales condicionantes de cualquier proceso de contratación y de todo modelo que pretenda apoyar la toma de decisiones. Las plataformas, algoritmos y herramientas predictivas requieren datos suficientemente claros, consistentes y actualizados. Cuando la información contiene errores, contradicciones o vacíos, el sistema puede

generar resultados formalmente correctos desde el punto de vista computacional, pero poco confiables para el propósito institucional.

En los procedimientos de contratación pública, las especificaciones técnicas cumplen una función central porque expresan las características que debe reunir el bien, servicio u obra requerida. Su adecuada formulación permite que los potenciales proveedores comprendan la necesidad institucional y preparen ofertas comparables. En cambio, una especificación ambigua, contradictoria o excesivamente general puede generar interpretaciones distintas y dificultar la evaluación posterior.

Los problemas en las especificaciones técnicas pueden manifestarse de diversas formas. Pueden presentarse unidades de medida inconsistentes, requisitos que no guardan relación directa con el objeto contractual, condiciones imposibles de verificar, rangos que se contradicen entre distintas secciones o descripciones que no permiten determinar con precisión el nivel de cumplimiento esperado. También pueden surgir dificultades cuando la información se encuentra dispersa entre varios documentos o cuando se utilizan términos sin una definición operacional.

Estas deficiencias no deben interpretarse únicamente como errores de redacción. Una especificación imprecisa puede alterar la comparación entre ofertas, afectar la comprensión de los participantes y trasladar incertidumbre a las fases posteriores del procedimiento. Si cada proveedor interpreta de manera diferente una condición, las propuestas resultantes pueden no responder a un mismo parámetro, lo que limita la posibilidad de compararlas de forma consistente.

La preparación de la información también resulta relevante para el funcionamiento de los modelos predictivos. Cada variable incorporada debe tener una definición clara. Debe conocerse qué representa, cómo fue medida, cuál es su unidad, de qué fuente proviene, cuándo fue actualizada y qué tratamiento recibió antes de ingresar al modelo. Sin esta documentación, una variable aparentemente sencilla puede adquirir significados diferentes según el registro o la persona que la utilice.

Por ejemplo, la experiencia de un proveedor puede expresarse mediante años de actividad, cantidad de contratos ejecutados, valor acumulado de contrataciones, número de proyectos similares o resultados obtenidos. Estas mediciones no son equivalentes. Seleccionar una de ellas sin explicar su relación con el objetivo del análisis puede conducir a interpretaciones erróneas.

Una situación semejante ocurre con variables como el costo, la capacidad operativa, el tiempo de respuesta, las certificaciones o el nivel de cumplimiento. El modelo necesita conocer la dirección del criterio. En algunas variables, un valor elevado puede considerarse favorable; en otras, como el costo o el tiempo de respuesta, un valor menor podría resultar preferible. No establecer esta relación puede producir comparaciones inadecuadas.

La calidad de los datos comprende, al menos, su integridad, consistencia, pertinencia y trazabilidad. La integridad se relaciona con la disponibilidad de los campos necesarios. La consistencia exige que los registros mantengan definiciones y formatos uniformes. La pertinencia implica que cada variable guarde una relación justificable con el problema analizado. La trazabilidad permite identificar el origen de la información y reconstruir las transformaciones realizadas.

Cuando existen valores faltantes, duplicados o atípicos, estos no deben corregirse de manera automática sin dejar constancia del procedimiento aplicado. Cada tratamiento puede modificar la información y, en consecuencia, influir en las predicciones. La eliminación de registros, la sustitución de valores y la normalización de escalas deben encontrarse documentadas para que los resultados puedan reproducirse y evaluarse.

La inteligencia artificial explicable puede contribuir a mostrar qué variables influyeron en una estimación, pero no puede garantizar por sí misma que esas variables sean correctas. Un modelo puede explicar con claridad una predicción basada en datos deficientes. La explicación permitiría conocer el razonamiento matemático seguido, pero no corregiría los problemas presentes en la información de origen.

Por este motivo, la calidad de los datos debe analizarse antes de valorar el desempeño del modelo. El desarrollo de una herramienta explicable requiere combinar dos responsabilidades: asegurar que la información utilizada reúna condiciones mínimas de confiabilidad y proporcionar mecanismos que permitan comprender cómo fue procesada. La explicabilidad no sustituye la validación de los datos, sino que la complementa.

En la presente obra, esta consideración obliga a diferenciar entre las variables utilizadas con fines demostrativos y aquellas que podrían emplearse en una implementación institucional. El prototipo puede mostrar el funcionamiento general de una herramienta, pero la selección definitiva de los datos requeriría un análisis jurídico, técnico y metodológico específico para cada tipo de procedimiento.

Transparencia, trazabilidad y confianza en la decisión

La transparencia en contratación pública posee varias dimensiones. Una primera dimensión se relaciona con el acceso a la información institucional, los documentos del procedimiento y los resultados de la contratación. Una segunda corresponde a la claridad de las reglas aplicadas para evaluar las propuestas. Una tercera se vincula con la posibilidad de reconstruir las actuaciones realizadas. Cuando intervienen modelos algorítmicos, se incorpora una dimensión adicional: la transparencia sobre el modo en que los datos fueron procesados y las razones que sustentan una predicción o recomendación.

La publicación de información representa una condición necesaria, pero no suficiente. Un procedimiento puede contener numerosos documentos disponibles públicamente y, aun así, resultar difícil de comprender si los criterios se encuentran dispersos, si las variables no poseen definiciones claras o si la justificación de la decisión no permite relacionar el resultado con las evidencias evaluadas.

En el ámbito tecnológico, la transparencia no significa que todos los usuarios deban conocer cada operación matemática ejecutada por el algoritmo. Significa que deben recibir información adecuada para comprender qué factores influyeron en el resultado, qué limitaciones posee el modelo y qué nivel de confianza puede atribuirse a la estimación. La explicación debe adaptarse a las necesidades de quienes utilizarán el sistema.

Un equipo técnico puede requerir información detallada sobre variables, parámetros, datos de entrenamiento y métricas de evaluación. Un responsable administrativo puede necesitar conocer los criterios que favorecieron o redujeron una puntuación. Un proveedor podría requerir una explicación comprensible sobre los elementos considerados, sin que ello implique revelar información reservada o afectar los derechos de terceros. Por esta razón, no existe una única explicación válida para todos los destinatarios.

La trazabilidad algorítmica debe permitir registrar, al menos, la versión del modelo utilizada, los datos ingresados, las transformaciones realizadas, la predicción generada, la explicación asociada y la persona que revisó el resultado. Este registro resulta indispensable cuando un modelo se actualiza o cuando diferentes versiones pueden producir resultados distintos.

La confianza tampoco debe asumirse como una consecuencia automática del uso de tecnología. Un sistema no se vuelve confiable por presentar una interfaz moderna o generar resultados numéricos. La confianza debe construirse mediante evidencias de funcionamiento, documentación, validación, posibilidad de revisión y reconocimiento explícito de las limitaciones.

La exactitud global de un modelo, considerada de manera aislada, tampoco resulta suficiente para demostrar su utilidad. Un sistema puede alcanzar un porcentaje aparentemente elevado y, al mismo tiempo, presentar dificultades para identificar determinadas categorías. También puede producir resultados estadísticamente aceptables, pero basados en variables que no deberían influir en la decisión.

En este sentido, la transparencia debe vincularse con la rendición de cuentas. Si una herramienta interviene en el análisis, debe quedar claramente establecido quién diseñó el modelo, quién autorizó su uso, quién revisó la información, quién interpretó el resultado y quién adoptó la decisión final. La presencia de un algoritmo no elimina la responsabilidad institucional.

La explicabilidad puede fortalecer la confianza cuando permite detectar errores, identificar variables influyentes y cuestionar resultados inesperados. No obstante, una explicación no convierte automáticamente una predicción en correcta o justa. Un modelo puede explicar de manera clara una relación construida sobre datos sesgados o criterios inadecuados. Por esta razón, la explicabilidad debe acompañarse de evaluaciones de calidad, consistencia y posibles efectos diferenciales.

También debe evitarse que la explicación se convierta en una justificación aparente. Una representación gráfica o una lista de variables puede dar la impresión de transparencia sin explicar realmente cómo se relacionan esos elementos con la predicción. Las explicaciones deben ser comprensibles, pero también fieles al funcionamiento del modelo.

La confianza en el sistema debe mantenerse dentro de límites razonables. El usuario necesita comprender que la herramienta puede cometer errores, que sus resultados dependen de la información disponible y que una predicción no equivale a una certeza. Esta comunicación de la incertidumbre es especialmente importante en contextos donde las decisiones pueden producir consecuencias económicas, jurídicas e institucionales.

La supervisión humana constituye, por tanto, una condición transversal. Su finalidad no consiste únicamente en confirmar el resultado generado por el sistema, sino en cuestionarlo, compararlo con otras evidencias y rechazarlo cuando existan razones justificadas. Un modelo explicable debe favorecer este ejercicio crítico, no reducirlo.

Riesgos asociados con el uso de modelos algorítmicos opacos

La incorporación de modelos algorítmicos en los procesos de análisis puede facilitar el tratamiento de grandes volúmenes de información, identificar relaciones entre variables y generar estimaciones que sirvan como apoyo para comparar diferentes alternativas. Sin embargo, estas capacidades no eliminan los riesgos asociados con la calidad de los datos, el diseño del modelo y la interpretación de los resultados. Cuando el funcionamiento de una herramienta no puede ser comprendido o reconstruido por sus usuarios, se produce una situación de opacidad que limita la posibilidad de examinar críticamente las recomendaciones generadas.

La opacidad algorítmica no depende únicamente de la complejidad matemática del modelo. También puede originarse por una documentación insuficiente, por la ausencia de definiciones claras sobre las variables utilizadas, por el desconocimiento de los criterios empleados durante el entrenamiento o por la presentación de resultados sin una explicación sobre los factores que influyeron en ellos. Incluso un modelo técnicamente sencillo puede resultar opaco cuando el usuario recibe una puntuación o clasificación sin conocer el procedimiento que permitió obtenerla.

En contratación pública, esta limitación adquiere especial relevancia porque las decisiones deben estar sustentadas en criterios verificables y vinculadas con las condiciones establecidas en cada procedimiento. Si una herramienta asigna una puntuación favorable a un proveedor, pero no permite identificar qué características determinaron ese resultado, la recomendación difícilmente podrá ser revisada, contrastada o incorporada de forma responsable al análisis institucional. La generación automática de un resultado no sustituye la obligación de justificar las decisiones.

Uno de los riesgos se encuentra en los datos utilizados para construir el modelo. Los registros históricos reflejan hechos ocurridos en contextos específicos, bajo determinadas reglas y con niveles variables de calidad documental. Por este motivo, no deben interpretarse automáticamente como una representación perfecta de las decisiones que deberían adoptarse en el futuro. Un modelo puede aprender regularidades presentes en la información histórica sin distinguir si esas

regularidades responden a criterios técnicamente pertinentes, a condiciones coyunturales o a limitaciones de los datos disponibles.

La utilización de antecedentes de adjudicación también puede favorecer la reproducción de patrones previos. Si el modelo interpreta que la adjudicación pasada constituye el principal indicador de idoneidad futura, los proveedores con mayor participación histórica podrían recibir sistemáticamente puntuaciones superiores. Esta situación podría reducir la posibilidad de considerar alternativas nuevas o empresas que, pese a no contar con un amplio historial contractual, poseen capacidades suficientes para cumplir las condiciones de un procedimiento.

Algo semejante puede ocurrir con variables que funcionan como aproximaciones indirectas de otras características. El tamaño de una empresa, la cantidad de contratos anteriores, su ubicación, la disponibilidad de determinados certificados o el número de empleados pueden aportar información relevante en ciertos contextos, pero no deben asumir automáticamente el valor de indicadores universales de calidad. Su inclusión necesita justificarse de acuerdo con el objeto contractual y con la finalidad específica del modelo.

Otro riesgo se relaciona con la asignación de pesos o niveles de importancia. Todo modelo incorpora decisiones previas sobre las variables que serán utilizadas, la forma en que serán representadas y el resultado que se desea estimar. Estas decisiones no son neutrales. La selección de un indicador, su transformación y su peso dentro del análisis pueden modificar la posición relativa de los proveedores. Por consiguiente, la aparente objetividad de una puntuación numérica no elimina la necesidad de examinar los supuestos que permitieron construirla.

También debe considerarse el riesgo de automatización excesiva. La presentación de porcentajes, rankings o probabilidades puede generar en los usuarios la impresión de que el resultado posee un grado de certeza superior al que realmente ofrece. Esta tendencia puede provocar que la recomendación algorítmica sea aceptada sin una revisión suficiente, especialmente cuando se percibe a la tecnología como más objetiva que el juicio humano. Sin embargo, los modelos también

pueden cometer errores, depender de información incompleta o producir estimaciones inadecuadas para casos distintos de aquellos considerados durante su diseño.

La supervisión humana no debe limitarse a confirmar la alternativa sugerida por el sistema. Debe incluir la posibilidad efectiva de cuestionar la recomendación, solicitar información adicional, revisar los datos de entrada y rechazar el resultado cuando existan argumentos técnicos, jurídicos o administrativos que lo justifiquen. Una herramienta de apoyo conserva su utilidad únicamente cuando fortalece la capacidad de análisis de los responsables institucionales y no cuando reemplaza su criterio.

La falta de actualización constituye otro factor de riesgo. Las condiciones de contratación, las necesidades institucionales, la disponibilidad de proveedores y las características del mercado pueden cambiar. Un modelo construido con información correspondiente a un período determinado puede perder capacidad para representar situaciones posteriores. Por ello, la revisión periódica de los datos, las variables y los resultados debe formar parte del ciclo de vida de la herramienta.

Asimismo, deben registrarse las versiones del modelo y las modificaciones realizadas. Una actualización en las variables, en los parámetros o en el procedimiento de procesamiento puede producir resultados diferentes ante una misma consulta. La trazabilidad requiere que sea posible identificar qué versión se utilizó, qué datos fueron ingresados y qué explicación acompañó a cada estimación.

Estos riesgos muestran que la incorporación de inteligencia artificial en contratación pública no puede abordarse únicamente como un problema de programación. También involucra decisiones sobre gobernanza de datos, documentación, responsabilidad, supervisión, transparencia y uso institucional. El desarrollo tecnológico debe integrarse dentro de un marco que permita reconocer sus posibilidades sin ocultar sus limitaciones.

Tabla 1

Dimensiones problemáticas asociadas con la toma de decisiones y el uso de modelos algorítmicos en contratación pública

Dimensión problemática	*Manifestación*	*Riesgo para la toma de decisiones*	*Orientación de respuesta*
Complejidad de los procedimientos	Concurrencia de criterios jurídicos, técnicos, económicos y documentales	Evaluaciones inconsistentes o insuficientemente justificadas	Organizar los criterios y conservar la trazabilidad del análisis
Calidad de la información	Datos incompletos, inconsistentes, ambiguos o poco estandarizados	Predicciones y comparaciones basadas en información poco confiable	Validar, depurar y documentar los datos antes de procesarlos
Errores en las especificaciones técnicas	Requisitos imprecisos, contradictorios o insuficientemente definidos	Dificultades para comparar ofertas y verificar su correspondencia con la necesidad institucional	Fortalecer la revisión técnica y la consistencia documental
Opacidad algorítmica	Resultados generados sin una explicación comprensible de los factores considerados	Dificultad para revisar, justificar o cuestionar una recomendación	Incorporar mecanismos de explicabilidad y registro de decisiones
Sesgos en los datos o en el modelo	Reproducción de patrones históricos o ponderaciones inadecuadas	Posibles efectos desiguales o recomendaciones injustificadas	Evaluar sesgos, documentar limitaciones y mantener supervisión humana
Insuficiente preparación de los usuarios	Uso de resultados algorítmicos sin comprensión de su alcance	Dependencia excesiva del sistema o interpretación incorrecta de las predicciones	Capacitar a los usuarios y establecer protocolos de uso responsable

Nota. *Elaboración propia de los autores. La tabla expone las causas y consecuencias de los problemas en la contratación pública del SERCOP y el impacto de la XAI. Los datos provienen de los resultados obtenidos en la presente investigación.*

La inteligencia artificial explicable como herramienta de apoyo

La inteligencia artificial explicable se plantea como un enfoque orientado a reducir la distancia existente entre el funcionamiento técnico de un modelo y la comprensión que los usuarios pueden alcanzar sobre sus resultados. Su finalidad no consiste únicamente en mostrar una predicción, sino en proporcionar información que permita interpretar los factores que influyeron en ella, reconocer sus limitaciones y valorar si resulta pertinente para el contexto en el que será utilizada.

La explicabilidad no debe confundirse con una descripción general del algoritmo. Saber que un sistema emplea un árbol de decisión, una regresión logística o un bosque aleatorio no permite, por sí solo, comprender por qué se produjo un resultado específico. Una explicación útil debe relacionar la salida del modelo con las características concretas de la información analizada.

En términos generales, la explicación puede abordarse desde dos niveles complementarios. El primero corresponde a la comprensión global del modelo y busca identificar cómo funciona el sistema en conjunto, qué variables suele considerar más relevantes y qué relaciones generales utiliza para producir sus estimaciones. El segundo corresponde a la explicación local y se concentra en una predicción particular, mostrando qué factores influyeron en el resultado obtenido para un proveedor, una propuesta o un caso específico.

Esta distinción resulta importante porque un modelo puede presentar un comportamiento general razonable y, aun así, generar una predicción particular que requiera revisión. De igual manera, una explicación local puede ayudar a comprender un caso concreto, pero no necesariamente permite evaluar el comportamiento completo del sistema. Por ello, una propuesta explicable debe combinar mecanismos que permitan analizar tanto el modelo general como las estimaciones individuales.

La explicación también debe adaptarse a los destinatarios. Los desarrolladores pueden requerir detalles sobre las variables, los parámetros, las transformaciones y las métricas de evaluación. Los responsables administrativos necesitan información comprensible acerca de los factores que

favorecieron o redujeron una estimación. Los especialistas jurídicos pueden requerir elementos que les permitan determinar si los criterios utilizados guardan relación con las reglas del procedimiento. Los proveedores, por su parte, podrían necesitar conocer de manera clara los elementos generales considerados, sin que esto implique revelar información protegida o perteneciente a otros participantes.

Una explicación orientada al usuario debería mostrar, como mínimo, el resultado generado, los principales factores que influyeron en él, la dirección de esa influencia y las limitaciones que deben considerarse. También debería advertir cuando la información disponible resulte insuficiente o cuando la consulta se aleje de las condiciones para las que fue diseñado el modelo.

La presentación visual puede contribuir a esta finalidad mediante gráficos, comparaciones y descripciones textuales. No obstante, la calidad de una explicación no depende únicamente de su apariencia. Un gráfico puede resultar atractivo y, al mismo tiempo, representar variables incompatibles o inducir a interpretaciones erróneas. Las visualizaciones deben respetar las escalas de las variables, diferenciar los criterios favorables y desfavorables y evitar que un valor elevado sea interpretado automáticamente como un mejor resultado.

Por ejemplo, una mayor experiencia o un mayor porcentaje de cumplimiento podrían considerarse favorables en determinadas circunstancias. En cambio, un costo más elevado o un mayor tiempo de respuesta no deberían interpretarse de la misma manera. El sistema explicativo debe identificar la dirección de cada criterio y mostrarla de forma comprensible. De lo contrario, la comparación visual podría conducir a conclusiones equivocadas.

La inteligencia artificial explicable tampoco garantiza que el modelo sea exacto, equitativo o jurídicamente adecuado. Un sistema puede ofrecer una explicación clara sobre una predicción incorrecta. También puede describir con fidelidad un resultado generado a partir de datos incompletos o criterios insuficientemente justificados. La explicabilidad permite examinar el funcionamiento del modelo, pero no reemplaza la validación de los datos, la evaluación técnica ni

la supervisión institucional.

En la propuesta desarrollada en esta obra, la explicabilidad debe concebirse como una capa de apoyo vinculada con el modelo predictivo y con la interfaz de usuario. Esta capa debería permitir que la estimación no aparezca como un resultado aislado, sino acompañada de información sobre las características que incidieron en ella. Su propósito sería facilitar la revisión humana, detectar resultados inesperados y mejorar la trazabilidad del análisis.

El enfoque adoptado no busca automatizar la adjudicación de contratos. Se orienta a explorar cómo una herramienta tecnológica puede organizar determinadas variables, producir una estimación experimental y presentar elementos explicativos que ayuden a comprenderla. La decisión definitiva continuaría bajo responsabilidad de los funcionarios y especialistas competentes.

Formulación y delimitación del problema

El problema central abordado por esta obra surge de la convergencia entre dos condiciones. Por una parte, los procedimientos de contratación pública requieren analizar información técnica, económica, operativa y documental para comparar proveedores y propuestas. Por otra, los modelos predictivos capaces de procesar esa información pueden generar resultados difíciles de interpretar cuando no incorporan mecanismos suficientes de transparencia.

Esta situación produce una tensión entre capacidad de procesamiento y comprensibilidad. Cuanto mayor es la cantidad de variables y registros analizados, más difícil puede resultar reconstruir manualmente las relaciones que sustentan una predicción. Sin embargo, la eficiencia computacional no puede considerarse suficiente cuando los resultados intervienen en un entorno que exige justificación, revisión y responsabilidad institucional.

El problema no consiste únicamente en que ciertos algoritmos sean complejos. También se relaciona con la ausencia de mecanismos que vinculen la predicción con los datos del caso analizado. Un sistema puede generar una probabilidad o una clasificación sin explicar por qué determinadas

características fueron consideradas favorables, desfavorables o irrelevantes. Esta ausencia limita la capacidad de los usuarios para evaluar la pertinencia de la recomendación.

A partir de esta problemática, la pregunta orientadora de la obra se formula de la siguiente manera:

¿Cómo puede diseñarse un prototipo de apoyo analítico que integre un modelo predictivo y mecanismos de explicabilidad para facilitar la comprensión de estimaciones relacionadas con proveedores y propuestas en escenarios de contratación pública?

La pregunta se formula en términos de diseño y exploración, no de demostración definitiva de impacto. La obra no cuenta actualmente con evidencia suficiente para afirmar que el prototipo reduce costos públicos, elimina errores, garantiza la equidad o mejora de forma comprobada las adjudicaciones. Tales efectos requerirían una aplicación con datos reales, un diseño de evaluación específico y la comparación de resultados antes y después de la implementación.

El ámbito demostrativo del prototipo se concentra en proveedores de servicios de tecnologías de la información. El manuscrito utiliza variables relacionadas con experiencia, proyectos completados, tiempo de respuesta, costo por hora, satisfacción del cliente, certificaciones, personal calificado, cobertura, soporte, especialización en el sector público y cumplimiento de niveles de servicio. Estas variables deben entenderse como componentes de un escenario experimental y no como criterios oficiales de adjudicación. El propio manuscrito presenta el prototipo mediante datos simulados de proveedores y propuestas de servicios tecnológicos.

La utilización de datos simulados permite observar el funcionamiento técnico general del código, probar la estructura del procesamiento y construir ejemplos de visualización. Sin embargo, no permite inferir el comportamiento real de los proveedores ni estimar con validez la probabilidad efectiva de que una empresa obtenga una adjudicación. Por esta razón, los resultados deben presentarse como demostraciones metodológicas.

El prototipo tampoco debe emplear nombres de empresas reales dentro de los datos simulados. La asociación de una empresa identificable con una puntuación, una probabilidad o una recomendación ficticia podría producir interpretaciones incorrectas. En la versión final, los proveedores deben denominarse de manera genérica, por ejemplo, Proveedor A, Proveedor B, Proveedor C y Proveedor D.

La delimitación de la obra comprende las siguientes condiciones:

- El prototipo tiene una finalidad académica y demostrativa.
- Los datos empleados en su ejecución son simulados.
- El ámbito de ejemplo corresponde a proveedores de servicios tecnológicos.
- Las variables utilizadas no constituyen criterios oficiales de adjudicación.
- Las predicciones no representan probabilidades reales de contratación.
- La herramienta no sustituye la evaluación técnica, jurídica o económica.
- El modelo no adjudica contratos ni emite decisiones administrativas.
- La utilidad del sistema se analiza en función de su diseño, procesamiento y capacidad potencial de explicación.

Estas delimitaciones no reducen el valor de la propuesta. Por el contrario, permiten evaluar el prototipo dentro de sus verdaderas posibilidades y diferenciar entre una demostración tecnológica y una implementación institucional validada.

Propósito y alcance de la obra

El propósito general de esta obra es presentar el diseño y funcionamiento de un prototipo que articula un modelo predictivo con elementos de explicabilidad para apoyar el análisis de información relacionada con proveedores y propuestas en escenarios simulados de contratación pública.

La propuesta busca mostrar cómo determinadas variables pueden ser organizadas, procesadas y representadas mediante una interfaz que facilite la interpretación de los resultados. El interés principal no se encuentra únicamente en producir una estimación, sino en explorar la forma en que esa estimación puede acompañarse de información comprensible para los usuarios.

A partir de este propósito general, la obra se orienta a describir el contexto institucional y tecnológico de la contratación pública; examinar los fundamentos de la inteligencia artificial explicable y de los sistemas de recomendación; presentar la estructura metodológica utilizada para desarrollar el prototipo; exponer sus componentes técnicos; y analizar sus posibilidades, riesgos y limitaciones.

El alcance del libro es tecnológico y académico. No se presenta una solución terminada para su implementación directa dentro del Sistema Nacional de Contratación Pública. Tampoco se propone modificar los procedimientos jurídicos vigentes ni establecer nuevos criterios de adjudicación. La obra desarrolla una aproximación experimental que puede servir como punto de partida para investigaciones posteriores.

Una implementación institucional requeriría, entre otros elementos, acceder a datos reales debidamente documentados, definir el propósito exacto de la herramienta, validar las variables con especialistas, establecer mecanismos de seguridad y protección de datos, comprobar el rendimiento del modelo y evaluar los posibles efectos sobre los distintos participantes. También sería necesario definir responsabilidades, procedimientos de auditoría, mecanismos de impugnación y formas de supervisión humana.

El libro reconoce que la innovación tecnológica debe mantenerse subordinada a los principios que orientan la gestión pública. La eficiencia no puede desvincularse de la transparencia; la automatización no puede sustituir la responsabilidad; y la capacidad predictiva no puede considerarse superior a la necesidad de justificar las decisiones.

En este sentido, el aporte de la obra se sitúa en la exploración de una arquitectura que combina procesamiento predictivo, presentación de resultados y explicación para el usuario. Su valor dependerá de la claridad con la que se documenten los procedimientos, de la honestidad con la que se expongan las limitaciones y de la posibilidad de reproducir y examinar críticamente el prototipo.

Los capítulos posteriores desarrollan los fundamentos conceptuales de la inteligencia artificial, el aprendizaje automático, la explicabilidad y los sistemas de recomendación. A continuación, se presenta la metodología empleada para la construcción del prototipo, sus componentes tecnológicos y las etapas de desarrollo. Finalmente, se analizan los resultados alcanzados dentro del escenario simulado y se establecen las condiciones que deberían considerarse en futuras investigaciones o implementaciones.

La incorporación de modelos predictivos en contratación pública representa una posibilidad de innovación que debe ser examinada desde una perspectiva técnica, institucional y ética. La capacidad para procesar información y generar estimaciones solo adquiere valor cuando los resultados pueden ser comprendidos, revisados y utilizados dentro de límites claramente definidos. Por ello, la propuesta desarrollada en esta obra parte de la necesidad de articular calidad de los datos, trazabilidad, explicabilidad y supervisión humana. Sobre esta base, el capítulo siguiente presenta los fundamentos conceptuales necesarios para comprender la estructura del prototipo y las decisiones metodológicas asociadas con su construcción.

CAPÍTULO II: INTELIGENCIA ARTIFICIAL EXPLICABLE Y SISTEMAS DE RECOMENDACIÓN

CHAPTER II: EXPLAINABLE ARTIFICIAL INTELLIGENCE AND RECOMMENDER SYSTEMS

CAPÍTULO II: INTELIGENCIA ARTIFICIAL EXPLICABLE Y SISTEMAS DE RECOMENDACIÓN

El desarrollo de sistemas computacionales destinados a apoyar la toma de decisiones exige integrar conocimientos procedentes de la inteligencia artificial, el aprendizaje automático, la analítica predictiva, los sistemas de recomendación y la interacción entre personas y tecnologías. Cada uno de estos campos responde a problemas específicos, pero su convergencia permite construir herramientas capaces de procesar información, identificar patrones, estimar resultados y presentar alternativas que pueden ser examinadas por los usuarios.

La inteligencia artificial comprende un campo amplio de investigación orientado a diseñar sistemas capaces de ejecutar actividades asociadas con el aprendizaje, el razonamiento, la resolución de problemas y la selección de acciones. La propuesta presentada por McCarthy, Minsky, Rochester y Shannon en 1955 constituyó uno de los antecedentes fundacionales del campo, al plantear la posibilidad de describir con suficiente precisión determinados aspectos del aprendizaje y la inteligencia para que pudieran ser simulados mediante máquinas. Esta formulación inicial no estableció una tecnología única, sino un programa interdisciplinario de investigación que posteriormente dio lugar a múltiples métodos computacionales (McCarthy et al., 1955).

Dentro de este campo, el aprendizaje automático se ocupa del desarrollo de modelos que utilizan datos para reconocer regularidades y generar predicciones sin depender exclusivamente de reglas programadas para cada situación particular. Sus aplicaciones abarcan problemas de clasificación, regresión, agrupamiento, detección de anomalías y priorización de alternativas. El modelo predictivo no reproduce de manera literal la realidad, sino que construye una representación matemática basada en las variables, ejemplos y criterios utilizados durante su entrenamiento.

Esta distinción resulta relevante para la presente obra, porque el prototipo no debe interpretarse como una reproducción integral de los procedimientos de contratación pública. Su función consiste en procesar un conjunto delimitado de variables, generar una estimación y presentar

información que facilite su análisis. El resultado producido por el modelo depende de la calidad de los datos, de la variable objetivo, de los supuestos incorporados y de las condiciones bajo las cuales fue desarrollado.

Los sistemas de recomendación, por su parte, surgieron como mecanismos destinados a reducir la sobrecarga de información y facilitar la selección entre múltiples alternativas. Estos sistemas estiman la utilidad o pertinencia de determinados objetos para un usuario y pueden utilizar información sobre preferencias, características de los elementos, interacciones anteriores o condiciones contextuales. Su propósito fundamental consiste en ordenar, filtrar o priorizar opciones, pero la recomendación producida no equivale necesariamente a una decisión definitiva (Adomavicius & Tuzhilin, 2005).

La diferencia entre predicción y recomendación debe mantenerse clara. Un modelo predictivo puede estimar la probabilidad de que ocurra un resultado, mientras que un sistema de recomendación transforma diferentes fuentes de información en una lista, puntuación o conjunto de opciones consideradas pertinentes para un usuario. La decisión implica una etapa posterior en la que una persona o una institución valora esa información junto con criterios jurídicos, técnicos, económicos y contextuales.

La evaluación de un sistema de recomendación tampoco puede limitarse a una única métrica de precisión. El desempeño debe analizarse en función de la tarea que se pretende apoyar, el tipo de usuario, la naturaleza de los datos y la utilidad real de la recomendación. La literatura especializada ha demostrado que diferentes tareas requieren métodos de evaluación distintos y que una mejora en la exactitud algorítmica no garantiza por sí sola una mejor experiencia o una decisión más adecuada (Herlocker et al., 2004).

Esta limitación se intensifica cuando se emplean modelos complejos cuyo funcionamiento interno no resulta fácilmente comprensible. Un sistema puede generar una clasificación o probabilidad con un desempeño estadístico aceptable y, al mismo tiempo, ofrecer poca información sobre las razones

que sustentan cada resultado. En contextos de bajo impacto, esta opacidad puede representar un inconveniente operativo. En entornos institucionales, donde las decisiones deben justificarse y someterse a revisión, puede convertirse en una limitación sustantiva.

La inteligencia artificial explicable, conocida como XAI por sus siglas en inglés, se orienta a desarrollar métodos y principios que permitan comprender o describir el funcionamiento de los modelos y las razones asociadas con sus resultados. La explicabilidad no implica necesariamente revelar cada operación matemática ejecutada por el sistema. Su propósito consiste en proporcionar evidencias o razones comprensibles, ajustadas a las necesidades de los usuarios y suficientemente fieles al proceso que produjo la salida.

El Instituto Nacional de Estándares y Tecnología de Estados Unidos ha propuesto cuatro principios generales para evaluar una explicación: el sistema debe proporcionar evidencia o razones para sus resultados; la explicación debe ser comprensible para su destinatario; debe representar correctamente el proceso mediante el cual se produjo la salida; y el sistema debe reconocer los límites dentro de los cuales puede operar de manera adecuada. Estos principios muestran que la explicación no constituye únicamente una propiedad técnica, sino una relación entre el modelo, el resultado y la persona que necesita comprenderlo (Phillips et al., 2021).

La relación entre inteligencia artificial explicable y sistemas de recomendación ha dado lugar al campo de la recomendación explicable. Este enfoque pretende que el sistema no se limite a presentar una opción, sino que proporcione información acerca de las razones que justifican su selección. Las explicaciones pueden servir para que los usuarios comprendan la recomendación, detecten posibles errores, revisen los factores considerados y formen una valoración más crítica sobre el resultado (Zhang & Chen, 2020).

En el contexto de esta obra, la recomendación explicable se plantea como un mecanismo de apoyo y no como una forma de decisión automatizada. La propuesta articula datos, un modelo predictivo, una estimación, una comparación entre alternativas y una explicación dirigida al usuario. Cada

uno de estos componentes posee requisitos y limitaciones propias. La calidad del modelo no garantiza automáticamente la calidad de la explicación, del mismo modo que una explicación clara no demuestra que la predicción sea correcta.

El presente capítulo establece los fundamentos conceptuales necesarios para analizar esta relación. Primero se revisa el desarrollo de los modelos predictivos, la inteligencia artificial explicable y los sistemas de recomendación. Después se examinan los principales paradigmas del aprendizaje automático, los algoritmos de clasificación y las técnicas empleadas para explicar sus resultados. Finalmente, se integran estos elementos dentro del contexto de los sistemas explicables de apoyo a decisiones.

Estado del conocimiento sobre inteligencia artificial explicable y recomendación

La evolución de los modelos predictivos y de los sistemas de recomendación ha estado acompañada por un cambio progresivo en los criterios utilizados para valorar su calidad. Durante las primeras etapas del desarrollo de estos sistemas, gran parte de la investigación se concentró en mejorar su capacidad para clasificar casos, predecir valores o recomendar objetos. Con el aumento de la complejidad de los modelos y su incorporación a contextos con consecuencias sociales, económicas e institucionales, la precisión dejó de considerarse el único criterio relevante.

Actualmente, el análisis de un sistema inteligente requiere considerar no solo cuánto acierta, sino también cómo utiliza los datos, qué variables influyen en sus resultados, qué limitaciones presenta, cómo se comporta ante casos distintos y qué información ofrece a las personas que deben interpretar sus salidas. Esta transformación explica el interés creciente por la transparencia algorítmica, la explicabilidad, la equidad, la robustez y la supervisión humana.

Evolución de los modelos predictivos aplicados a la toma de decisiones

Los modelos predictivos se construyen con el propósito de establecer relaciones entre un conjunto de variables de entrada y un resultado que se desea estimar. En los problemas de clasificación, el

objetivo consiste en asignar los casos a una o varias categorías. En los problemas de regresión, se busca estimar un valor continuo. Aunque ambos procedimientos pueden utilizarse para apoyar decisiones, sus resultados deben interpretarse como estimaciones condicionadas por la información disponible.

Los primeros sistemas de inteligencia artificial se apoyaron principalmente en reglas definidas de manera explícita por especialistas. Este enfoque permitió representar conocimiento y establecer secuencias de razonamiento, pero dependía de la posibilidad de anticipar y codificar las condiciones relevantes. El aprendizaje automático modificó este esquema al permitir que el sistema identificara relaciones a partir de ejemplos y datos.

Esta transición amplió la capacidad de los sistemas para trabajar con conjuntos de información de mayor tamaño y con relaciones difíciles de expresar mediante reglas simples. Sin embargo, también generó una nueva dependencia: el comportamiento del modelo comenzó a estar condicionado por la estructura y calidad de los datos de entrenamiento. Si los registros contienen errores, omisiones o relaciones que no deberían proyectarse hacia nuevos casos, el sistema puede reproducir esas limitaciones.

Entre los modelos utilizados para clasificación se encuentran la regresión logística, los árboles de decisión, las máquinas de vectores de soporte, las redes neuronales y los métodos de conjunto. Cada familia ofrece diferentes niveles de flexibilidad, rendimiento e interpretabilidad. Algunos modelos permiten observar directamente la relación entre variables y resultados; otros alcanzan un mayor nivel de complejidad, pero requieren técnicas adicionales para explicar sus predicciones.

Los bosques aleatorios constituyen un ejemplo de método de conjunto. Breiman los definió como colecciones de clasificadores estructurados en forma de árbol, construidos mediante elementos aleatorios y combinados a través de votación. Su funcionamiento busca reducir la dependencia de un único árbol y mejorar la capacidad de generalización mediante la diversidad entre los modelos individuales (Breiman, 2001).

La combinación de múltiples árboles puede mejorar el desempeño predictivo, pero dificulta la representación del resultado mediante una única secuencia de reglas. Mientras un árbol aislado puede leerse desde la raíz hasta una hoja, un bosque aleatorio integra las decisiones de numerosos árboles. Esta característica ayuda a comprender por qué un modelo puede ofrecer un rendimiento adecuado y, al mismo tiempo, requerir métodos específicos para interpretar sus predicciones.

La aplicación de modelos predictivos a la toma de decisiones debe distinguir dos procesos. El primero es la generación de una estimación matemática. El segundo es la valoración humana e institucional de esa estimación. El modelo puede identificar relaciones entre características, pero no determina por sí mismo si esas relaciones son jurídicamente procedentes, éticamente aceptables o técnicamente relevantes para un contexto determinado.

Por esta razón, el diseño de un modelo de apoyo debe comenzar con una definición clara del problema. Es necesario establecer qué resultado se desea predecir, qué variables pueden utilizarse, qué significado posee cada una y cómo se evaluará el desempeño. También deben definirse las consecuencias de los errores. En algunos escenarios, un falso positivo puede resultar más problemático que un falso negativo; en otros, ocurre lo contrario.

La evaluación del modelo debe responder a estas condiciones. La exactitud global puede resultar insuficiente cuando las clases se encuentran desbalanceadas, debido a que un sistema podría obtener un porcentaje aparentemente elevado al predecir siempre la categoría mayoritaria. En consecuencia, la interpretación del rendimiento requiere considerar métricas complementarias y examinar la matriz de confusión, la sensibilidad, la especificidad y la capacidad de discriminación.

En los sistemas destinados a apoyar decisiones institucionales, la evaluación debe incluir además la estabilidad de los resultados, la posibilidad de reproducirlos y la capacidad de documentar el proceso seguido. Un modelo no debe valorarse únicamente por su rendimiento en un conjunto de prueba, sino también por la claridad con la que se han definido sus datos, objetivos, limitaciones y condiciones de uso.

Desarrollo de la inteligencia artificial explicable

El crecimiento de los modelos de aprendizaje automático produjo sistemas capaces de representar relaciones complejas, pero también incrementó la distancia entre la predicción generada y la comprensión humana del proceso. Este problema ha sido descrito mediante la metáfora de la caja negra, utilizada para señalar aquellos modelos cuyas entradas y salidas pueden observarse, aunque las razones internas que conectan unas con otras resulten difíciles de interpretar.

La explicabilidad surgió como respuesta a esta limitación. Su desarrollo no implica rechazar los modelos complejos, sino construir mecanismos que permitan examinarlos y utilizar sus resultados con mayor conocimiento. La explicación puede referirse al comportamiento general del sistema o concentrarse en una predicción específica. También puede formar parte de un modelo diseñado para ser interpretable o generarse después del entrenamiento mediante un procedimiento adicional.

Una de las contribuciones relevantes en este campo fue el desarrollo de LIME, un método destinado a explicar predicciones individuales mediante la aproximación local del comportamiento de un clasificador. Su lógica consiste en analizar cómo cambia la predicción alrededor de un caso particular y construir una representación más sencilla que permita identificar los factores influyentes en esa zona específica (Ribeiro et al., 2016).

Posteriormente, Lundberg y Lee propusieron SHAP como un marco unificado de atribución de características. Este enfoque asigna a cada variable un valor que representa su contribución a una predicción particular, a partir de principios derivados de los valores de Shapley. SHAP permite examinar la dirección y magnitud con que las variables contribuyen al resultado del modelo (Lundberg & Lee, 2017).

Estos métodos muestran que la explicación puede construirse alrededor de una predicción sin reemplazar el modelo original. No obstante, una explicación posterior al entrenamiento debe

evaluarse con cuidado. Una representación simplificada puede ser comprensible, pero no necesariamente reflejar con exactitud todas las operaciones del sistema. La calidad de la explicación depende de su fidelidad, estabilidad, alcance y adecuación al destinatario.

La literatura contemporánea también reconoce que no existe una explicación universalmente adecuada. Una explicación útil para un desarrollador puede resultar demasiado técnica para un funcionario o para un usuario externo. La comprensibilidad depende de la experiencia, los conocimientos, las responsabilidades y las preguntas del destinatario. Por ello, la explicabilidad debe diseñarse en función de las personas que utilizarán la información y del propósito que persigue la explicación.

Los principios propuestos por Phillips et al. (2021) permiten evaluar esta relación desde cuatro dimensiones. La primera exige que la salida se acompañe de evidencia o razones. La segunda establece que esas razones deben poseer significado para el usuario. La tercera requiere que la explicación refleje correctamente el proceso del sistema. La cuarta demanda que la herramienta comunique los límites de su conocimiento y evite producir resultados fuera de las condiciones para las que fue diseñada.

Estas dimensiones resultan especialmente pertinentes para una obra centrada en apoyo a decisiones. La presencia de un gráfico, una puntuación o una lista de variables no garantiza por sí misma que exista una explicación adecuada. Debe comprobarse que el recurso permita comprender el resultado, que represente fielmente la predicción y que advierta sobre sus limitaciones.

La explicabilidad tampoco debe confundirse con la confianza automática. Su propósito no consiste en persuadir al usuario para que acepte el resultado, sino en proporcionarle elementos para evaluarlo. Una explicación responsable puede aumentar la confianza cuando el sistema actúa de manera coherente, pero también debe permitir identificar errores o razones para rechazar la recomendación.

Evolución de los sistemas de recomendación

Los sistemas de recomendación se desarrollaron para ayudar a los usuarios a seleccionar objetos dentro de conjuntos amplios de alternativas. Su funcionamiento tradicional se ha basado en tres grandes enfoques: recomendación mediante contenido, filtrado colaborativo y métodos híbridos. Los sistemas basados en contenido relacionan las características de los objetos con el perfil del usuario. Los modelos colaborativos utilizan interacciones o preferencias de diferentes usuarios para identificar patrones de semejanza. Los enfoques híbridos combinan distintas fuentes y técnicas con el propósito de compensar las limitaciones de cada método (Adomavicius & Tuzhilin, 2005).

La recomendación puede expresarse mediante una predicción de utilidad, una puntuación, un ordenamiento o una lista de elementos. La forma específica depende de la tarea. Un sistema puede predecir la valoración que un usuario asignaría a un producto, ordenar varias alternativas según su relevancia o seleccionar un subconjunto de opciones consideradas compatibles con determinadas preferencias.

Esta variedad obliga a evaluar los sistemas de acuerdo con la tarea que pretenden resolver. Herlocker et al. (2004) señalaron que la evaluación debe considerar el objetivo del usuario, la naturaleza de los datos, el protocolo experimental y las propiedades que se desean medir. No existe una única métrica que represente de manera completa la calidad de todos los sistemas de recomendación.

Los primeros trabajos se concentraron principalmente en la precisión de las predicciones y en la capacidad de encontrar elementos relevantes. Con posterioridad, la investigación incorporó otras dimensiones, como diversidad, novedad, cobertura, utilidad, confianza, control del usuario y transparencia. Esta ampliación refleja el reconocimiento de que una recomendación técnicamente precisa puede no resultar útil cuando el usuario no comprende su origen, cuando presenta opciones demasiado semejantes o cuando no responde a la tarea real.

La recomendación tradicional se ha aplicado con frecuencia a productos, contenidos audiovisuales, noticias, servicios y recursos digitales. Sin embargo, los principios generales pueden adaptarse a otros dominios siempre que se definan correctamente el usuario, el objeto recomendado, las interacciones, el contexto y la función de utilidad.

En contratación pública, esta adaptación requiere precaución. Los proveedores no deben equipararse automáticamente con productos de consumo y la selección de una oferta no responde únicamente a preferencias individuales. Intervienen normas, requisitos, condiciones técnicas, restricciones presupuestarias y responsabilidades institucionales. Por ello, el sistema propuesto debe presentarse como una herramienta de apoyo basada en puntuaciones o estimaciones, no como una reproducción directa de los recomendadores comerciales.

Surgimiento de los sistemas de recomendación explicables

La expansión de modelos de recomendación basados en factores latentes, aprendizaje automático y redes neuronales incrementó su capacidad para identificar relaciones complejas, pero redujo en numerosos casos la posibilidad de comprender por qué se seleccionaba una opción. Esta tensión favoreció el desarrollo de los sistemas de recomendación explicables.

La recomendación explicable busca producir dos salidas relacionadas: una alternativa considerada pertinente y una razón comprensible que permita interpretar esa selección. Las explicaciones pueden estar integradas en el propio modelo o generarse después de obtener la recomendación. También pueden basarse en características del objeto, preferencias del usuario, reglas, ejemplos semejantes, comparaciones, información textual o contribuciones de variables.

Zhang y Chen (2020) organizan este campo alrededor de la necesidad de responder no solo qué se recomienda, sino por qué se recomienda. Desde esta perspectiva, la explicación puede beneficiar tanto a los usuarios como a los diseñadores. Para los primeros, facilita la comprensión y la

evaluación de la sugerencia; para los segundos, puede contribuir a diagnosticar errores, examinar el comportamiento del sistema y mejorar su diseño.

La explicación puede perseguir objetivos diferentes. Puede favorecer la transparencia al mostrar los factores considerados; contribuir a la eficacia al ayudar al usuario a elegir una alternativa; facilitar la revisión del sistema; o mejorar la capacidad para detectar resultados inesperados. Sin embargo, estos objetivos no siempre coinciden. Una explicación persuasiva puede llevar al usuario a aceptar una recomendación sin reflejar fielmente el funcionamiento del modelo.

Por este motivo, debe distinguirse entre explicar y justificar. Una explicación intenta describir los factores o procesos que produjeron la salida. Una justificación presenta razones que hacen parecer adecuada la alternativa. Ambas pueden coincidir, pero no son equivalentes. Un sistema responsable debe evitar explicaciones diseñadas únicamente para aumentar la aceptación del resultado.

En contextos institucionales, la fidelidad de la explicación adquiere especial importancia. La herramienta debe permitir relacionar la recomendación con los datos y criterios efectivamente utilizados. No sería suficiente presentar una razón genérica, como afirmar que un proveedor es más adecuado por poseer mayor experiencia, si esa variable no tuvo una influencia relevante en el resultado del modelo.

La recomendación explicable también requiere reconocer la diversidad de usuarios. El desarrollador necesita comprender el comportamiento técnico del algoritmo; el analista requiere examinar variables y comparaciones; la autoridad necesita identificar los criterios relevantes; y los participantes externos pueden necesitar una explicación accesible sobre las bases generales del análisis. Diseñar una única explicación para todos estos perfiles puede reducir su utilidad.

En la propuesta de este libro, la explicación debe funcionar como un componente de revisión. Su finalidad no será convencer al usuario de aceptar automáticamente la alternativa con mayor

puntuación, sino proporcionarle elementos para analizarla, contrastarla con la información disponible y determinar si el resultado resulta coherente con las condiciones del caso.

Fundamentos de inteligencia artificial y aprendizaje automático

La inteligencia artificial constituye un campo interdisciplinario dedicado al estudio y desarrollo de sistemas capaces de realizar procesos vinculados con la percepción, el razonamiento, el aprendizaje, la resolución de problemas y la selección de acciones. Su desarrollo integra conocimientos procedentes de la informática, las matemáticas, la estadística, la lógica, la psicología cognitiva, la lingüística y otras disciplinas interesadas en comprender o reproducir determinadas capacidades asociadas con el comportamiento inteligente.

Uno de los documentos fundacionales del campo fue la propuesta elaborada por McCarthy, Minsky, Rochester y Shannon para el proyecto de investigación de Dartmouth. En ese documento se planteó que ciertos componentes del aprendizaje y de la inteligencia podían describirse con suficiente precisión como para ser simulados por una máquina. Esta formulación no ofrecía una definición definitiva de inteligencia artificial, pero estableció un programa de investigación centrado en el aprendizaje, el lenguaje, la formación de conceptos, la resolución de problemas y el mejoramiento de las máquinas mediante la experiencia (McCarthy et al., 1955).

Desde entonces, la inteligencia artificial ha evolucionado a través de diferentes enfoques. Algunos se han orientado a representar reglas y conocimientos definidos por especialistas; otros se han concentrado en desarrollar sistemas que aprenden relaciones a partir de los datos. También se han formulado aproximaciones basadas en el razonamiento lógico, la optimización, la probabilidad, la búsqueda, la planificación y la interacción de agentes con su entorno.

Una definición contemporánea debe reconocer esta diversidad. La Organización para la Cooperación y el Desarrollo Económicos caracteriza un sistema de inteligencia artificial como un sistema basado en máquinas que, a partir de objetivos explícitos o implícitos, infiere cómo generar

resultados tales como predicciones, contenidos, recomendaciones o decisiones capaces de influir en entornos físicos o virtuales. Asimismo, estos sistemas pueden diferenciarse por su grado de autonomía y por su capacidad de adaptación después de haber sido implementados (OECD, 2024).

Esta conceptualización resulta pertinente para la presente obra porque distingue entre varios tipos de salida. Una predicción, una recomendación y una decisión no son equivalentes. El sistema puede estimar una probabilidad o priorizar determinadas alternativas, pero la decisión institucional implica una valoración posterior en la que deben considerarse disposiciones jurídicas, criterios técnicos, condiciones económicas y responsabilidades administrativas.

Por consiguiente, la inteligencia artificial no debe entenderse como una tecnología única ni como una entidad que reproduce de manera completa la inteligencia humana. Comprende un conjunto de métodos diseñados para resolver problemas delimitados. La capacidad del sistema depende de la tarea para la que fue construido, de la representación del problema, de los datos disponibles y de los criterios utilizados para evaluar su funcionamiento.

Enfoques para comprender la inteligencia artificial

Russell y Norvig (2021) organizan históricamente las definiciones de inteligencia artificial en cuatro enfoques: sistemas que piensan como los seres humanos, sistemas que actúan como los seres humanos, sistemas que piensan racionalmente y sistemas que actúan racionalmente. Esta clasificación permite conservar uno de los recursos conceptuales más valiosos del manuscrito original, pero requiere diferenciar claramente las finalidades de cada perspectiva.

El enfoque orientado a pensar como los seres humanos se interesa por representar los procesos cognitivos que intervienen en actividades como la memoria, el aprendizaje, el razonamiento y la resolución de problemas. Su propósito no consiste solamente en obtener una respuesta correcta, sino en aproximarse a la forma en que las personas procesan información y construyen sus decisiones.

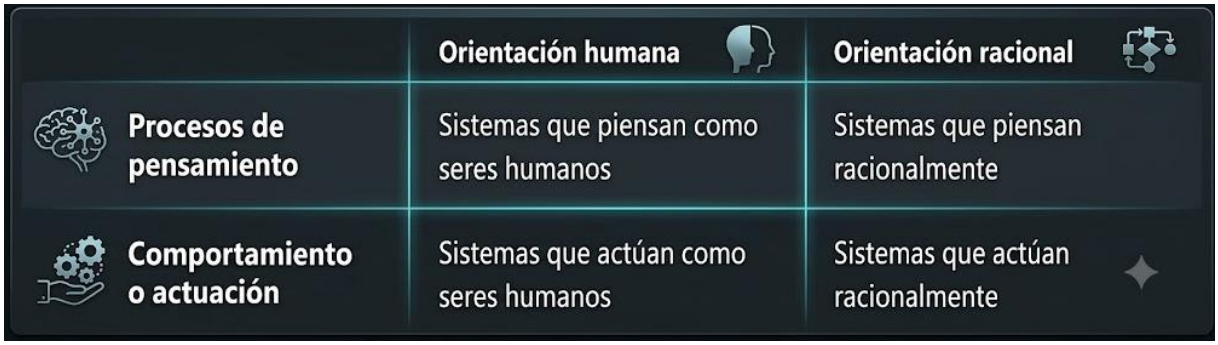
El enfoque centrado en actuar como los seres humanos evalúa el comportamiento observable del sistema. Desde esta perspectiva, una máquina puede considerarse inteligente cuando su actuación presenta características semejantes a las humanas en tareas relacionadas con el lenguaje, la percepción, el razonamiento o la interacción.

En el contexto de los sistemas de apoyo a decisiones, el enfoque de los agentes racionales resulta particularmente útil. El sistema recibe información, procesa variables y produce una salida destinada a contribuir al cumplimiento de un objetivo. No obstante, esa racionalidad es limitada: el modelo solo puede utilizar los datos y criterios que fueron incorporados en su diseño. Si el objetivo se encuentra mal definido o la información es insuficiente, la salida puede ser técnicamente coherente y, al mismo tiempo, inadecuada para la necesidad institucional.

La inteligencia artificial aplicada a contratación pública debe entenderse, por tanto, como un mecanismo delimitado de procesamiento y apoyo. El sistema no posee comprensión jurídica autónoma ni determina por sí mismo el interés público. Su operación depende de decisiones humanas previas relacionadas con la definición del problema, la selección de variables, la construcción de los datos y la interpretación de los resultados.

Figura 1

Enfoques históricos para comprender la inteligencia artificial



Nota. *Elaboración propia basada en la clasificación presentada por Russell y Norvig (2021).*

Relación entre inteligencia artificial, aprendizaje automático y aprendizaje profundo

La inteligencia artificial representa el campo general dentro del cual se desarrollan métodos destinados a producir comportamientos o resultados asociados con capacidades inteligentes. El aprendizaje automático constituye una de sus áreas y se enfoca en construir sistemas que mejoran su desempeño a partir de datos o experiencias. El aprendizaje profundo, a su vez, corresponde a una familia particular de métodos de aprendizaje automático basada en redes neuronales con múltiples niveles de representación.

Esta relación debe interpretarse como una organización de campos y no como una secuencia histórica en la que una tecnología reemplaza completamente a la anterior. Existen sistemas de inteligencia artificial que no utilizan aprendizaje automático, como determinados sistemas basados en reglas, lógica o búsqueda. Del mismo modo, numerosos modelos de aprendizaje automático no emplean redes neuronales profundas.

El aprendizaje automático permite construir una función o modelo que relaciona información de entrada con una salida esperada. El proceso de aprendizaje consiste en ajustar los parámetros del modelo utilizando ejemplos, de manera que el sistema pueda generalizar lo aprendido y producir resultados para casos que no formaron parte del entrenamiento. Por tanto, el objetivo no es memorizar los datos disponibles, sino identificar relaciones que puedan mantenerse en observaciones nuevas (Hastie et al., 2009; Goodfellow et al., 2016).

La generalización constituye una condición esencial. Un modelo puede describir con gran precisión los datos con los que fue entrenado y presentar un desempeño deficiente al recibir información nueva. Esta situación se conoce como sobreajuste. En el extremo contrario, el subajuste ocurre cuando el modelo resulta demasiado simple para representar las relaciones relevantes y presenta un rendimiento insuficiente incluso con los datos de entrenamiento.

El desempeño depende también de la representación utilizada. Los datos deben convertirse en características que el modelo pueda procesar. En el caso de proveedores de servicios tecnológicos, estas características podrían incluir experiencia, cantidad de proyectos, costo, tiempos de respuesta, certificaciones o cumplimiento. Sin embargo, la existencia de una variable disponible no significa que deba incorporarse automáticamente. Su selección debe responder a una justificación técnica y a una relación pertinente con el objetivo del análisis.

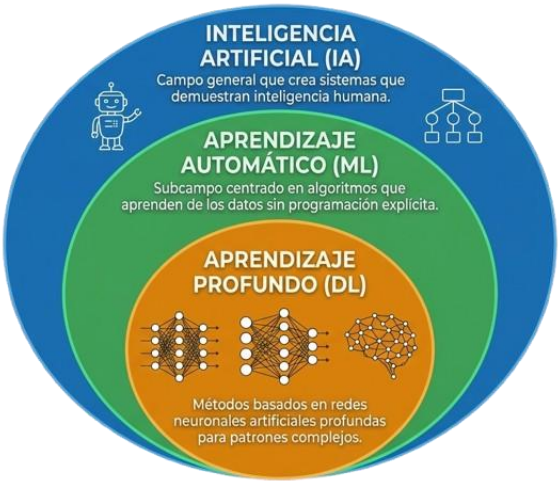
El aprendizaje profundo utiliza redes neuronales organizadas en múltiples capas, capaces de construir representaciones progresivamente más complejas. Las primeras capas pueden identificar relaciones elementales, mientras que las posteriores combinan esas relaciones para representar estructuras de mayor nivel. Esta capacidad ha sido especialmente relevante para imágenes, lenguaje, sonido y otros tipos de información de alta dimensionalidad (Goodfellow et al., 2016).

No obstante, el aprendizaje profundo no constituye necesariamente la mejor opción para todos los problemas. Estos modelos suelen requerir grandes cantidades de datos, recursos computacionales y procedimientos de ajuste más complejos. En problemas tabulares con una cantidad moderada de observaciones, pueden resultar apropiados otros modelos, como la regresión logística, los árboles de decisión o los métodos de conjunto.

El prototipo desarrollado en esta obra no emplea aprendizaje profundo. Utiliza un modelo de bosque aleatorio aplicado a datos tabulares. Por esta razón, el aprendizaje profundo se presenta como parte del contexto general de la inteligencia artificial, pero no debe ocupar una extensión desproporcionada ni sugerir que forma parte de la arquitectura implementada.

Figura 2

Relación entre inteligencia artificial, aprendizaje automático y aprendizaje profundo



Nota. *Elaboración propia basada en Goodfellow et al. (2016) y Russell y Norvig (2021).*

Paradigmas del aprendizaje automático

Los métodos de aprendizaje automático pueden organizarse de acuerdo con la información disponible durante el entrenamiento y con el tipo de tarea que se espera resolver. Las categorías más conocidas son aprendizaje supervisado, aprendizaje no supervisado, aprendizaje semisupervisado y aprendizaje por refuerzo. Cada paradigma responde a una relación distinta entre datos, objetivos y retroalimentación.

Aprendizaje supervisado

El aprendizaje supervisado utiliza observaciones que contienen variables de entrada y un resultado previamente conocido. El modelo analiza la relación entre ambos componentes y ajusta sus parámetros para predecir la salida correspondiente a nuevos casos. Cuando el resultado pertenece a una categoría, la tarea se denomina clasificación. Cuando corresponde a una medida numérica continua, se denomina regresión (Hastie et al., 2009; Goodfellow et al., 2016).

En un problema de clasificación binaria, cada observación pertenece a una de dos clases. El modelo puede asignar directamente una categoría o estimar una probabilidad que posteriormente se transforma en clase mediante un umbral. Esta distinción es importante porque la probabilidad

conserva más información que la etiqueta final y permite examinar el grado de seguridad de la estimación.

El aprendizaje supervisado requiere que la variable objetivo sea coherente con la finalidad del modelo. Si las etiquetas se encuentran mal definidas, contienen errores o fueron generadas sin relación con las variables predictoras, el sistema no podrá aprender un patrón válido. La cantidad de datos no compensa una variable objetivo carente de significado.

En el prototipo del libro, el bosque aleatorio se presenta como un clasificador supervisado. Las variables relacionadas con las características del proveedor constituyen las entradas, mientras que la condición que se desea predecir constituye la variable objetivo. Por ello, la calidad de esa etiqueta y la relación que mantenga con las variables son determinantes para evaluar el modelo.

Aprendizaje no supervisado

El aprendizaje no supervisado trabaja con datos que no incluyen una variable objetivo previamente etiquetada. Su finalidad consiste en reconocer estructuras, agrupaciones, relaciones o representaciones internas dentro del conjunto de observaciones. Entre sus tareas se encuentran el agrupamiento, la reducción de dimensionalidad, la estimación de densidades y la detección de casos atípicos (Hastie et al., 2009).

En los métodos de agrupamiento, el algoritmo intenta formar conjuntos cuyos integrantes presenten mayor semejanza entre sí que respecto de las observaciones pertenecientes a otros grupos. La utilidad de los agrupamientos depende de las variables seleccionadas, de la medida de similitud y del método empleado. Los grupos encontrados por el algoritmo no representan necesariamente categorías naturales ni poseen automáticamente un significado institucional.

En contratación pública, un método no supervisado podría utilizarse de manera exploratoria para identificar perfiles semejantes de proveedores, patrones de comportamiento o registros atípicos.

Sin embargo, estos resultados necesitarían interpretación humana y no deberían transformarse de manera directa en criterios de elegibilidad o exclusión.

Aprendizaje semisupervisado

El aprendizaje semisupervisado combina una cantidad limitada de observaciones etiquetadas con un conjunto mayor de datos sin etiqueta. Su utilidad aparece cuando producir etiquetas confiables requiere tiempo, recursos o intervención especializada. El modelo utiliza la información disponible en ambos grupos para intentar mejorar su capacidad de generalización (Goodfellow et al., 2016).

Su aplicación exige precaución, porque los datos no etiquetados pueden contener distribuciones o patrones que no correspondan con el problema de interés. La incorporación de información adicional solo resulta beneficiosa cuando existe una relación adecuada entre los datos etiquetados, los no etiquetados y la tarea que se desea resolver.

Aprendizaje por refuerzo

El aprendizaje por refuerzo se centra en un agente que interactúa con un entorno, selecciona acciones y recibe señales de recompensa o penalización. El objetivo consiste en aprender una estrategia que maximice la recompensa acumulada a lo largo del tiempo. A diferencia del aprendizaje supervisado, el agente no recibe necesariamente la respuesta correcta para cada situación; debe aprender mediante las consecuencias de sus acciones (Sutton & Barto, 2018).

Este paradigma se aplica a problemas secuenciales en los que una acción puede influir en estados y recompensas posteriores. Aunque constituye un área relevante de la inteligencia artificial, no forma parte del prototipo desarrollado en el libro. Su inclusión en este marco teórico tiene una finalidad delimitadora y permite evitar que todas las formas de aprendizaje automático sean tratadas como equivalentes.

Tabla 2

Principales paradigmas del aprendizaje automático

Paradigma	*Información disponible*	*Propósito general*	*Ejemplos de tareas*	*Relación con el prototipo*
Aprendizaje supervisado	Datos de entrada y resultados conocidos	Predecir categorías o valores para nuevos casos	Clasificación y regresión	Es el paradigma utilizado
Aprendizaje no supervisado	Datos sin una variable objetivo etiquetada	Identificar estructuras o patrones internos	Agrupamiento, reducción de dimensionalidad y detección de anomalías	Puede utilizarse como análisis exploratorio
Aprendizaje semisupervisado	Pocos datos etiquetados y numerosos datos sin etiqueta	Aprovechar información no etiquetada para mejorar el aprendizaje	Clasificación con etiquetado limitado	No se implementa
Aprendizaje por refuerzo	Interacción con un entorno y señales de recompensa	Aprender una política de acciones secuenciales	Control, planificación y optimización dinámica	No se implementa

Nota. *Elaboración propia basada en Goodfellow et al. (2016), Hastie et al. (2009) y Sutton y Barto (2018).*

Datos, variables y representación del problema

El aprendizaje automático comienza antes de seleccionar un algoritmo. La primera decisión consiste en transformar una necesidad real en un problema que pueda ser representado mediante datos. Este proceso requiere definir la unidad de análisis, las variables de entrada, el resultado esperado y los criterios que permitirán valorar si el modelo cumple su finalidad.

La unidad de análisis determina qué representa cada registro. Puede corresponder a un proveedor, una oferta, un contrato, una interacción o una combinación de estos elementos. Mezclar unidades distintas dentro de un mismo conjunto de datos puede producir relaciones difíciles de interpretar y errores en la evaluación.

Las variables de entrada, también denominadas características o predictores, representan los atributos utilizados por el modelo. Pueden ser numéricas, categóricas, binarias, ordinales o textuales. Cada tipo requiere un tratamiento coherente. Las transformaciones aplicadas deben documentarse porque modifican la forma en que el sistema interpreta la información.

La variable objetivo representa el resultado que el modelo intenta estimar. Su definición determina el sentido de todo el proceso de aprendizaje. En un contexto de contratación pública, predecir la adjudicación histórica no equivale a estimar capacidad técnica, pertinencia de una oferta, cumplimiento futuro o riesgo contractual. Cada objetivo responde a una pregunta diferente y requiere variables, datos y criterios de evaluación propios.

También debe examinarse la posibilidad de que ciertas variables introduzcan información indebida o produzcan una fuga de datos. La fuga ocurre cuando el conjunto de predictores contiene información que no estaría disponible en el momento real de generar la predicción o que revela indirectamente el resultado esperado. En esas condiciones, el modelo puede presentar un rendimiento elevado durante la evaluación y fracasar al ser utilizado en un escenario real.

Los datos históricos tampoco deben considerarse neutrales. Reflejan las condiciones, reglas y decisiones existentes en el período en que fueron generados. Un modelo entrenado con dichos registros puede reproducir patrones previos, incluso cuando esos patrones no resulten pertinentes para decisiones futuras. Por este motivo, la selección de las variables debe combinar criterios estadísticos con revisión técnica, institucional y jurídica.

La preparación de los datos comprende actividades como identificación de duplicados, tratamiento de valores ausentes, revisión de valores extremos, transformación de categorías, normalización y separación entre conjuntos de entrenamiento y evaluación. Estas decisiones deben realizarse antes de interpretar el rendimiento del modelo y deben quedar registradas para favorecer la reproducibilidad.

Entrenamiento, evaluación y generalización

El conjunto de datos utilizado para desarrollar un modelo suele dividirse en información de entrenamiento y de evaluación. El primer grupo permite ajustar los parámetros del algoritmo. El segundo se reserva para analizar su comportamiento ante observaciones que no fueron utilizadas directamente durante el aprendizaje. Esta separación busca aproximarse a la situación en la que el sistema recibe casos nuevos.

La evaluación debe evitar que la información del conjunto de prueba influya durante el entrenamiento. Cuando las decisiones de ajuste se toman repetidamente observando los resultados del mismo conjunto de prueba, este deja de representar una evaluación independiente y puede producir una estimación optimista del rendimiento.

La validación cruzada permite repetir el proceso de entrenamiento y evaluación utilizando diferentes particiones de los datos. Su finalidad es obtener una estimación más estable del comportamiento del modelo y reducir la dependencia de una única división. No obstante, la validación cruzada tampoco corrige problemas derivados de etiquetas incorrectas, variables inadecuadas o datos que no representan el contexto de aplicación.

Las métricas deben seleccionarse de acuerdo con la tarea y con las consecuencias de los errores. La exactitud indica la proporción total de predicciones correctas, pero puede ser engañosa cuando una clase es mucho más frecuente que la otra. En esos casos deben considerarse medidas que permitan observar por separado el comportamiento del sistema frente a cada categoría.

El rendimiento estadístico representa solo una dimensión de la evaluación. En un sistema de apoyo institucional también deben analizarse la estabilidad, la trazabilidad, la capacidad de explicación, la posibilidad de auditoría y la correspondencia entre las variables utilizadas y la finalidad declarada.

La generalización no puede establecerse únicamente mediante el desempeño alcanzado con datos simulados. Un conjunto sintético permite comprobar el funcionamiento del código y desarrollar escenarios de demostración, pero no representa necesariamente la variabilidad, las inconsistencias y las relaciones existentes en los datos reales. Por ello, los resultados obtenidos en el prototipo deben interpretarse dentro de su alcance experimental.

Alcance del aprendizaje profundo en la presente obra

El aprendizaje profundo debe conservarse en el marco teórico como una línea relevante del desarrollo contemporáneo de la inteligencia artificial, pero no como componente central del prototipo. Su utilidad se ha demostrado principalmente en tareas que requieren aprender representaciones complejas a partir de imágenes, lenguaje, sonido y grandes volúmenes de información.

Las redes neuronales profundas integran múltiples capas que transforman progresivamente las entradas y permiten representar relaciones no lineales. Esta capacidad puede ofrecer ventajas en tareas complejas, aunque también puede dificultar la interpretación directa de los resultados y aumentar las exigencias de datos y recursos computacionales (Goodfellow et al., 2016).

Para el problema desarrollado en el libro, las observaciones se representan mediante variables tabulares y el modelo implementado corresponde a un bosque aleatorio. No existe evidencia en el manuscrito que justifique la incorporación de una red neuronal profunda. Incluirla únicamente por su relevancia general podría aumentar innecesariamente la complejidad sin demostrar una mejora en el desempeño o en la utilidad de la herramienta.

En consecuencia, se elimina la narración anterior sobre Google Brain, DeepFace, juegos de Atari y la adquisición de DeepMind. Esos antecedentes no contribuyen de manera directa a explicar la arquitectura del prototipo ni las decisiones metodológicas adoptadas. El marco teórico debe

concentrarse en los modelos que permiten comprender y evaluar la solución realmente desarrollada.

El aprendizaje supervisado proporciona el marco general para construir modelos que relacionan características de entrada con resultados conocidos. No obstante, los algoritmos disponibles presentan diferencias importantes en su estructura, complejidad, capacidad predictiva e interpretabilidad. El apartado siguiente examina los principales modelos de clasificación relacionados con el prototipo, con especial atención a la regresión logística, los árboles de decisión y los bosques aleatorios.

Modelos predictivos para clasificación y apoyo a decisiones

Los modelos predictivos de clasificación buscan establecer una relación entre un conjunto de variables de entrada y una categoría que representa el resultado de interés. A partir de observaciones previamente etiquetadas, el algoritmo identifica patrones que posteriormente utiliza para clasificar casos nuevos o estimar la probabilidad de pertenencia a una determinada categoría. La salida generada no debe entenderse como una certeza, sino como una estimación condicionada por los datos, las variables, el algoritmo y el procedimiento de evaluación empleado.

En un problema de clasificación binaria existen dos categorías posibles. Una de ellas suele definirse como clase positiva y la otra como clase negativa. Esta designación no implica que una categoría sea favorable o desfavorable en sentido ético; solamente permite organizar el cálculo de las métricas. En un escenario experimental relacionado con proveedores, por ejemplo, podrían distinguirse dos resultados históricos. Sin embargo, la etiqueta utilizada debe corresponder con precisión al fenómeno que se desea estudiar.

La definición de la variable objetivo constituye una de las decisiones más importantes del modelado. Predecir si una empresa obtuvo previamente una adjudicación no equivale a estimar su capacidad técnica, su nivel de cumplimiento, la pertinencia de su propuesta o el riesgo de ejecución

contractual. Cada uno de estos objetivos representa un problema distinto y necesita datos, variables y criterios de evaluación específicos.

Los modelos de clasificación pueden producir directamente una categoría o generar una probabilidad que posteriormente se transforma en clase mediante un umbral. Cuando se utiliza el umbral convencional de 0,50, una probabilidad igual o superior puede asignarse a la clase positiva. No obstante, este punto de corte no es obligatorio ni necesariamente adecuado para todos los problemas. Su selección debe considerar la distribución de las clases y las consecuencias asociadas con cada tipo de error.

En los sistemas de apoyo a decisiones, la probabilidad estimada puede resultar más informativa que la etiqueta final. Dos observaciones clasificadas dentro de la misma categoría pueden presentar probabilidades muy diferentes. Esta información permite identificar casos cercanos al umbral, reconocer niveles de incertidumbre y priorizar aquellos resultados que requieren una revisión humana más detallada.

La regresión logística, los árboles de decisión y los bosques aleatorios representan tres aproximaciones relevantes para comprender el prototipo de esta obra. Estos modelos poseen estructuras diferentes y ofrecen distintos niveles de flexibilidad e interpretabilidad. Su comparación permite entender que la selección de un algoritmo no debe basarse únicamente en la precisión, sino también en la naturaleza de los datos, la estabilidad, la capacidad de explicación y la finalidad institucional de la herramienta.

Regresión logística binaria

La regresión logística constituye un método estadístico utilizado para modelar la relación entre varias características predictoras y una variable categórica. En su modalidad binaria, estima la probabilidad de que una observación pertenezca a una de dos categorías. Para ello, transforma una

combinación de las variables de entrada mediante una función logística que mantiene la probabilidad estimada dentro del intervalo comprendido entre cero y uno (Hastie et al., 2009).

A diferencia de una regresión lineal convencional, la regresión logística no predice directamente un valor continuo sin restricciones. Modela el logaritmo de la razón entre la probabilidad de que ocurra un resultado y la probabilidad de que no ocurra. Esta formulación permite relacionar las variables predictoras con la probabilidad de pertenencia a una categoría.

Una de sus principales ventajas es la relativa facilidad con que pueden interpretarse sus coeficientes. El signo de cada parámetro indica la dirección de la relación con el resultado, mientras que su magnitud permite examinar el cambio asociado con la variable cuando las demás permanecen constantes. Esta característica convierte a la regresión logística en un modelo de referencia útil para comparar algoritmos más complejos.

No obstante, la interpretación de los coeficientes requiere precaución. La presencia de variables correlacionadas, transformaciones, interacciones o escalas diferentes puede modificar su significado. Además, el modelo supone una relación lineal entre los predictores y el logaritmo de las probabilidades relativas. Cuando la estructura real es altamente no lineal, la regresión logística puede necesitar términos adicionales o presentar un rendimiento inferior al de métodos más flexibles.

La regresión logística multinomial amplía el procedimiento a variables objetivo con más de dos categorías. Sin embargo, el prototipo del libro se estructura como un problema binario, por lo que no resulta necesario dedicar una extensión amplia al modelo multinomial. Su inclusión debe limitarse a reconocer que existe como alternativa para problemas de clasificación con múltiples clases.

En un estudio aplicado, la regresión logística puede utilizarse como línea de base. Comparar un modelo complejo con una alternativa más sencilla permite determinar si el aumento de complejidad

produce una mejora suficientemente relevante. Si ambos modelos alcanzan un comportamiento semejante, podría resultar preferible el que ofrezca mayor transparencia, estabilidad y facilidad de implementación.

Árboles de decisión

Los árboles de decisión son modelos que dividen progresivamente el espacio de los datos mediante reglas organizadas en una estructura jerárquica. Cada nodo interno representa una condición aplicada a una variable, las ramas muestran los posibles resultados de esa condición y los nodos terminales o hojas contienen la clasificación o estimación final (Breiman et al., 1984; Hastie et al., 2009).

El proceso comienza en el nodo raíz, que contiene el conjunto inicial de observaciones. El algoritmo selecciona una variable y un punto de división para formar subconjuntos más homogéneos respecto de la variable objetivo. Este procedimiento se repite en cada rama hasta alcanzar una condición de detención relacionada con la profundidad, la cantidad de observaciones o el grado de pureza de los nodos.

En los problemas de clasificación, la selección de las divisiones puede realizarse mediante criterios como el índice de Gini o la entropía. Ambos buscan identificar particiones que reduzcan la mezcla de categorías dentro de los nodos resultantes. Un nodo se considera más puro cuando contiene una proporción elevada de observaciones pertenecientes a una misma clase.

La estructura jerárquica facilita la interpretación del modelo. Un usuario puede seguir la secuencia desde el nodo raíz hasta una hoja e identificar las condiciones que condujeron a una clasificación. Esta propiedad resulta valiosa cuando se necesita presentar reglas comprensibles o revisar el razonamiento aplicado a un caso particular.

La interpretabilidad disminuye cuando el árbol alcanza una profundidad elevada. Un árbol excesivamente grande puede generar numerosas reglas, depender de pequeñas variaciones en los

datos y aprender características específicas del conjunto de entrenamiento. Esta situación puede producir sobreajuste y reducir el desempeño ante nuevas observaciones.

La poda constituye un procedimiento destinado a disminuir la complejidad del árbol. Consiste en eliminar divisiones o ramas que aportan una mejora limitada y que pueden afectar la capacidad de generalización. También pueden establecerse restricciones durante el crecimiento, como una profundidad máxima, una cantidad mínima de observaciones por nodo o un número mínimo de casos requeridos para realizar una nueva división.

Una limitación importante de los árboles individuales es su inestabilidad. Pequeñas modificaciones en los datos pueden producir una estructura distinta, porque una división diferente en los primeros niveles modifica todas las ramas posteriores. Esta sensibilidad favoreció el desarrollo de métodos de conjunto que combinan numerosos árboles para generar predicciones más estables.

Figura 3

Estructura general de un árbol de decisión para clasificación

Nota. *Elaboración propia basada en Breiman et al. (1984) y Hastie et al. (2009).*

Métodos de conjunto

Los métodos de conjunto combinan las predicciones de varios modelos con el propósito de obtener un resultado más estable o preciso que el producido por cada componente de manera individual. La lógica general consiste en construir múltiples modelos y agregar sus salidas mediante votación, promedio u otro procedimiento de combinación.

La efectividad del conjunto depende de que sus modelos individuales presenten cierto grado de diversidad. Si todos producen exactamente los mismos errores, su combinación ofrecerá pocas ventajas. En cambio, cuando cada modelo representa aspectos diferentes de los datos, los errores individuales pueden compensarse parcialmente.

El método conocido como bagging construye varias versiones del modelo utilizando muestras obtenidas del conjunto original. Cada muestra se genera mediante selección aleatoria con reemplazo, por lo que algunas observaciones pueden aparecer varias veces y otras no ser incluidas. Los resultados de los modelos se agregan posteriormente para generar la predicción final.

En clasificación, la agregación puede realizarse mediante votación. Cada modelo asigna una categoría y el conjunto selecciona la clase que recibe la mayor cantidad de votos. También puede calcularse el promedio de las probabilidades producidas por los clasificadores y aplicar un umbral al valor resultante.

La combinación de modelos puede reducir la variabilidad de algoritmos inestables, como los árboles de decisión. Sin embargo, el aumento del número de componentes también puede dificultar la interpretación directa. Mientras que un árbol individual puede representarse mediante una secuencia de reglas, un conjunto integrado por numerosos árboles requiere procedimientos adicionales para analizar su comportamiento.

Bosques aleatorios

Los bosques aleatorios constituyen un método de conjunto que combina numerosos árboles de decisión. Cada árbol se construye a partir de una muestra aleatoria de los datos y, en cada división, considera solamente un subconjunto de las variables disponibles. La predicción final se obtiene mediante la agregación de las salidas de todos los árboles (Breiman, 2001).

La aleatoriedad se incorpora en dos niveles. El primero corresponde a la selección de las observaciones utilizadas para construir cada árbol. El segundo se relaciona con el subconjunto de

variables que puede examinarse en cada nodo. Esta combinación reduce la semejanza entre los árboles y evita que una misma característica domine sistemáticamente todas las estructuras.

En un problema de clasificación, cada árbol genera una predicción y el bosque determina el resultado mediante votación o promedio de probabilidades. La combinación suele presentar menor variabilidad que un árbol individual y puede representar relaciones no lineales e interacciones entre variables sin necesidad de especificarlas previamente.

Los bosques aleatorios pueden trabajar con numerosas características, capturar relaciones complejas y ofrecer un buen desempeño en datos tabulares. También requieren menos decisiones de especificación que determinados modelos paramétricos. No obstante, estas ventajas no significan que funcionen correctamente en cualquier conjunto de datos ni que puedan compensar etiquetas aleatorias, variables mal definidas o información insuficiente.

El número de árboles controla la cantidad de clasificadores incluidos en el conjunto. Incrementar este valor puede estabilizar las estimaciones, aunque también aumenta el tiempo de procesamiento. La profundidad máxima, el número mínimo de observaciones requerido para dividir un nodo y la cantidad de variables examinadas en cada partición influyen en la complejidad de los árboles y en el riesgo de sobreajuste.

Una ventaja del procedimiento es la posibilidad de estimar el error mediante las observaciones que no fueron seleccionadas en la muestra utilizada para construir un árbol. Estas observaciones, conocidas como casos out-of-bag, pueden evaluarse con el árbol correspondiente y proporcionar una estimación interna del rendimiento. Este mecanismo no reemplaza necesariamente un conjunto de prueba independiente, pero constituye un recurso adicional para examinar el modelo (Breiman, 2001).

La principal limitación, desde la perspectiva de esta obra, se relaciona con la interpretabilidad. Un bosque aleatorio es menos transparente que un árbol individual porque la predicción integra

numerosas estructuras. No es posible explicar el resultado mostrando una sola ruta de decisión. Para comprender el comportamiento del modelo se necesitan herramientas como la importancia de variables, las medidas por permutación, los gráficos de dependencia o los métodos de explicación local.

El bosque aleatorio no debe describirse como un modelo completamente inexplicable. Su estructura permite realizar diferentes análisis, aunque estos requieren procedimientos adicionales y poseen limitaciones. La explicabilidad depende tanto del algoritmo como de la forma en que se documentan los datos, se presentan las variables y se comunican los resultados.

Comparación entre los modelos considerados

La regresión logística, el árbol de decisión y el bosque aleatorio representan distintos niveles de complejidad. La primera ofrece una estructura relativamente estable y parámetros interpretables, pero puede presentar dificultades para representar relaciones no lineales. El árbol permite visualizar reglas de clasificación, aunque puede ser inestable y susceptible al sobreajuste. El bosque aleatorio reduce parte de la variabilidad del árbol individual y captura relaciones complejas, pero dificulta la explicación directa.

No existe un modelo universalmente superior. La elección depende de la finalidad, la cantidad y calidad de los datos, la relación entre las variables, las consecuencias de los errores y la necesidad de interpretación. En contextos institucionales, una mejora limitada en el rendimiento no siempre justifica una pérdida considerable de transparencia.

Por esta razón, es recomendable comparar el modelo elegido con una línea de base. La regresión logística y un árbol de profundidad controlada pueden cumplir esta función. La comparación debe realizarse utilizando las mismas particiones de datos y métricas, de manera que la diferencia observada pueda atribuirse al comportamiento de los modelos y no a procedimientos de evaluación distintos.

Tabla 3

Comparación conceptual de modelos predictivos de clasificación

Modelo	*Características principales*	*Ventajas*	*Limitaciones*	*Función dentro de la obra*
Regresión logística	Estima probabilidades mediante una relación paramétrica entre predictores y resultado	Relativa facilidad de interpretación, estabilidad y utilidad como línea de base	Puede representar insuficientemente relaciones no lineales si no se incorporan transformaciones	Modelo de referencia para comparar desempeño e interpretabilidad
Árbol de decisión	Divide los datos mediante reglas jerárquicas	Visualización intuitiva, identificación de reglas e interacciones	Inestabilidad, sensibilidad a los datos y riesgo de sobreajuste	Recurso para comprender la lógica de clasificación
Bosque aleatorio	Combina numerosos árboles construidos con observaciones y variables seleccionadas aleatoriamente	Flexibilidad, reducción de la variabilidad y capacidad para modelar relaciones complejas	Menor transparencia directa y necesidad de métodos complementarios de explicación	Modelo principal implementado en el prototipo
Red neuronal profunda	Utiliza múltiples capas de procesamiento y representaciones no lineales	Capacidad para modelar datos complejos y de alta dimensionalidad	Elevados requisitos de datos, ajuste y explicabilidad	Se presenta como contexto, pero no se implementa

Nota. *Elaboración propia basada en Breiman (2001), Goodfellow et al. (2016) y Hastie et al. (2009).*

Matriz de confusión y tipos de error

La evaluación de un clasificador binario puede organizarse mediante una matriz de confusión. Esta matriz compara la categoría real de cada observación con la clase predicha por el modelo y permite identificar cuatro resultados: verdaderos positivos, verdaderos negativos, falsos positivos y falsos negativos.

Los verdaderos positivos corresponden a los casos pertenecientes a la clase positiva que fueron correctamente identificados. Los verdaderos negativos representan los casos de la clase negativa correctamente clasificados. Un falso positivo ocurre cuando el modelo asigna la clase positiva a una

observación que pertenece a la clase negativa. Un falso negativo se produce cuando un caso positivo es clasificado como negativo.

La relevancia de cada error depende del problema. En determinados contextos, un falso positivo puede generar costos o riesgos más elevados. En otros, el mayor problema puede encontrarse en los falsos negativos. Por tanto, la evaluación no debe limitarse a contabilizar el número total de aciertos, sino que debe examinar las consecuencias asociadas con cada tipo de equivocación.

En el contexto experimental de este libro, la categoría positiva debe definirse de manera explícita. Si se utiliza una variable como "ganador", debe aclararse que la etiqueta representa un resultado simulado y no una evaluación integral de la capacidad del proveedor. Además, no debe interpretarse que un falso positivo o un falso negativo equivalen directamente a una decisión administrativa incorrecta, porque el prototipo no ejecuta adjudicaciones reales.

Métricas para evaluar clasificadores

La exactitud representa la proporción de observaciones clasificadas correctamente respecto del total. Aunque ofrece una visión general, puede resultar engañosa cuando las clases presentan frecuencias muy diferentes. Un modelo que asigna siempre la categoría mayoritaria podría alcanzar una exactitud elevada sin identificar adecuadamente los casos minoritarios.

La precisión indica qué proporción de las observaciones clasificadas como positivas pertenece realmente a esa categoría. Esta métrica permite examinar la confiabilidad de las predicciones positivas y disminuye cuando el modelo genera numerosos falsos positivos.

La sensibilidad, también denominada exhaustividad o recall, muestra qué proporción de los casos positivos reales fue identificada por el modelo. Su valor disminuye cuando existen falsos negativos. La especificidad, en cambio, indica la capacidad para reconocer correctamente los casos negativos.

La medida F1 combina precisión y sensibilidad mediante una media armónica. Puede ser útil cuando se busca mantener un equilibrio entre ambas, especialmente cuando la distribución de las clases no es uniforme. No obstante, tampoco representa todos los aspectos del desempeño y debe interpretarse junto con otras medidas.

La exactitud balanceada calcula el promedio del desempeño alcanzado en cada clase. Su utilización evita que la categoría mayoritaria determine de manera desproporcionada el resultado global. En problemas desbalanceados puede ofrecer una representación más informativa que la exactitud convencional.

La curva ROC relaciona la tasa de verdaderos positivos con la tasa de falsos positivos para diferentes umbrales de clasificación. El área bajo esta curva, conocida como ROC-AUC, resume la capacidad del modelo para distinguir entre las clases a través de esos umbrales. Sin embargo, su interpretación requiere considerar la distribución de las categorías y la finalidad del análisis (Fawcett, 2006).

En conjuntos desbalanceados, la curva de precisión y sensibilidad puede ofrecer información más directa sobre el comportamiento del modelo frente a la clase minoritaria. Saito y Rehmsmeier (2015) demostraron que las curvas de precisión-recall pueden resultar más informativas que las curvas ROC cuando existe una gran diferencia entre la cantidad de observaciones positivas y negativas.

La evaluación debe incluir varias métricas y no seleccionar únicamente la que presente el valor más favorable. También es necesario informar la distribución de las clases, la partición de los datos, el umbral aplicado y, cuando corresponda, la variabilidad obtenida mediante validación cruzada.

Tabla 4

Métricas principales para evaluar un clasificador binario

Métrica	*Pregunta que responde*	*Precaución interpretativa*

Exactitud	¿Qué proporción total fue clasificada correctamente?	Puede ocultar un desempeño deficiente en la clase minoritaria
Precisión	¿Cuántas predicciones positivas fueron correctas?	No informa cuántos casos positivos reales fueron omitidos
Sensibilidad	¿Cuántos casos positivos reales fueron detectados?	Puede aumentar a costa de generar más falsos positivos
Especificidad	¿Cuántos casos negativos reales fueron reconocidos?	Debe interpretarse junto con la sensibilidad
F1	¿Cuál es el equilibrio entre precisión y sensibilidad?	No incorpora directamente los verdaderos negativos
Exactitud balanceada	¿Cuál es el desempeño promedio en las dos clases?	No sustituye el análisis de cada categoría
ROC-AUC	¿Qué capacidad posee el modelo para discriminar clases en distintos umbrales?	Puede ofrecer una percepción optimista con clases muy desbalanceadas
PR-AUC	¿Cómo se relacionan precisión y sensibilidad para la clase positiva?	Depende de la prevalencia de la clase positiva

Elaboración propia basada en Fawcett (2006) y Saito y Rehmsmeier (2015).

Desbalance de clases

Existe desbalance cuando una categoría contiene muchas más observaciones que otra. Esta situación es frecuente en problemas donde el resultado de interés ocurre con menor frecuencia. El desbalance no constituye por sí mismo un error, pero modifica la forma en que deben entrenarse y evaluarse los modelos.

Cuando la clase negativa representa, por ejemplo, el 70 % de las observaciones, un clasificador que predice siempre esa categoría alcanzaría una exactitud de 70 % sin identificar ningún caso positivo. Por esta razón, un porcentaje elevado de aciertos no demuestra automáticamente que el sistema haya aprendido relaciones útiles.

La primera medida consiste en informar la distribución de las clases y examinar la matriz de confusión. También pueden utilizarse particiones estratificadas, que conservan aproximadamente la proporción de categorías en los conjuntos de entrenamiento y evaluación.

Existen estrategias adicionales, como ponderar las clases, modificar el umbral, aumentar la representación de la categoría minoritaria o reducir la clase mayoritaria. Estas técnicas no deben aplicarse mecánicamente. Cada una modifica el proceso de aprendizaje y puede introducir nuevos sesgos o producir sobreajuste si se utiliza antes de separar correctamente los datos.

En el caso del prototipo, el desbalance debe analizarse conjuntamente con la forma en que fue generada la variable objetivo. Ninguna técnica de balanceo puede crear una relación real entre predictores y resultado cuando la etiqueta se asignó aleatoriamente. Corregir las proporciones de las clases no sustituye la necesidad de definir una variable objetivo con significado.

Importancia de variables en los bosques aleatorios

Los bosques aleatorios pueden proporcionar medidas generales sobre la contribución de las variables. Una de las más utilizadas se basa en la reducción acumulada de impureza producida por cada característica dentro de los árboles. Las variables que generan divisiones capaces de separar mejor las clases reciben valores de importancia más elevados.

Esta medida ofrece una visión global, pero no explica necesariamente una predicción individual. Tampoco debe interpretarse como evidencia de causalidad. Una variable puede ser útil para predecir un resultado sin ser su causa, y su importancia puede cambiar cuando existen características correlacionadas.

Las medidas basadas en impureza pueden favorecer variables continuas o características con numerosos valores posibles. Strobl et al. (2007) identificaron que estas medidas pueden presentar sesgos cuando los predictores difieren en escala o en cantidad de categorías. Por ello, su interpretación debe acompañarse de análisis adicionales.

La importancia por permutación representa una alternativa. Este procedimiento modifica aleatoriamente los valores de una variable y observa cuánto disminuye el desempeño del modelo. Si

la alteración produce una reducción considerable, se considera que la característica contribuía de forma importante a la predicción.

La permutación también presenta limitaciones. Cuando dos variables contienen información semejante, modificar una de ellas puede producir una disminución pequeña porque la otra conserva parte de la capacidad predictiva. Por consiguiente, las medidas de importancia deben utilizarse como instrumentos de inspección y no como conclusiones definitivas sobre el valor institucional de cada criterio.

La diferencia entre importancia global y explicación local debe conservarse. La importancia global resume el comportamiento del modelo en un conjunto de datos. La explicación local busca identificar qué características influyeron en la predicción de una observación específica. Esta segunda dimensión será desarrollada en el apartado sobre inteligencia artificial explicable.

Relación de los modelos predictivos con el prototipo

La selección del bosque aleatorio en el prototipo puede justificarse por su capacidad para trabajar con datos tabulares, representar relaciones no lineales y combinar numerosos árboles. Sin embargo, la elección del algoritmo no demuestra por sí misma que el modelo sea adecuado para la contratación pública.

La validez de la propuesta depende, en primer lugar, de que la variable objetivo represente un resultado con significado. También requiere que los predictores posean una relación justificable con el fenómeno analizado y que los datos reproduzcan, dentro de límites razonables, las condiciones del escenario de aplicación.

El bosque aleatorio debe compararse con modelos más sencillos. Una regresión logística puede funcionar como línea de base y un árbol individual puede aportar una referencia interpretable. Esta comparación permitiría determinar si la complejidad del bosque produce una mejora real en el desempeño.

El modelo también debe evaluarse mediante métricas adecuadas para la distribución de sus clases. La exactitud aislada no es suficiente. Deben presentarse la matriz de confusión, precisión, sensibilidad, especificidad, F1, exactitud balanceada y, cuando resulte pertinente, las áreas bajo las curvas ROC y de precisión-recall.

El prototipo utiliza datos simulados, por lo que sus resultados deben interpretarse como una prueba de funcionamiento del código y de la interfaz. No pueden utilizarse para afirmar que el modelo predice adjudicaciones reales, mejora la eficiencia institucional, reduce el gasto público o selecciona correctamente proveedores.

La inclusión de mecanismos de explicabilidad tampoco corrige por sí sola estas limitaciones. Un sistema puede explicar de manera fiel una predicción generada por un modelo que no aprendió relaciones válidas. Por ello, la explicabilidad debe integrarse después de comprobar que el problema, los datos, la variable objetivo y la evaluación poseen coherencia metodológica.

El apartado siguiente aborda los fundamentos de la inteligencia artificial explicable, las diferencias entre transparencia, interpretabilidad y explicación, así como los principales métodos disponibles para analizar modelos predictivos y comunicar sus resultados a diferentes usuarios.

Inteligencia artificial explicable

La inteligencia artificial explicable comprende un campo de investigación orientado a desarrollar conceptos, métodos y recursos que permitan a las personas comprender el funcionamiento de los sistemas de inteligencia artificial y examinar las razones asociadas con sus resultados. Su importancia ha aumentado con la utilización de modelos capaces de representar relaciones complejas, pero cuyo comportamiento no puede describirse fácilmente mediante reglas observables o coeficientes directamente interpretables.

La explicabilidad no constituye una propiedad única ni posee el mismo significado en todos los contextos. Puede referirse a la comprensión de la estructura general de un modelo, a la

identificación de las variables que influyen en sus predicciones, a la reconstrucción del proceso aplicado a un caso específico o a la comunicación de razones adaptadas a las necesidades de un usuario. Esta diversidad ha generado dificultades para establecer una definición universal y ha llevado a proponer que toda explicación se evalúe de acuerdo con su finalidad, su destinatario y las características de la tarea (Doshi-Velez & Kim, 2017; Lipton, 2018).

En los sistemas de apoyo a decisiones, la explicación cumple una función distinta de la predicción. El modelo predictivo estima una categoría, un valor o una probabilidad; el mecanismo explicativo aporta información destinada a comprender cómo se relaciona esa salida con los datos analizados. La explicación no sustituye el resultado ni demuestra automáticamente que este sea correcto. Su finalidad consiste en facilitar la revisión, detectar posibles inconsistencias y permitir que el usuario valore la pertinencia de la estimación.

La inteligencia artificial explicable tampoco debe entenderse únicamente como un recurso para incrementar la aceptación de la tecnología. Una explicación responsable puede fortalecer la confianza cuando el sistema se comporta de manera coherente, pero también debe proporcionar razones para desconfiar, cuestionar o rechazar una predicción. Si el recurso explicativo se limita a persuadir al usuario, pierde su función crítica y puede contribuir a una dependencia excesiva del modelo.

En consecuencia, la explicabilidad debe integrarse con la evaluación predictiva, la calidad de los datos, la documentación técnica y la supervisión humana. Un sistema puede presentar una explicación comprensible y, al mismo tiempo, utilizar variables inadecuadas, reproducir patrones problemáticos o alcanzar un rendimiento insuficiente. La claridad de una explicación no corrige por sí sola las debilidades del modelo.

Interpretabilidad, explicabilidad y transparencia

Los términos interpretabilidad, explicabilidad y transparencia se utilizan con frecuencia como equivalentes, aunque representan dimensiones relacionadas pero no idénticas. Distinguirlos permite evaluar de manera más precisa qué información ofrece un sistema y qué nivel de comprensión puede alcanzar el usuario.

La interpretabilidad se relaciona con la posibilidad de atribuir un significado comprensible al funcionamiento o a los resultados de un modelo. Un sistema interpretable permite establecer conexiones entre sus entradas, su estructura y sus salidas. Esta comprensión puede derivarse directamente de la forma del modelo o construirse mediante métodos adicionales de análisis.

Murdoch et al. (2019) plantean que una interpretación debe examinarse en función de tres dimensiones complementarias: su capacidad predictiva, la precisión con la que describe lo aprendido por el modelo y la relevancia de la información para una audiencia humana. Esta perspectiva destaca que una explicación técnicamente detallada puede resultar poco útil cuando no responde a las necesidades de quien debe utilizarla.

La explicabilidad se refiere a la capacidad de proporcionar razones o evidencias relacionadas con un resultado. Mientras la interpretabilidad puede describir una propiedad del modelo, la explicación suele materializarse en una salida concreta, como una regla, una comparación, una lista de factores influyentes, un gráfico o una descripción textual.

La transparencia posee un alcance más amplio. Incluye la disponibilidad de información sobre el diseño del sistema, el origen de los datos, las variables utilizadas, el procedimiento de entrenamiento, las métricas, las versiones del modelo y las condiciones en las que puede aplicarse. Un sistema puede ofrecer una explicación de una predicción y continuar siendo poco transparente respecto de su construcción o mantenimiento.

La auditabilidad se relaciona con la posibilidad de revisar y reconstruir el funcionamiento del sistema. Requiere conservar registros sobre los datos procesados, las versiones del algoritmo, las

configuraciones empleadas, las predicciones generadas y las intervenciones humanas. Mientras la explicación responde a preguntas sobre una salida, la auditoría examina el proceso completo y las responsabilidades asociadas.

Estos conceptos no deben evaluarse de manera aislada. Un sistema puede ser transparente en cuanto a su documentación, pero presentar un modelo difícil de interpretar. También puede ofrecer explicaciones locales sin revelar información suficiente sobre la calidad de sus datos. La implementación responsable requiere combinar documentación, interpretación, explicación y trazabilidad.

El problema de los modelos de caja negra

La expresión caja negra se utiliza para describir modelos cuyo comportamiento interno resulta difícil de comprender para las personas, aunque sus datos de entrada y resultados de salida puedan observarse. Esta dificultad puede originarse en la cantidad de parámetros, en la presencia de relaciones no lineales, en la combinación de numerosos componentes o en la imposibilidad de representar el proceso mediante una secuencia breve de reglas.

La condición de caja negra no depende exclusivamente de una familia de algoritmos. Un modelo lineal con centenares de variables y transformaciones puede ser difícil de comprender, mientras que una red compleja puede proporcionar información útil mediante herramientas adecuadas. Por ello, la interpretabilidad no debe atribuirse automáticamente a la denominación del algoritmo, sino evaluarse según la estructura, la presentación y las capacidades del destinatario (Lipton, 2018).

En un bosque aleatorio, la opacidad surge porque la predicción final integra las decisiones de numerosos árboles. Cada árbol puede seguir reglas relativamente comprensibles, pero el conjunto completo no puede resumirse mediante una sola ruta. La explicación requiere examinar cómo las variables influyen en el comportamiento general y en la predicción de cada observación.

La opacidad también puede ser organizacional. Aunque el algoritmo sea conocido, los usuarios pueden desconocer el origen de los datos, la variable objetivo, las transformaciones aplicadas o el criterio utilizado para establecer el umbral. En este caso, la dificultad no se encuentra exclusivamente en la matemática, sino en la ausencia de documentación.

El problema se intensifica cuando las salidas intervienen en decisiones con consecuencias económicas, jurídicas o sociales. En estos contextos, conocer solamente la categoría o la puntuación resulta insuficiente. Es necesario examinar qué información influyó, si las variables son pertinentes, si existen errores y si el sistema operó dentro de las condiciones para las que fue diseñado.

Rudin (2019) advierte que, en decisiones de alto impacto, debe considerarse prioritariamente el uso de modelos inherentemente interpretables antes de recurrir a sistemas opacos acompañados de explicaciones posteriores. Esta posición se fundamenta en que una explicación aproximada puede no reproducir con fidelidad el comportamiento de la caja negra.

Este planteamiento no implica que todos los modelos complejos deban descartarse. Exige justificar por qué su utilización resulta necesaria y comprobar si la mejora predictiva obtenida compensa la pérdida de interpretabilidad. Si un modelo sencillo alcanza un comportamiento semejante, puede ofrecer una alternativa más adecuada para contextos en los que la comprensión y la rendición de cuentas son esenciales.

En el prototipo de esta obra, la elección del bosque aleatorio debe evaluarse frente a modelos más simples. La comparación con una regresión logística y un árbol de decisión permitiría determinar si existe una mejora predictiva significativa y si esa diferencia justifica la incorporación de métodos adicionales de explicación.

Principios de una explicación adecuada

Una explicación útil no se limita a mostrar información relacionada con la predicción. Debe cumplir condiciones que permitan determinar si esa información representa correctamente al sistema y si puede ser comprendida por su destinatario.

Phillips et al. (2021), en el marco desarrollado por el Instituto Nacional de Estándares y Tecnología de Estados Unidos, proponen cuatro principios generales para los sistemas explicables: ofrecer evidencia o razones para sus resultados, producir explicaciones comprensibles para el destinatario, asegurar que la explicación represente correctamente el proceso del sistema y reconocer los límites dentro de los cuales la herramienta puede operar adecuadamente.

El primer principio exige que el sistema no entregue únicamente una categoría, probabilidad o recomendación. Debe acompañar la salida con información que permita relacionarla con los factores considerados. En un caso aplicado a proveedores, podría mostrar qué características aumentaron o redujeron una estimación.

El segundo principio establece que la explicación debe adaptarse al usuario. Una presentación destinada a un desarrollador puede incluir parámetros, distribuciones y medidas estadísticas. Una explicación dirigida a un funcionario debería priorizar las variables relevantes, la dirección de sus efectos, las limitaciones y las advertencias necesarias.

El tercer principio se relaciona con la fidelidad. La explicación debe representar el funcionamiento real del modelo y no ofrecer razones meramente plausibles. Una descripción puede parecer razonable para el usuario y, sin embargo, no corresponder con los factores que determinaron la predicción. En tal caso, funcionaría como una justificación aparente y no como una explicación fiel.

El cuarto principio exige reconocer los límites del sistema. El modelo debe evitar producir o presentar como confiables resultados correspondientes a condiciones muy diferentes de aquellas

consideradas durante su desarrollo. También debe comunicar cuando los datos son insuficientes, cuando la predicción se encuentra cerca del umbral o cuando existe un nivel elevado de incertidumbre.

Estos principios muestran que la explicación posee dimensiones técnicas y comunicativas. No basta con calcular contribuciones matemáticas; también es necesario presentarlas de forma comprensible y acompañarlas de información sobre el alcance y las restricciones del modelo.

Modelos interpretables y explicaciones posteriores

Los métodos de inteligencia artificial explicable pueden organizarse en dos grandes grupos. El primero comprende modelos diseñados para que su funcionamiento sea interpretable. El segundo incluye métodos que explican, después del entrenamiento, las predicciones de modelos que no son directamente comprensibles.

Los modelos intrínsecamente interpretables integran la comprensión dentro de su propia estructura. Entre ellos pueden encontrarse regresiones con una cantidad controlada de variables, sistemas de puntuación, árboles de profundidad limitada, listas de reglas y determinados modelos aditivos. En estos casos, la predicción y la explicación proceden esencialmente del mismo mecanismo.

La interpretabilidad intrínseca depende de la complejidad real del modelo. Un árbol pequeño puede ser examinado con facilidad, mientras que uno con centenares de nodos deja de ser comprensible. Del mismo modo, una regresión con pocas variables puede interpretarse directamente, pero una estructura con numerosas interacciones y transformaciones puede resultar difícil de analizar.

Las explicaciones post hoc se generan después de que el modelo ha sido entrenado. Su función consiste en analizar la relación entre entradas y salidas sin modificar necesariamente el algoritmo

original. Este enfoque permite aplicar herramientas de explicación a bosques aleatorios, máquinas de vectores de soporte, redes neuronales y otros modelos complejos.

La ventaja de los métodos posteriores consiste en que pueden incorporarse a sistemas ya desarrollados y, en ciertos casos, utilizarse con diferentes familias de algoritmos. Su principal limitación es que la explicación puede constituir una aproximación y no una reproducción exacta del razonamiento interno.

Doshi-Velez y Kim (2017) sostienen que la evaluación de la interpretabilidad debe realizarse con procedimientos rigurosos y adaptados al contexto. No resulta suficiente afirmar que una representación es interpretable por contener pocas variables o adoptar una forma visual. Debe comprobarse si permite a los usuarios realizar correctamente la tarea para la cual fue construida.

La elección entre un modelo interpretable y una explicación posterior debe responder al riesgo, al rendimiento y a la finalidad. En un escenario demostrativo puede justificarse el uso de una caja negra para explorar técnicas de explicación. En una implementación institucional con consecuencias relevantes, debe analizarse primero si un modelo transparente puede alcanzar un desempeño adecuado.

Explicaciones globales y locales

Las explicaciones también pueden clasificarse según su alcance. Las globales buscan describir el comportamiento general del modelo, mientras que las locales se concentran en una predicción específica.

Una explicación global intenta responder preguntas como: ¿qué variables utiliza con mayor frecuencia el modelo?, ¿qué relaciones generales ha aprendido?, ¿cómo cambia la predicción cuando se modifica una característica?, ¿qué reglas predominan en el conjunto de datos? Estas explicaciones permiten examinar si el comportamiento global coincide con el conocimiento técnico y con la finalidad declarada.

Entre los recursos globales se encuentran las medidas de importancia de variables, los gráficos de dependencia parcial, los modelos sustitutos y determinadas representaciones de reglas. Su utilidad radica en ofrecer una visión general, aunque pueden ocultar diferencias entre observaciones individuales.

Una explicación local busca responder por qué el sistema produjo un resultado para un caso determinado. Puede mostrar las características que aumentaron la probabilidad, aquellas que la redujeron y la magnitud aproximada de cada contribución. Este tipo de explicación resulta especialmente relevante cuando el usuario debe revisar una estimación individual.

Las explicaciones globales y locales no son intercambiables. Una variable puede ser importante para el comportamiento general y tener poca influencia en un caso específico. De manera inversa, una característica poco relevante en promedio puede ser determinante para una observación particular.

En el prototipo, la importancia general de las variables podría utilizarse para examinar el comportamiento agregado del bosque aleatorio. Para comparar dos proveedores concretos, sería necesaria una explicación local que identifique la contribución de las características en cada predicción.

Métodos específicos y métodos independientes del modelo

Los métodos específicos se diseñan para una familia particular de algoritmos y aprovechan su estructura interna. En un árbol de decisión, por ejemplo, la ruta seguida desde la raíz hasta la hoja puede utilizarse como explicación. En una regresión, los coeficientes permiten examinar la dirección y magnitud de las relaciones bajo los supuestos del modelo.

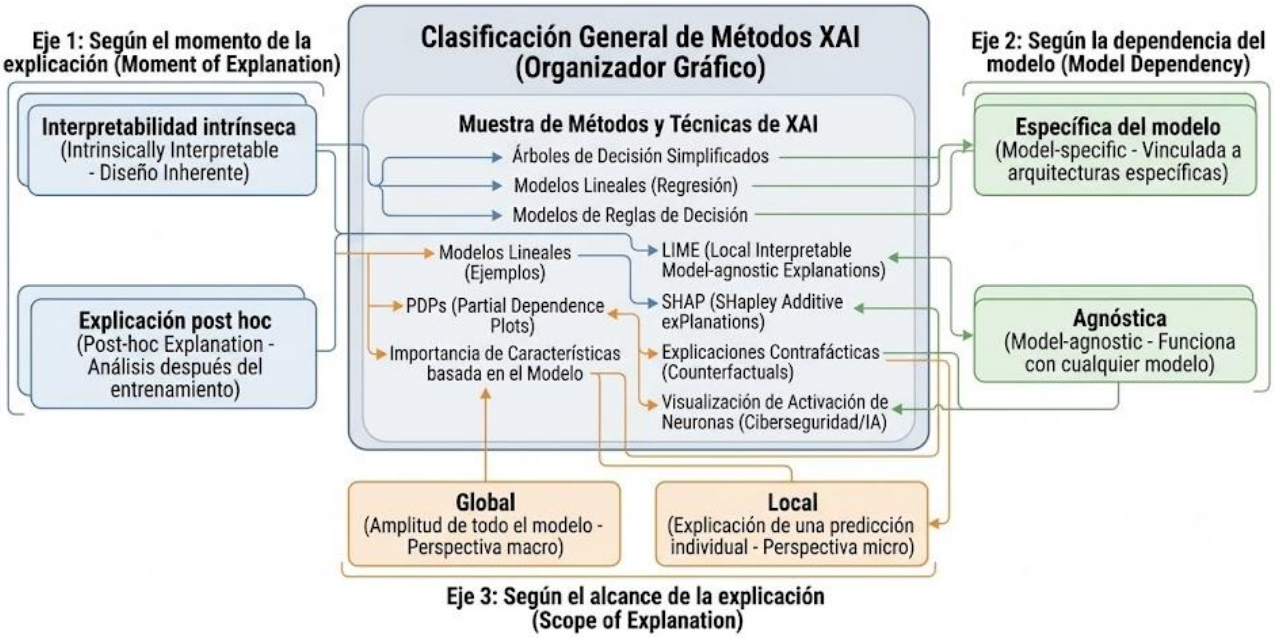
Los métodos independientes o agnósticos analizan el comportamiento del sistema principalmente a través de sus entradas y salidas. Pueden aplicarse a diferentes algoritmos siempre que sea posible generar predicciones. Esta flexibilidad resulta útil para comparar modelos o explicar sistemas cuya estructura interna no está disponible.

La condición de agnóstico no elimina las limitaciones. Estos métodos pueden requerir numerosas consultas al modelo, depender de perturbaciones artificiales de los datos o producir aproximaciones que varían según la configuración. La flexibilidad debe acompañarse de análisis de estabilidad y fidelidad.

Barredo Arrieta et al. (2020) organizan las técnicas de XAI mediante criterios como el momento de aplicación, el alcance de la explicación, la dependencia respecto del modelo y el tipo de resultado producido. Esta taxonomía permite seleccionar el método según la pregunta que se desea responder y no únicamente por su popularidad.

Figura 4

Clasificación general de los métodos de inteligencia artificial explicable



Nota. *Elaboración propia basada en Barredo Arrieta et al. (2020) y Doshi-Velez y Kim (2017).*

Importancia por permutación

La importancia por permutación evalúa cuánto disminuye el rendimiento del modelo cuando se altera aleatoriamente una variable. El procedimiento mantiene las demás características y rompe la

relación entre el predictor seleccionado y la variable objetivo. Si el desempeño se reduce de manera considerable, se interpreta que la característica aportaba información relevante.

A diferencia de la importancia basada en la reducción de impureza, la permutación puede aplicarse a diferentes modelos y calcularse sobre datos que no fueron utilizados durante el entrenamiento. Esta característica permite examinar la contribución de las variables en condiciones más próximas a la evaluación.

La importancia por permutación ofrece una explicación global. Indica cuánto depende el modelo de una variable dentro de un conjunto de observaciones, pero no explica necesariamente la predicción de un caso particular. Tampoco establece causalidad ni demuestra que la característica sea jurídicamente pertinente.

Cuando existen variables correlacionadas, la interpretación puede complicarse. Si dos predictores contienen información semejante, modificar uno puede producir una reducción limitada porque el modelo conserva parte de la información a través del otro. Por ello, los valores bajos no siempre significan que la variable carezca de relevancia.

En el prototipo, la importancia por permutación podría complementar la medida interna del bosque aleatorio. La comparación entre ambas permitiría identificar variables cuya importancia depende del procedimiento utilizado y detectar resultados que requieren una revisión adicional.

Dependencia parcial y expectativa condicional individual

Los gráficos de dependencia parcial muestran cómo cambia, en promedio, la predicción del modelo cuando se modifica una variable y se mantienen las demás características según los valores observados. Su finalidad consiste en representar la relación que el modelo ha aprendido entre una característica y la salida.

Este recurso puede revelar tendencias no lineales, umbrales o zonas en las que la variable produce cambios relevantes. Sin embargo, el resultado corresponde a un promedio y puede ocultar comportamientos diferentes entre observaciones.

Los gráficos de expectativa condicional individual, conocidos como ICE, representan la relación para cada observación por separado. En lugar de mostrar una única curva promedio, generan varias trayectorias que permiten identificar heterogeneidad e interacciones. Goldstein et al. (2015) desarrollaron este procedimiento para observar diferencias que los gráficos de dependencia parcial pueden ocultar.

Estos gráficos requieren precaución cuando las variables se encuentran fuertemente correlacionadas. Modificar una característica sin alterar las demás puede generar combinaciones poco plausibles o inexistentes en los datos reales. La visualización debe interpretarse como una descripción del comportamiento del modelo y no necesariamente como un efecto causal.

En la propuesta del libro, la dependencia parcial podría emplearse para analizar cómo el modelo relaciona variables como experiencia, cumplimiento o tiempo de respuesta con la probabilidad estimada. No obstante, la interpretación debe verificar que los valores representados sean coherentes con el escenario y que la dirección del criterio haya sido definida correctamente.

LIME

LIME, acrónimo de Local Interpretable Model-Agnostic Explanations, es un método destinado a explicar una predicción individual mediante la construcción de un modelo interpretable alrededor del caso analizado. El procedimiento genera variaciones próximas a la observación, consulta las predicciones de la caja negra y ajusta una representación sencilla que aproxima localmente su comportamiento (Ribeiro et al., 2016).

La lógica del método se fundamenta en que un modelo complejo puede comportarse de manera relativamente sencilla dentro de una región limitada, aunque su estructura global sea difícil de

interpretar. La explicación muestra qué características favorecieron o redujeron la predicción dentro de esa zona.

LIME es independiente del algoritmo y puede utilizarse con clasificadores diversos. Su flexibilidad ha favorecido su aplicación en datos tabulares, texto e imágenes. Sin embargo, la explicación puede variar según la forma en que se generan las perturbaciones, la medida de proximidad, la cantidad de observaciones sintéticas y la configuración del modelo local.

La estabilidad constituye una consideración importante. Dos ejecuciones sobre el mismo caso podrían producir explicaciones distintas cuando existe aleatoriedad o cuando pequeños cambios modifican la aproximación local. Por ello, no debe presentarse una única explicación sin examinar su consistencia.

En datos tabulares, las observaciones sintéticas deben respetar las características del dominio. Generar combinaciones imposibles entre variables puede producir una explicación basada en casos que no podrían ocurrir. La configuración del método necesita incorporar restricciones y conocimientos del contexto.

LIME podría utilizarse para explicar la estimación asignada a un proveedor específico. La salida podría mostrar qué características aumentaron o redujeron localmente la probabilidad. Sin embargo, su incorporación exigiría evaluar la estabilidad y fidelidad de la aproximación.

SHAP

SHAP, acrónimo de SHapley Additive exPlanations, constituye un marco para atribuir a cada característica una contribución dentro de una predicción. Se fundamenta en los valores de Shapley, desarrollados originalmente en teoría de juegos para distribuir de forma equitativa el resultado de una coalición entre sus participantes.

En el contexto del aprendizaje automático, las variables se interpretan como participantes que contribuyen a desplazar la predicción desde un valor de referencia hasta el resultado obtenido. Cada valor SHAP indica la dirección y magnitud de la contribución de una característica para un caso específico (Lundberg & Lee, 2017).

Una contribución positiva aumenta la salida respecto del valor de referencia, mientras que una contribución negativa la reduce. La suma de las contribuciones y el valor base permite reconstruir la predicción bajo el marco aditivo del método.

SHAP puede producir explicaciones locales y agregarse para construir análisis globales. Las explicaciones individuales muestran la contribución de las variables en un caso. La agregación de múltiples observaciones permite identificar patrones generales, distribuciones de efectos e interacciones potenciales.

Para los modelos basados en árboles existen procedimientos especializados que permiten calcular los valores con mayor eficiencia. Esta característica convierte a SHAP en una alternativa pertinente para bosques aleatorios.

A pesar de sus ventajas, los valores SHAP no representan efectos causales. Muestran cómo el modelo utiliza las variables y no cómo una modificación real produciría necesariamente un cambio en el fenómeno estudiado. También dependen del valor de referencia y de los supuestos empleados para representar la ausencia de una característica.

Las variables correlacionadas pueden dificultar la distribución de las contribuciones. Cuando dos características contienen información semejante, la asignación entre ellas puede variar según el procedimiento y el conjunto de referencia. Los resultados necesitan interpretarse junto con el conocimiento del dominio.

En el prototipo, SHAP podría emplearse para explicar por qué el bosque aleatorio asignó una probabilidad determinada a cada proveedor. Una visualización local permitiría diferenciar

variables que aumentaron y redujeron la estimación. Las visualizaciones globales podrían mostrar qué características influyeron con mayor frecuencia en el conjunto de casos.

Explicaciones contrafactuales

Las explicaciones contrafactuales responden a una pregunta diferente de la importancia de variables: ¿qué cambios mínimos en las características habrían producido un resultado distinto? En lugar de describir únicamente por qué se obtuvo una predicción, muestran una situación alternativa cercana en la que la salida del modelo cambia.

Una explicación podría indicar, por ejemplo, que la estimación habría superado un umbral si una característica modificable hubiese alcanzado determinado valor. Este tipo de presentación puede ser más comprensible para algunos usuarios porque relaciona el resultado con condiciones concretas.

No todas las variables pueden modificarse libremente. Algunas representan antecedentes históricos, características estructurales o datos que no deben alterarse. Una explicación contrafactual responsable debe respetar restricciones, mantener combinaciones plausibles y diferenciar factores modificables de aquellos que no lo son.

También debe evitarse presentar el contrafactual como una garantía. Alcanzar los valores señalados no asegura que el resultado real cambie, porque la explicación describe el comportamiento del modelo y no una relación causal comprobada.

En contratación pública, los contrafactuales requieren especial cautela. No deberían convertirse en instrucciones para adaptar artificialmente una oferta con el único propósito de superar el algoritmo. Su utilidad potencial se encuentra en la revisión interna del modelo y en la comprensión de los umbrales aprendidos.

Evaluación de la calidad de las explicaciones

La existencia de una explicación no demuestra su calidad. Los métodos deben evaluarse según criterios relacionados con fidelidad, estabilidad, consistencia, comprensibilidad, relevancia y utilidad.

La fidelidad expresa el grado en que la explicación representa correctamente el comportamiento del modelo. Una explicación sencilla puede resultar clara y, al mismo tiempo, omitir relaciones esenciales. La evaluación debe determinar cuánto se pierde al simplificar.

La estabilidad se relaciona con la semejanza entre las explicaciones producidas para observaciones cercanas o entre ejecuciones repetidas. Cambios pequeños en los datos no deberían generar razones completamente diferentes, salvo que el modelo presente una frontera de decisión muy próxima.

La consistencia permite examinar si dos modelos que utilizan una variable de manera semejante generan explicaciones comparables. También puede referirse a la aplicación uniforme de criterios a casos equivalentes.

La comprensibilidad depende del destinatario. Una explicación con numerosos valores y gráficos puede resultar informativa para un especialista y producir sobrecarga para otro usuario. La evaluación debe incorporar personas representativas del contexto de aplicación.

La relevancia exige que la explicación responda a la pregunta real del usuario. Un funcionario puede necesitar conocer por qué se produjo la estimación, mientras que un desarrollador puede requerir información sobre errores o inestabilidad. Una misma salida no satisface necesariamente ambos objetivos.

Doshi-Velez y Kim (2017) distinguen evaluaciones basadas en aplicaciones reales, estudios con personas y análisis funcionales sin usuarios. Esta clasificación muestra que la calidad no puede

determinarse únicamente mediante una propiedad matemática y que, en determinados contextos, debe comprobarse cómo las personas utilizan la explicación.

Explicabilidad centrada en el usuario

La explicación debe diseñarse a partir de las responsabilidades y conocimientos de sus destinatarios. Un sistema explicable no necesita mostrar la misma información a todas las personas.

Los desarrolladores requieren detalles relacionados con variables, parámetros, distribución de datos, estabilidad, errores y comportamiento del algoritmo. Su finalidad principal consiste en diagnosticar y mejorar el modelo.

Los analistas técnicos necesitan examinar cómo se relaciona la predicción con la información del caso. Pueden requerir comparaciones, contribuciones locales, valores de referencia y advertencias sobre datos atípicos.

Los responsables institucionales necesitan comprender los factores principales, el grado de incertidumbre y las limitaciones. La explicación debe permitirles integrar la estimación con evidencias externas y criterios normativos.

Los auditores requieren trazabilidad. Deben poder identificar el origen de los datos, la versión del modelo, las transformaciones, el resultado y la intervención humana.

Los proveedores o participantes externos podrían requerir explicaciones comprensibles sobre los criterios generales utilizados. Esta información debe equilibrarse con la protección de datos, la seguridad del sistema y los derechos de terceros.

Murdoch et al. (2019) destacan que la relevancia de una interpretación depende de la audiencia. Esto implica que la explicación no puede evaluarse sin definir quién la utilizará y qué tarea necesita realizar.

Aplicación de la XAI en el prototipo

La capa explicativa del prototipo debe construirse sobre un modelo previamente evaluado. Antes de explicar una predicción, es necesario comprobar que el sistema aprendió relaciones válidas y que su comportamiento supera una línea de base razonable.

La primera explicación debería mostrar la probabilidad estimada y advertir que corresponde a un escenario simulado. La segunda debería identificar las variables que contribuyeron positiva y negativamente. La tercera podría comparar las contribuciones de dos proveedores sin asumir que el valor más alto siempre representa una condición favorable.

Las variables deben orientarse según su significado. Una mayor experiencia puede interpretarse como favorable dentro del escenario, mientras que un costo o tiempo de respuesta más alto podría reducir la estimación. La interfaz debe comunicar esta dirección y evitar comparaciones basadas únicamente en la magnitud numérica.

La importancia general de variables puede utilizarse para examinar el modelo completo. La importancia por permutación permitiría comprobar cuánto depende el rendimiento de cada característica. SHAP podría emplearse para explicar casos individuales y construir una síntesis global.

LIME podría incorporarse como método complementario, aunque sería necesario evaluar su estabilidad. No resulta conveniente utilizar simultáneamente numerosos métodos sin explicar sus diferencias. La selección debe responder a las preguntas concretas del usuario.

La interfaz también debe comunicar limitaciones. Debe indicar que los datos son simulados, que la predicción no representa una adjudicación, que las variables no corresponden necesariamente a criterios oficiales y que la salida requiere revisión humana.

La explicación no debe afirmar que un proveedor es “el mejor”. Es preferible indicar que, bajo las variables y el modelo empleados, una alternativa recibió una estimación superior y que determinados factores contribuyeron a esa diferencia.

La supervisión humana debe quedar registrada. El usuario debería poder aceptar el resultado como información de apoyo, solicitar una revisión o señalar que la recomendación no resulta aplicable debido a factores no considerados por el modelo.

La inteligencia artificial explicable proporciona herramientas para examinar las predicciones y comunicar los factores que influyen en ellas. Sin embargo, el prototipo no se limita a clasificar observaciones, sino que utiliza las estimaciones para comparar y priorizar alternativas. Por ello, resulta necesario analizar los fundamentos de los sistemas de recomendación, sus componentes, sus principales enfoques y las diferencias entre predecir, puntuar, ordenar y recomendar.

Fundamentos de los sistemas de recomendación

Los sistemas de recomendación constituyen herramientas de filtrado y apoyo a decisiones diseñadas para identificar, ordenar o presentar alternativas que pueden resultar pertinentes para un usuario, una organización o una situación determinada. Su desarrollo responde a la dificultad de examinar manualmente grandes conjuntos de opciones y seleccionar aquellas que guardan mayor relación con una necesidad, preferencia o criterio previamente establecido.

Resnick y Varian (1997) describieron estos sistemas como mecanismos que utilizan las opiniones y experiencias de otras personas para orientar la selección de información, productos o servicios. Esta formulación temprana permitió reconocer que la recomendación no consiste únicamente en recuperar información disponible, sino en procesarla con el propósito de reducir el conjunto de alternativas que el usuario necesita revisar.

Con el desarrollo del campo, los sistemas de recomendación incorporaron diferentes fuentes de información, métodos estadísticos y algoritmos de aprendizaje automático. Las recomendaciones

pueden basarse en las características de los objetos, en las interacciones previas de los usuarios, en las preferencias de personas semejantes, en reglas de conocimiento o en combinaciones de estas estrategias. La tercera edición del Recommender Systems Handbook organiza el campo alrededor de métodos colaborativos, contenido, retroalimentación implícita, contexto, redes neuronales, evaluación, impacto, explicaciones y apoyo a decisiones individuales y grupales (Ricci et al., 2022).

El propósito de un sistema de recomendación no es necesariamente determinar una única opción correcta. En muchos casos, su función consiste en reducir el espacio de búsqueda, asignar puntuaciones de relevancia y presentar un conjunto ordenado de alternativas. La utilidad de la recomendación depende de la relación entre los datos disponibles, la finalidad de la herramienta, las necesidades del usuario y el contexto en el que se utilizará.

Esta precisión resulta especialmente importante en la contratación pública. Una herramienta puede ordenar proveedores, identificar oportunidades compatibles con determinadas características o mostrar alternativas que requieren revisión. Sin embargo, la recomendación no sustituye la evaluación jurídica, técnica y económica correspondiente, ni posee por sí misma la capacidad de emitir una decisión administrativa.

Recomendación y sobrecarga de información

La sobrecarga de información se presenta cuando la cantidad de alternativas disponibles supera la capacidad práctica del usuario para revisarlas, compararlas y seleccionar una opción. El problema no radica únicamente en la cantidad de datos, sino en la dificultad para determinar qué información resulta pertinente y cómo debe relacionarse con la necesidad existente.

Los sistemas de búsqueda tradicionales suelen recuperar elementos que coinciden con una consulta expresada por el usuario. Los sistemas de recomendación incorporan una etapa adicional, porque intentan estimar la relevancia o utilidad de las alternativas a partir de información previa. Esta

diferencia permite generar sugerencias incluso cuando el usuario no formula una consulta completamente detallada.

La recomendación puede disminuir el esfuerzo necesario para explorar un catálogo, pero también puede limitar la visibilidad de ciertas opciones. Si el sistema prioriza sistemáticamente los elementos más conocidos, más evaluados o semejantes a elecciones anteriores, puede reducir la diversidad y dificultar el descubrimiento de alternativas nuevas. Por esta razón, la función de filtrado debe examinarse junto con criterios como cobertura, novedad, diversidad y equidad.

En el contexto institucional, la sobrecarga no se presenta únicamente por la cantidad de proveedores. También puede originarse en el volumen de documentos, requisitos, especificaciones, antecedentes, certificaciones y variables que deben examinarse. Una herramienta de apoyo podría contribuir a organizar esta información, pero el procesamiento automatizado no elimina la necesidad de comprobar la calidad y pertinencia de cada dato.

Usuario, objeto e interacción

Los sistemas de recomendación se estructuran tradicionalmente a partir de tres entidades fundamentales: los usuarios, los objetos y las interacciones. Esta representación permite organizar la información y establecer la tarea que debe realizar el sistema.

El usuario es la persona o entidad para la cual se genera la recomendación. En aplicaciones comerciales puede ser un consumidor que busca un producto, una película o una canción. En un contexto institucional, el usuario puede ser un analista, una unidad administrativa, una entidad contratante o un proveedor interesado en identificar oportunidades compatibles con su actividad.

El objeto, también denominado ítem, corresponde a la alternativa que puede ser recomendada. Dependiendo del dominio, puede tratarse de un producto, servicio, documento, contenido, proveedor, oferta o procedimiento de contratación. La correcta definición del objeto resulta

esencial porque determina qué características serán almacenadas y qué significa que una recomendación sea relevante.

La interacción representa una relación observable entre el usuario y el objeto. Puede expresarse mediante una valoración, selección, consulta, compra, lectura, clic, tiempo de permanencia o cualquier otra acción que proporcione información sobre el posible interés del usuario.

El modelo clásico usuario-objeto resulta útil para numerosas aplicaciones, pero no describe completamente todos los contextos. Una recomendación también puede depender del momento, la ubicación, el dispositivo, la finalidad de la consulta, las restricciones o las características de otros participantes. Estos factores adicionales conforman el contexto de la recomendación.

En contratación pública, la identificación de las entidades requiere una delimitación específica. Si la herramienta recomienda procedimientos a los proveedores, el proveedor funciona como usuario y la oportunidad contractual como objeto. Si el sistema organiza proveedores para una entidad, la entidad o el analista constituye el usuario y los proveedores u ofertas representan los objetos. Ambas configuraciones responden a finalidades diferentes y no deben confundirse.

El prototipo desarrollado en el manuscrito compara características de proveedores y estima una probabilidad asociada con cada caso. Por ello, su estructura se aproxima a un sistema de puntuación predictiva utilizado para ordenar alternativas. No corresponde plenamente a un recomendador clásico basado en una matriz de usuarios y objetos.

Función de utilidad y relevancia

La función de utilidad representa el criterio mediante el cual el sistema estima el grado en que un objeto puede resultar pertinente para un usuario. Esta utilidad puede derivarse de una valoración prevista, una probabilidad, una medida de similitud, el cumplimiento de restricciones o una combinación de diferentes factores.

Adomavicius y Tuzhilin (2005) plantean el problema de recomendación como la estimación de la utilidad de los objetos para los usuarios y la selección de aquellos que alcanzan los valores más elevados. Esta formulación admite múltiples métodos, pero exige definir con claridad qué representa la utilidad dentro del dominio analizado.

En un sistema de entretenimiento, la utilidad puede representar la probabilidad de que una persona disfrute una película. En una plataforma de comercio, podría aproximarse mediante la probabilidad de compra. En una herramienta institucional, la utilidad debe construirse a partir de criterios técnicos y normativos explícitos, no de preferencias informales.

Una probabilidad histórica de adjudicación no puede interpretarse automáticamente como utilidad institucional. Un proveedor puede haber obtenido numerosos contratos debido a circunstancias, requisitos o períodos específicos que no se aplican al procedimiento actual. Convertir directamente esa frecuencia en recomendación podría reproducir patrones anteriores sin comprobar la pertinencia de la alternativa.

La función de utilidad también puede incorporar varios criterios. Un sistema podría considerar experiencia, cumplimiento, capacidad operativa, costo y tiempo de respuesta. Sin embargo, cada variable necesita una definición, una dirección y una ponderación justificadas. La suma de valores heterogéneos sin un procedimiento metodológico puede producir una puntuación numérica cuya apariencia de precisión no corresponde con su fundamento.

En la presente obra, el término utilidad debe emplearse de forma general. Las variables simuladas permiten construir una demostración computacional, pero no constituyen una función oficial para evaluar proveedores ni pueden emplearse directamente en una adjudicación real.

Datos de entrada

Los sistemas de recomendación requieren información que permita representar a los usuarios, los objetos y sus relaciones. La naturaleza de estos datos depende del método utilizado y de la tarea que se desea resolver.

Entre las principales fuentes de información se encuentran:

- Valoraciones explícitas.
- Interacciones implícitas.
- Características de los usuarios.
- Características de los objetos.
- Información contextual.
- Reglas, restricciones y conocimiento del dominio.

Las valoraciones explícitas son manifestaciones directas de preferencia. Pueden expresarse mediante escalas numéricas, elecciones binarias, listas de intereses o respuestas proporcionadas en formularios. Su ventaja es que reflejan una declaración consciente del usuario. Sin embargo, requieren participación activa y pueden verse afectadas por escalas interpretadas de manera diferente.

La retroalimentación implícita se obtiene a partir del comportamiento. Acciones como consultar, seleccionar, adquirir, descargar o permanecer durante cierto tiempo en un contenido pueden funcionar como señales indirectas. Estas interacciones suelen estar disponibles en mayor cantidad, pero su significado es ambiguo. La ausencia de una acción no implica necesariamente rechazo, y una consulta no demuestra siempre preferencia.

Hu et al. (2008) diferencian la retroalimentación implícita de las valoraciones explícitas porque la primera expresa principalmente niveles de confianza derivados de comportamientos observados,

no evaluaciones directas de agrado. Por ello, el modelo debe distinguir entre preferencia declarada e interacción registrada.

Las características del usuario describen atributos relevantes para la personalización. Pueden incluir necesidades, conocimientos, experiencia, restricciones o información demográfica. Su utilización debe limitarse a datos pertinentes y respetar principios de privacidad, proporcionalidad y no discriminación.

Las características de los objetos permiten representar su contenido o propiedades. Un documento puede describirse mediante términos; un servicio, mediante funciones y condiciones; una oportunidad contractual, mediante su objeto, presupuesto, ubicación, plazo y requisitos; un proveedor, mediante atributos técnicos y operativos.

La información contextual incorpora condiciones bajo las cuales ocurre la interacción. El mismo usuario puede considerar pertinente una alternativa en una situación y descartarla en otra. El contexto puede incluir tiempo, ubicación, dispositivo, finalidad, restricciones presupuestarias o fase del procedimiento.

Las reglas y el conocimiento del dominio permiten establecer relaciones que no dependen exclusivamente de patrones estadísticos. En ámbitos regulados, ciertas condiciones deben cumplirse obligatoriamente y no pueden sustituirse por una puntuación elevada en otra variable. Un proveedor que incumple una condición habilitante no debería compensar esa limitación mediante mayor experiencia o menor costo.

Tabla 5

Principales fuentes de información en los sistemas de recomendación

Fuente de información	*Descripción*	*Ejemplos*	*Principal precaución*
Valoración explícita	Preferencia declarada directamente por el usuario	Puntuación, selección, interés o rechazo	Requiere participación y puede contener respuestas inconsistentes

Interacción implícita	Señal obtenida mediante el comportamiento registrado	Clic, consulta, descarga, compra o tiempo de uso	La acción observada no equivale necesariamente a preferencia
Perfil del usuario	Características o necesidades de la persona o entidad	Intereses, experiencia, actividad o restricciones	Riesgo de utilizar datos irrelevantes, sensibles o discriminatorios
Características del objeto	Atributos descriptivos de la alternativa	Categoría, contenido, precio, ubicación o requisitos	Depende de la calidad y estandarización de la descripción
Información contextual	Condiciones presentes durante la interacción	Tiempo, ubicación, dispositivo, finalidad o presupuesto	Aumenta la complejidad y puede producir mayor dispersión de datos
Reglas y conocimiento	Criterios definidos por especialistas o por el dominio	Restricciones, compatibilidad y requisitos obligatorios	Requiere actualización y validación especializada

Nota. *Elaboración propia basada en Adomavicius y Tuzhilin (2005), Hu et al. (2008) y Ricci et al. (2022).*

Valoraciones explícitas y comportamiento implícito

La distinción entre información explícita e implícita afecta directamente la interpretación de las recomendaciones. Una valoración explícita indica que el usuario expresó una opinión mediante el mecanismo proporcionado por el sistema. Una interacción implícita únicamente demuestra que ocurrió una acción.

Por ejemplo, consultar una oportunidad contractual no significa necesariamente que el proveedor la considere adecuada. La consulta pudo realizarse para verificar un requisito, observar a la competencia o descartar la participación. Del mismo modo, no consultar una oportunidad puede deberse a falta de conocimiento y no a desinterés.

Los sistemas basados en información implícita necesitan modelar la intensidad y confiabilidad de las señales. Varias consultas, descargas de documentos o acciones posteriores pueden aportar más evidencia que una interacción aislada. Sin embargo, ninguna combinación elimina completamente la ambigüedad.

En el ámbito público, el comportamiento histórico también debe interpretarse con cuidado. La participación frecuente de un proveedor puede reflejar capacidad y experiencia, pero también mayor conocimiento del sistema, disponibilidad de recursos administrativos o concentración en determinadas categorías. El modelo no debe convertir automáticamente la frecuencia en un indicador de superioridad.

Cuando sea posible, resulta conveniente combinar señales explícitas e implícitas. Las preferencias declaradas pueden aportar intención, mientras que las interacciones permiten observar comportamiento. Esta integración requiere resolver posibles contradicciones y documentar el significado asignado a cada fuente.

Perfil del usuario

El perfil del usuario reúne la información que el sistema utiliza para representar sus necesidades, intereses o características. Puede construirse de manera explícita, mediante formularios y configuraciones, o inferirse a partir de interacciones anteriores.

Un perfil estático contiene información que cambia con poca frecuencia, como actividad principal, ubicación o categoría institucional. Un perfil dinámico incorpora señales que evolucionan con el tiempo, como consultas recientes, cambios en las preferencias o nuevas necesidades.

La inferencia automática del perfil puede mejorar la personalización, pero también generar representaciones incompletas. El usuario puede quedar asociado con intereses anteriores que ya no resultan pertinentes. Por ello, debería existir la posibilidad de revisar, actualizar o corregir la información.

En una herramienta dirigida a proveedores, el perfil podría incorporar actividad económica, experiencia, ubicación, capacidad operativa y tipos de contratos de interés. La recomendación consistiría en identificar oportunidades compatibles con esas características.

En una herramienta dirigida a entidades contratantes, el concepto de perfil debería sustituirse o complementarse con la descripción de la necesidad institucional. Esta descripción incluiría objeto contractual, condiciones técnicas, restricciones, presupuesto, plazo y criterios aplicables.

El perfil no debe incluir información simplemente porque se encuentra disponible. Cada atributo necesita demostrar su relevancia para la finalidad de la recomendación. Las variables sensibles o indirectamente relacionadas con características protegidas requieren un análisis reforzado para evitar efectos discriminatorios.

Representación de los objetos

Los objetos deben transformarse en una representación que permita compararlos o relacionarlos con los usuarios. Esta representación puede construirse mediante atributos estructurados, textos, categorías, vectores o factores latentes.

Los atributos estructurados incluyen campos definidos, como precio, ubicación, duración o categoría. Su ventaja es la facilidad de comparación, siempre que se utilicen definiciones y unidades uniformes.

La información textual puede procesarse para identificar términos, temas o representaciones semánticas. Esta posibilidad resulta relevante en contratación pública porque una parte importante de la información se encuentra en descripciones, especificaciones y documentos.

Las representaciones latentes son construidas automáticamente por el modelo y pueden reflejar relaciones que no aparecen como variables observables. Aunque permiten capturar patrones complejos, su significado puede ser difícil de interpretar.

La calidad de la representación determina qué semejanzas puede reconocer el sistema. Si dos oportunidades se describen mediante textos distintos pero corresponden a necesidades semejantes,

una representación exclusivamente literal podría no relacionarlas. En cambio, una representación demasiado general puede considerar equivalentes objetos que presentan diferencias esenciales.

En el prototipo actual, los objetos son proveedores descritos mediante variables tabulares simuladas. Esta representación permite comparar valores numéricos, pero no incorpora documentos, especificaciones ni características particulares de un procedimiento. Por ello, la recomendación resultante se limita a la información artificialmente generada.

Predicción, puntuación, ranking, recomendación y decisión

Estos cinco conceptos deben diferenciarse a lo largo de toda la obra:

1. La predicción es la estimación de un resultado desconocido. Puede consistir en una categoría, un valor o una probabilidad.
2. La puntuación es un valor utilizado para representar el grado de relevancia, compatibilidad, utilidad o probabilidad asociado con una alternativa.
3. El ranking es el ordenamiento de los objetos según sus puntuaciones o según una regla de prioridad.
4. La recomendación es la selección y presentación de una o varias alternativas al usuario. Puede derivarse de un ranking, pero también incorporar restricciones, diversidad y contexto.
5. La decisión es el acto mediante el cual una persona o institución selecciona una acción y asume la responsabilidad correspondiente.

Un clasificador binario puede producir una probabilidad para cada proveedor. Estas probabilidades pueden utilizarse para construir un ranking. Presentar los primeros lugares transforma el resultado en una recomendación. Sin embargo, ninguna de estas operaciones constituye por sí misma la decisión de adjudicar.

La distinción también permite evaluar correctamente el prototipo. El código desarrollado genera estimaciones de probabilidad y compara alternativas. Por consiguiente, debe describirse como un modelo predictivo utilizado para construir una priorización experimental, no como un sistema completo de adjudicación o un recomendador validado en producción.

Tabla 6

Diferencias entre predicción, puntuación, ranking, recomendación y decisión

Concepto	*Resultado producido*	*Ejemplo dentro del prototipo*	*Responsable*
Predicción	Categoría, valor o probabilidad	Probabilidad simulada asociada con un proveedor	Modelo
Puntuación	Medida de relevancia o utilidad	Valor asignado a cada alternativa	Modelo o regla
Ranking	Lista ordenada	Proveedores organizados de mayor a menor puntuación	Sistema
Recomendación	Alternativa o conjunto presentado al usuario	Proveedores priorizados para revisión	Sistema de apoyo
Decisión	Selección institucional con consecuencias	Determinación adoptada dentro del procedimiento	Autoridad o responsable humano

Nota. *La predicción y la recomendación constituyen insumos para el análisis. No sustituyen la decisión administrativa ni trasladan la responsabilidad institucional al sistema.*

Componentes de un sistema de recomendación

Un sistema de recomendación puede organizarse mediante varios componentes relacionados:

1. Fuentes de datos, donde se almacenan perfiles, atributos, interacciones, reglas y contexto.
2. Preparación de información, que comprende limpieza, transformación, integración y validación.
3. Modelo de recomendación, encargado de estimar puntuaciones o relaciones.
4. Mecanismo de selección, que ordena y filtra las alternativas.
5. Interfaz, mediante la cual el usuario recibe la recomendación.
6. Explicación, que comunica los factores asociados con el resultado.
7. Retroalimentación, donde se registran nuevas interacciones o evaluaciones.

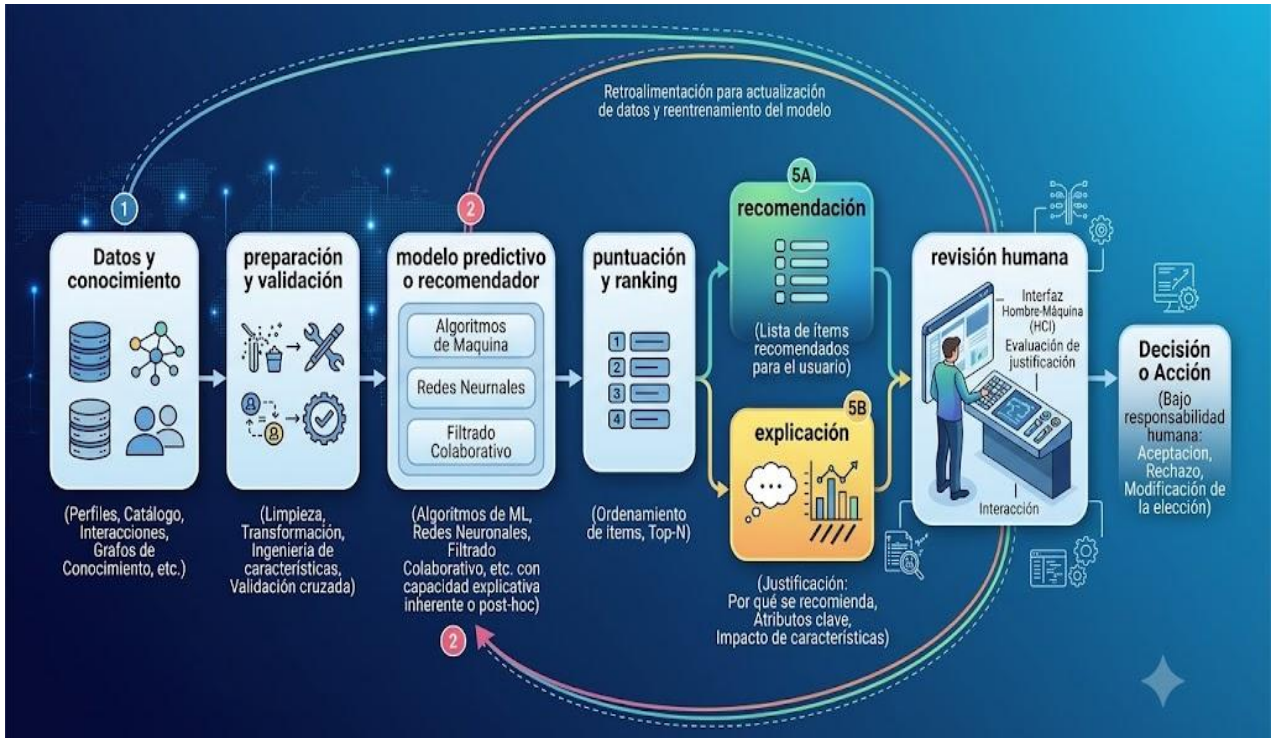
8. Supervisión, que permite revisar, corregir y auditar el funcionamiento.

La descripción original del manuscrito señalaba únicamente datos de entrada, algoritmo y base de datos. Estos elementos se conservan, pero resultan insuficientes para representar la arquitectura completa. La interfaz, la explicación, la retroalimentación y la supervisión son indispensables en una herramienta orientada a decisiones institucionales.

El sistema tampoco debe considerarse una secuencia completamente automática. En diferentes etapas puede requerir validación humana, especialmente durante la definición de reglas, revisión de datos, interpretación de explicaciones y evaluación de resultados.

Figura 5

Arquitectura general de un sistema de recomendación explicable



Nota. *Elaboración propia basada en Adomavicius y Tuzhilin (2005), Ricci et al. (2022) y Zhang y Chen (2020).*

Sistemas personalizados y no personalizados

No todas las recomendaciones se adaptan a una persona específica. Los sistemas no personalizados presentan las mismas alternativas a todos los usuarios, por ejemplo, elementos más consultados, mejor evaluados o recientemente incorporados.

Estas recomendaciones pueden funcionar como una línea de base. Su implementación es sencilla y no requiere construir perfiles individuales. Sin embargo, ofrecen poca adaptación a necesidades particulares.

Los sistemas personalizados utilizan información relacionada con el usuario, sus preferencias, historial o contexto. Su objetivo consiste en producir recomendaciones diferentes para personas distintas.

Una recomendación personalizada no es necesariamente superior. Cuando existe poca información sobre el usuario, la personalización puede resultar inestable. También puede limitar la diversidad si el sistema reproduce continuamente elecciones anteriores.

En un entorno institucional, parte de la recomendación puede ser no personalizada y responder a criterios generales, mientras otra parte puede adaptarse a la necesidad particular del procedimiento. Esta combinación debe ser transparente para que el usuario identifique qué elementos proceden de reglas comunes y cuáles dependen de información específica.

Recomendación y apoyo a decisiones

Los sistemas de recomendación pueden formar parte de una arquitectura de apoyo a decisiones, pero ambos conceptos no son equivalentes. El recomendador selecciona o prioriza alternativas. El sistema de apoyo puede incorporar además simulaciones, reglas, análisis comparativos, documentación y mecanismos de revisión.

En dominios de consumo, la consecuencia de una recomendación incorrecta puede limitarse a una experiencia insatisfactoria. En contratación pública, una recomendación inadecuada puede afectar el análisis de recursos, oportunidades y participantes. Por ello, la herramienta requiere mayores niveles de trazabilidad y control.

La recomendación debe presentarse como una propuesta revisable. El usuario necesita conocer qué información fue considerada, qué criterios influyeron, qué elementos no fueron analizados y qué nivel de incertidumbre posee el resultado.

Un diseño responsable también debe evitar la automatización por defecto. La interfaz no debería inducir a aceptar la primera opción sin revisar alternativas. Puede ser conveniente mostrar varias posibilidades, sus diferencias y explicaciones, en lugar de presentar un único resultado como indiscutible.

Evaluación de un sistema de recomendación

La evaluación de un sistema de recomendación depende de la tarea. Cuando el objetivo es predecir valoraciones, pueden utilizarse medidas de error. Cuando se genera una lista ordenada, resultan pertinentes métricas que examinan si los elementos relevantes aparecen en las primeras posiciones.

La evaluación también puede incorporar cobertura, diversidad, novedad y utilidad. La cobertura expresa qué proporción de usuarios u objetos puede atender el sistema. La diversidad examina las diferencias entre las alternativas recomendadas. La novedad se relaciona con la capacidad de presentar opciones que el usuario probablemente no conocía.

Herlocker et al. (2004) señalan que las métricas deben seleccionarse de acuerdo con la tarea real y que no existe una única medida adecuada para todos los sistemas. Un algoritmo puede mejorar una métrica y no producir una experiencia o decisión más útil.

En sistemas explicables, la evaluación debe extenderse a la calidad de las razones proporcionadas. La recomendación puede ser técnicamente relevante, pero su explicación resultar confusa o infiel. También puede existir una explicación clara para una recomendación inadecuada.

Para evaluar el prototipo serían necesarias, al menos, tres dimensiones:

- Desempeño predictivo.
- Calidad del ranking o priorización.
- Fidelidad y utilidad de las explicaciones.

La evaluación actual del manuscrito se concentra principalmente en la exactitud del clasificador. Esta medida no es suficiente para afirmar que existe un sistema de recomendación efectivo. Deben definirse los usuarios, la tarea, los objetos, el criterio de relevancia y el procedimiento utilizado para valorar la priorización.

Adaptación del concepto al prototipo del libro

El prototipo desarrollado utiliza características simuladas de proveedores, un clasificador de bosque aleatorio y una interfaz que compara resultados. En consecuencia, combina elementos de predicción, puntuación y apoyo a decisiones.

No utiliza una matriz clásica de usuarios y objetos ni registra preferencias de múltiples usuarios. Tampoco aprende a partir de valoraciones colaborativas. Por ello, no debe definirse estrictamente como un sistema de filtrado colaborativo.

Su estructura puede describirse con mayor precisión como un prototipo experimental de apoyo analítico que emplea un modelo de clasificación para asignar puntuaciones a proveedores simulados, ordenar alternativas y presentar información explicativa para su revisión humana. Esta definición permite conservar la relación con los sistemas de recomendación, pero evita atribuirle funciones que no están implementadas.

El término recomendación se justifica únicamente en la medida en que las puntuaciones se utilizan para priorizar alternativas. La obra debe aclarar que esta priorización se desarrolla en un escenario simulado y no representa una recomendación institucional validada.

La evolución futura del prototipo podría orientarse hacia dos configuraciones distintas. La primera recomendaría oportunidades contractuales a los proveedores mediante compatibilidad entre perfiles y requisitos. La segunda apoyaría a las entidades en la organización de información de las ofertas. Cada configuración necesitaría datos, usuarios, objetos, reglas y evaluaciones diferentes.

Los sistemas de recomendación pueden construirse mediante diferentes estrategias según la información disponible y el tipo de relación que se desea modelar. Los enfoques basados en contenido utilizan características descriptivas; los colaborativos aprovechan patrones de interacción; los sistemas basados en conocimiento aplican reglas y restricciones; y los modelos híbridos combinan varias fuentes. El apartado siguiente examina estos enfoques, sus ventajas, sus limitaciones y su posible relación con el prototipo.

Principales enfoques de los sistemas de recomendación

Los sistemas de recomendación pueden clasificarse según la información que emplean, la forma en que representan a los usuarios y objetos, y el procedimiento mediante el cual generan una puntuación o selección. Las categorías más consolidadas comprenden la recomendación basada en contenido, el filtrado colaborativo, los modelos basados en conocimiento y los enfoques híbridos. A estas categorías pueden añadirse extensiones que incorporan información contextual, múltiples criterios o interacción conversacional.

La clasificación no debe confundirse con las fuentes utilizadas para recoger información. Una encuesta, una valoración, una actividad registrada o una variable económica no constituyen necesariamente tipos independientes de sistemas de recomendación. Son mecanismos de obtención

de preferencias, datos de interacción, atributos de los objetos o restricciones que pueden integrarse dentro de diferentes modelos.

Esta distinción permite reorganizar varios conceptos del manuscrito original. Los sistemas denominados “basados en inspección” corresponden, en realidad, a la obtención explícita de preferencias mediante formularios o preguntas. Los sistemas “basados en actividades” utilizan retroalimentación implícita. El “filtrado económico” incorpora el costo como atributo o restricción. Los sistemas basados en reglas pertenecen al campo de la recomendación basada en conocimiento.

Sistemas de recomendación basados en contenido

Los sistemas basados en contenido recomiendan objetos cuyas características son semejantes a las de aquellos que el usuario ha considerado relevantes anteriormente. Su funcionamiento exige representar los objetos mediante atributos y construir un perfil que describa las preferencias identificadas.

Pazzani y Billsus (2007) explican que estos sistemas aprenden un perfil a partir de las características de los elementos asociados con respuestas positivas o negativas del usuario. Posteriormente, comparan ese perfil con la representación de nuevos objetos para estimar cuáles podrían resultar pertinentes.

La representación del contenido depende del dominio. En un sistema de noticias, puede construirse mediante palabras, temas o categorías. En una plataforma de productos, puede incluir marca, precio, funciones y especificaciones. En contratación pública, una oportunidad podría representarse mediante su objeto, categoría, presupuesto, ubicación, plazo, requisitos técnicos y experiencia solicitada.

El perfil del usuario puede elaborarse de forma explícita o implícita. En la primera modalidad, la persona declara sus intereses, necesidades o restricciones. En la segunda, el sistema infiere preferencias a partir de objetos consultados, seleccionados o valorados anteriormente.

Una ventaja de los modelos basados en contenido es que pueden generar recomendaciones sin depender de las valoraciones de otros usuarios. También permiten relacionar la explicación con atributos observables. Un sistema puede señalar que una oportunidad se recomienda debido a su categoría, ubicación, presupuesto o correspondencia con la experiencia declarada.

Sin embargo, la calidad de la recomendación depende de la representación utilizada. Cuando los atributos son incompletos, ambiguos o excesivamente generales, el modelo puede considerar semejantes objetos que presentan diferencias relevantes.

También existe el riesgo de sobreespecialización. Si el sistema recomienda únicamente objetos semejantes a los seleccionados anteriormente, puede reducir la diversidad y limitar el descubrimiento de nuevas alternativas. Un proveedor que ha participado principalmente en una categoría podría recibir continuamente oportunidades del mismo tipo, aunque posea capacidad para ampliar su actividad.

Los modelos basados en contenido encuentran dificultades cuando el usuario es nuevo y no existe información suficiente para construir su perfil. Este problema puede reducirse mediante formularios iniciales, incorporación de características declaradas o reglas basadas en conocimiento.

Aplicación potencial al contexto del libro

Un enfoque basado en contenido sería especialmente pertinente si el sistema se orientara a recomendar oportunidades contractuales a proveedores. El perfil de la empresa podría compararse con los atributos de los procedimientos publicados. La recomendación se basaría en compatibilidad y no en la predicción de adjudicaciones anteriores.

Esta configuración sería diferente de la desarrollada actualmente en el prototipo. El modelo existente clasifica proveedores simulados y no compara directamente un perfil empresarial con las características de una licitación específica.

Filtrado colaborativo

El filtrado colaborativo utiliza patrones de interacción compartidos por múltiples usuarios. Su principio consiste en aprovechar las valoraciones o comportamientos registrados para identificar semejanzas y predecir qué objetos podrían resultar relevantes.

Uno de los antecedentes del enfoque fue GroupLens, sistema que utilizaba las valoraciones de lectores de noticias para predecir el interés de otros usuarios. El sistema demostró que la información distribuida entre una comunidad podía utilizarse para producir estimaciones personalizadas (Resnick et al., 1994).

A diferencia de la recomendación basada en contenido, el filtrado colaborativo no necesita conocer necesariamente las propiedades internas de los objetos. Puede identificar relaciones a partir de las interacciones. Dos documentos, productos o servicios pueden aparecer vinculados porque usuarios semejantes interactuaron con ambos, aunque sus descripciones textuales sean diferentes.

La información suele representarse mediante una matriz en la que las filas corresponden a usuarios, las columnas a objetos y las celdas contienen valoraciones o interacciones. En la práctica, esta matriz suele ser dispersa, porque cada usuario interactúa con una proporción reducida del catálogo.

El filtrado colaborativo puede organizarse en métodos basados en memoria y métodos basados en modelos. Los primeros generan recomendaciones directamente a partir de relaciones de semejanza. Los segundos aprenden una representación o función predictiva utilizando los datos de interacción.

Pertinencia para el prototipo

El filtrado colaborativo no se encuentra realmente implementado en el código descrito en el manuscrito. El prototipo no contiene una matriz de múltiples usuarios y proveedores, ni valoraciones históricas mediante las cuales puedan identificarse preferencias compartidas.

Por este motivo, el filtrado colaborativo debe conservarse como parte del marco teórico, pero no debe presentarse como técnica utilizada en la solución desarrollada.

Filtrado colaborativo basado en usuarios

El filtrado basado en usuarios identifica personas o entidades con patrones de valoración semejantes. La predicción para un usuario se construye a partir de las preferencias observadas en sus vecinos más próximos.

El procedimiento general comprende cuatro etapas:

1. Representar las valoraciones o interacciones de cada usuario.
2. Calcular semejanzas entre usuarios.
3. Seleccionar un conjunto de vecinos.
4. Agregar sus valoraciones para estimar la relevancia de objetos no evaluados.

La semejanza puede calcularse mediante correlación de Pearson, similitud de coseno o coeficiente de Jaccard, dependiendo de la naturaleza de los datos. La correlación de Pearson considera la relación entre patrones de valoración y puede compensar diferencias en las escalas personales. La similitud de coseno examina el ángulo entre vectores. Jaccard resulta pertinente cuando las interacciones se representan mediante conjuntos o variables binarias.

La recomendación depende de la cantidad y calidad de interacciones compartidas. Dos usuarios pueden parecer semejantes si coinciden en pocos objetos, pero esa relación podría ser inestable. Por

ello, suele exigirse un número mínimo de coincidencias o aplicarse una corrección según la cantidad de evidencia disponible.

Una ventaja del enfoque es su carácter intuitivo. La recomendación puede explicarse mediante la semejanza con otros usuarios. Sin embargo, este tipo de explicación debe proteger la privacidad y evitar la exposición de información individual.

El filtrado basado en usuarios enfrenta problemas de escalabilidad cuando el número de personas aumenta y se requiere calcular numerosas relaciones. También presenta dificultades cuando los perfiles cambian rápidamente o cuando la matriz contiene pocas interacciones.

En contratación pública, este enfoque tendría una aplicabilidad limitada si se pretendiera comparar entidades contratantes o proveedores mediante comportamientos históricos. Las semejanzas encontradas no demostrarían que dos procedimientos sean jurídicamente o técnicamente equivalentes.

Filtrado colaborativo basado en objetos

El filtrado basado en objetos calcula semejanzas entre los elementos del catálogo a partir de los patrones de interacción de los usuarios. Para generar una recomendación, identifica objetos similares a aquellos con los que el usuario ya ha interactuado.

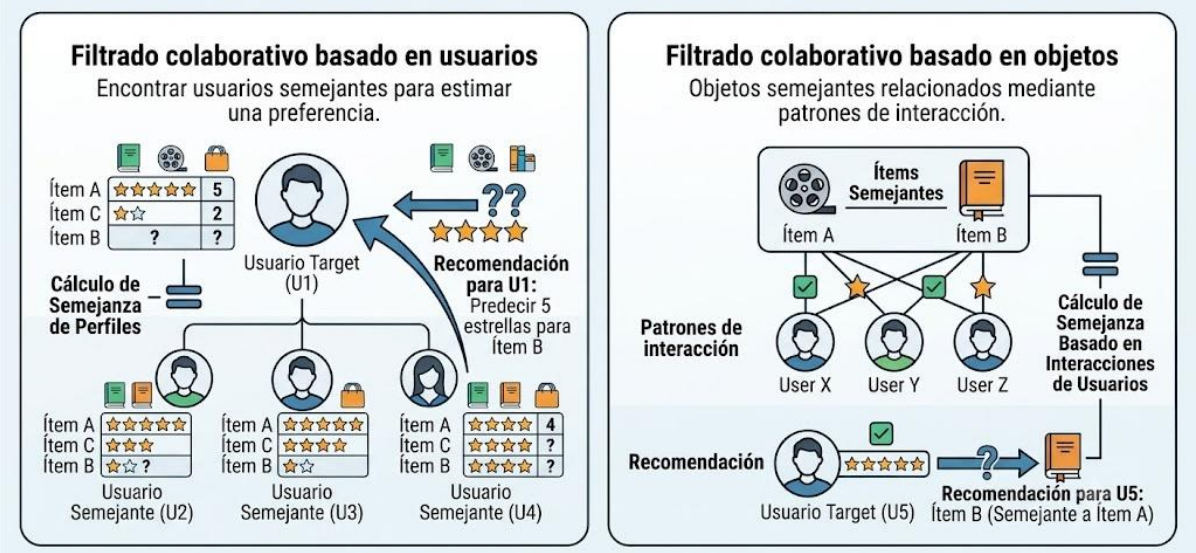
Sarwar et al. (2001) desarrollaron y evaluaron algoritmos de filtrado colaborativo basados en objetos, mostrando que podían ofrecer ventajas de calidad y escalabilidad frente a ciertos métodos centrados en usuarios. La semejanza se calcula mediante las columnas de la matriz de interacción, y las predicciones se obtienen combinando las valoraciones de objetos relacionados.

No obstante, la semejanza colaborativa refleja comportamiento y no equivalencia técnica. Dos procedimientos pueden aparecer relacionados porque fueron consultados por los mismos

proveedores, aunque sus requisitos sean diferentes. La recomendación debería combinarse con contenido o reglas que verifiquen su compatibilidad.

Figura 6

Filtrado colaborativo basado en usuarios y basado en objetos



Nota. *Elaboración propia basada en Resnick et al. (1994) y Sarwar et al. (2001).*

Filtrado colaborativo basado en modelos

Los métodos basados en modelos utilizan algoritmos estadísticos o de aprendizaje automático para aprender una representación de las interacciones. En lugar de calcular directamente las relaciones entre usuarios u objetos cada vez que se genera una recomendación, construyen un modelo capaz de estimar valores faltantes o producir un ranking.

Entre estos métodos se encuentran:

- Factorización matricial.
- Modelos probabilísticos.
- Agrupamiento.
- Redes neuronales.

- Métodos de clasificación y regresión.
- Modelos de factores latentes.

La factorización matricial representa a usuarios y objetos en un espacio de menor dimensionalidad. Cada usuario y cada objeto se describen mediante factores latentes aprendidos a partir de las interacciones. La predicción se obtiene mediante la compatibilidad entre ambas representaciones.

Koren et al. (2009) explican que estos modelos permiten descubrir dimensiones latentes que caracterizan a los usuarios y objetos, y alcanzaron especial relevancia por su capacidad para trabajar con matrices dispersas de valoraciones.

Los factores latentes pueden mejorar el rendimiento, pero no poseen necesariamente un significado directamente observable. Una dimensión aprendida por el modelo puede combinar múltiples propiedades sin que resulte posible asignarle una interpretación sencilla.

La factorización matricial tampoco resuelve automáticamente el problema de usuarios u objetos nuevos. Si no existen interacciones, el sistema dispone de poca información para estimar sus factores. Una alternativa consiste en incorporar atributos de contenido o utilizar modelos híbridos.

Los sistemas basados en modelos requieren un proceso de entrenamiento y actualización. Su calidad depende de la función objetivo, regularización, cantidad de factores y procedimiento de evaluación. También pueden perder vigencia cuando cambian los intereses de los usuarios o las características del catálogo.

Relación con el bosque aleatorio

El bosque aleatorio utilizado en el prototipo también es un modelo de aprendizaje automático, pero no constituye filtrado colaborativo basado en modelos. No aprende factores a partir de una matriz usuario-objeto, sino que clasifica registros descritos mediante características tabulares.

Por tanto, debe evitarse afirmar que el prototipo implementa filtrado colaborativo solamente porque utiliza aprendizaje automático.

Sistemas de recomendación basados en conocimiento

Los sistemas basados en conocimiento generan recomendaciones mediante información explícita sobre las necesidades del usuario, las características de los objetos y las relaciones que determinan cuándo una alternativa resulta apropiada. Su funcionamiento puede apoyarse en reglas, restricciones, ontologías, razonamiento basado en casos o conocimiento proporcionado por especialistas.

Este enfoque es útil cuando las decisiones son poco frecuentes, los objetos son complejos o no existe suficiente historial de interacciones. En estas circunstancias, el filtrado colaborativo puede disponer de poca evidencia y la recomendación basada en contenido puede resultar insuficiente para representar condiciones obligatorias.

Los sistemas basados en conocimiento pueden organizarse en dos modalidades principales:

1. Basados en restricciones, cuando verifican si los objetos satisfacen condiciones establecidas.
2. Basados en casos, cuando comparan la necesidad actual con situaciones o soluciones anteriores.

Las restricciones pueden ser obligatorias o preferenciales. Las obligatorias determinan si una alternativa puede ser considerada. Las preferenciales permiten ordenar opciones que ya cumplen las condiciones mínimas.

Esta distinción es especialmente relevante en contratación pública. Determinados requisitos no deben compensarse mediante una puntuación elevada en otros criterios. Si una condición habilitante es obligatoria, su incumplimiento no puede neutralizarse mediante experiencia, precio u otra variable.

Una ventaja de este enfoque es que las recomendaciones pueden explicarse mediante reglas explícitas. El sistema puede indicar qué condiciones se cumplieron, cuáles no se verificaron y qué atributos determinaron la compatibilidad.

La principal limitación consiste en la adquisición y mantenimiento del conocimiento. Las reglas deben ser definidas, revisadas y actualizadas por especialistas. Una norma modificada o una condición técnica desactualizada puede producir recomendaciones incorrectas.

En la evolución futura del prototipo, un componente basado en conocimiento podría utilizarse para aplicar restricciones antes de generar el ranking predictivo. De esta manera, el modelo estadístico actuaría únicamente sobre alternativas que ya cumplen condiciones verificables.

Sistemas basados en reglas

Los sistemas basados en reglas constituyen una modalidad de recomendación basada en conocimiento. Utilizan estructuras condicionales que relacionan las características del usuario o la necesidad con los atributos de los objetos.

Una regla simple puede adoptar la forma “si la alternativa cumple las condiciones A, B y C, puede incluirse en el conjunto de opciones”.

Las reglas pueden representar requisitos obligatorios, relaciones de compatibilidad, preferencias o criterios de exclusión. Su ventaja es la trazabilidad, porque el usuario puede identificar la condición aplicada.

Sin embargo, una gran cantidad de reglas puede producir contradicciones, duplicaciones o dificultades de mantenimiento. También existe el riesgo de simplificar situaciones que requieren interpretación especializada.

Las reglas no deben inferirse únicamente a partir de patrones históricos cuando representan requisitos jurídicos. Deben construirse con participación de especialistas y vincularse con una fuente normativa o técnica verificable.

El apartado original sobre "filtrado basado en reglas" se conserva conceptualmente dentro de esta categoría. No debe permanecer como un enfoque separado de los sistemas basados en conocimiento.

Sistemas híbridos

Los sistemas híbridos combinan dos o más estrategias de recomendación con el propósito de aprovechar sus fortalezas y compensar sus limitaciones. Una configuración frecuente integra información de contenido con patrones colaborativos. También pueden combinarse reglas, conocimiento, contexto y modelos predictivos.

Burke (2002) identificó diferentes formas de hibridación, entre ellas:

- Combinación ponderada de puntuaciones.
- Selección o cambio entre métodos.
- Presentación conjunta de resultados.
- Aplicación secuencial o en cascada.
- Uso de la salida de un modelo como entrada de otro.
- Incorporación de características de un método dentro de otro.

El enfoque ponderado combina los resultados mediante pesos. El sistema de selección utiliza un método u otro según la disponibilidad de información. La estrategia mixta presenta recomendaciones procedentes de varios componentes. En una cascada, un método genera un conjunto inicial y otro refina el orden.

La hibridación puede disminuir los problemas de inicio en frío y dispersión. Un usuario nuevo podría recibir recomendaciones basadas en contenido o conocimiento hasta que existan interacciones suficientes para incorporar filtrado colaborativo.

Sin embargo, un sistema híbrido aumenta la complejidad. Cada componente requiere datos, entrenamiento, evaluación y mantenimiento. También debe explicarse cómo se combinaron las salidas y qué método influyó en cada recomendación.

Una posible evolución del prototipo combinaría:

1. Reglas para verificar condiciones obligatorias.
2. Compatibilidad basada en contenido.
3. Un modelo predictivo para asignar puntuaciones.
4. Una capa XAI para explicar el resultado.
5. Revisión humana antes de cualquier uso.

Esta arquitectura sería híbrida, pero requeriría una evaluación específica de cada componente y del sistema completo.

Problemas comunes de los enfoques de recomendación

Inicio en frío

El inicio en frío aparece cuando el sistema no dispone de información suficiente sobre un usuario u objeto nuevo. Los métodos colaborativos son especialmente sensibles porque dependen de interacciones anteriores.

Las estrategias de mitigación incluyen formularios iniciales, atributos de contenido, reglas, conocimiento del dominio y modelos híbridos. La solución debe equilibrar la obtención de información con el esfuerzo exigido al usuario.

Dispersión de datos

La dispersión ocurre cuando la matriz contiene numerosas celdas vacías. Esta situación dificulta identificar vecinos y aprender representaciones estables.

La factorización matricial y otros modelos latentes pueden trabajar con información dispersa, pero no eliminan la necesidad de contar con interacciones suficientes y representativas (Koren et al., 2009).

Escalabilidad

El crecimiento del número de usuarios, objetos e interacciones aumenta los requisitos de procesamiento. Los métodos basados en vecinos pueden requerir numerosos cálculos de similitud.

Los modelos basados en objetos, las aproximaciones y el procesamiento distribuido pueden mejorar la escalabilidad, aunque introducen nuevas decisiones de diseño.

Sobreespecialización

Los sistemas basados en contenido pueden recomendar alternativas excesivamente semejantes al historial del usuario. Esto reduce la diversidad y puede impedir el descubrimiento de nuevas opciones.

Popularidad y concentración

Los algoritmos pueden favorecer objetos con muchas interacciones, porque disponen de mayor evidencia. Esta tendencia puede reducir la visibilidad de alternativas nuevas o menos conocidas.

En contratación pública, un sesgo de popularidad podría favorecer a proveedores con mayor historial y reducir la exposición de nuevos participantes, aun cuando posean capacidad suficiente.

Vulnerabilidad a manipulaciones

Las valoraciones y comportamientos pueden ser alterados con el propósito de influir en las recomendaciones. Los sistemas necesitan mecanismos para detectar interacciones anómalas y proteger la integridad de los datos.

Privacidad

La personalización requiere información sobre usuarios y comportamientos. La recolección debe limitarse a los datos necesarios y acompañarse de controles de acceso, información de finalidad y políticas de conservación.

Explicabilidad

Los modelos colaborativos y latentes pueden generar resultados difíciles de relacionar con atributos observables. La explicación debe evitar razones genéricas o construidas únicamente para persuadir.

Enfoque más coherente con la evolución del prototipo

El prototipo actual no implementa recomendación basada en contenido, filtrado colaborativo ni un modelo híbrido completo. Utiliza características tabulares simuladas y un clasificador que asigna probabilidades.

Una evolución coherente podría adoptar un enfoque híbrido y explicable:

- Las reglas verificarían requisitos obligatorios.
- El contenido mediría compatibilidad entre perfiles y oportunidades.
- El modelo predictivo proporcionaría una puntuación complementaria.
- La capa explicativa mostraría las contribuciones de las variables.
- La interfaz presentaría varias alternativas.

- La revisión humana determinaría la utilidad del resultado.

Los distintos enfoques muestran que la recomendación puede construirse mediante características, interacciones, reglas, contexto o combinaciones de estos elementos. Sin embargo, la presentación de una alternativa no garantiza que el usuario comprenda las razones que la sustentan. El apartado siguiente integra los sistemas de recomendación con la inteligencia artificial explicable y examina las finalidades, formas y riesgos de las explicaciones asociadas con una recomendación.

Sistemas de recomendación basados en conocimiento

Realizan recomendaciones partiendo del conocimiento que da el usuario sobre sus necesidades, y del conocimiento de los objetos a recomendar, buscando los que mejor se adapten a las necesidades de los usuarios, apoyados fundamentalmente en el razonamiento basado en casos. (Valenzuela, Correa, & Rendon, 2022)

Figura 7

Tipos de sistemas de recomendación

Nota. *Elaboración propia basada en Recommender Systems Handbook (Ricci et al., 2015) y Adomavicius y Tuzhilin (2005).*

Sistemas de Recomendación Basado en Conocimientos

Los sistemas de recomendación basado en conocimiento proporcionan sugerencias deduciendo las necesidades y preferencias del usuario. Este enfoque se destaca por utilizar información específica sobre cómo un objeto puede satisfacer las necesidades del usuario, lo que permite razonar sobre la relación entre una necesidad y la recomendación potencial.

Estos sistemas se apoyan en la creación de perfiles de usuarios y pueden enriquecerse utilizando expresiones en lengua natura. (Valenzuela, Correa, & Rendon, 2022)

Sistemas de recomendación basado en contenido

Los sistemas de recomendaciones basado en contenidos se enfocan única y exclusivamente en el usuario activo, es decir sugiere artículos que comparten características con aquellos que el usuario ha valorado positivamente anteriormente. Este tipo de sistemas se dedica a elaborar perfiles de usuarios y a analizar la correspondencia entre sus características y los atributos de los objetos. (Couñago, 2021)

Sistemas de recomendación filtrado colaborativo

Según el autor (Fonseca & Cornelio, 2022), podemos concluir que el funcionamiento de estos sistemas se basa en la información disponible sobre los usuarios. Analizar las compras, preferencias o calificaciones de todos los usuarios y los agrupan según sus características, formando conjuntos de usuarios con preferencias similares. Para un usuario específico, se recomiendan productos que han sido bien recibidos por los usuarios con gustos similares, pero que él aún no ha consumido.

Aplicación Web

Una aplicación web se refiere a un programa en línea que opera a través de Internet. En este entorno, numerosos usuarios interactúan mediante un navegador, realizan solicitudes remotas y aguardan respuestas que abarcan aspectos diversos como la presentación de información impresa,

el desarrollo de software, estrategias de marketing, aspectos informáticos, comunicación interna y relaciones externas, así como la combinación de arte y tecnología. Mendoza, M., & Barrios, J. (2004). Propuestas metodológicas para el desarrollo de aplicaciones web: una evaluación según la ingeniería de métodos. Revista ciencia e ingeniería, 25(2), 89-90.

Python

Es un lenguaje de programación versátil utilizado en aplicaciones web, desarrollo de software y machine learning. Su popularidad se debe a su sintaxis, que facilita la escritura y lectura del código. Además, es compatible con diversos sistemas operativos y es de acceso libre, lo que contribuye a su amplio uso.

MySQL

Es un sistema gestor de base de datos o SGBD relacional, multihilo y multiusuario. Se adapta a diferentes entornos de desarrollo, y permite que la interacción con los lenguajes de programación sea factible, así como la integración con los distintos sistemas operativos. También destaca por ser open source, su configuración y administración menos compleja.

CAPITULO III: DISEÑO METODOLÓGICO Y GOBERNANZA DE DATOS

CHAPTER III: METHODOLOGICAL DESIGN AND DATA GOVERNANCE

CAPITULO III: DISEÑO METODOLÓGICO Y GOBERNANZA DE DATOS

El presente capítulo describe el diseño metodológico adoptado para la construcción de un prototipo experimental de apoyo analítico basado en inteligencia artificial explicable. La propuesta integra la generación programática de datos sintéticos, el procesamiento de variables relacionadas con proveedores tecnológicos, el entrenamiento de un modelo de clasificación supervisada y la presentación de puntuaciones, comparaciones y elementos explicativos destinados a facilitar la revisión humana.

El propósito del capítulo no consiste en demostrar que el sistema puede utilizarse directamente para evaluar ofertas o seleccionar proveedores dentro de procedimientos reales de contratación pública. Su función es documentar la manera en que fueron articulados los componentes del prototipo, explicar las decisiones adoptadas durante su desarrollo y establecer las condiciones bajo las cuales deben interpretarse sus resultados.

La metodología se estructura alrededor de siete dimensiones. La primera delimita el enfoque y el alcance del prototipo. La segunda describe la estrategia iterativa utilizada durante el desarrollo y examina sus condiciones básicas de factibilidad. La tercera identifica la procedencia, naturaleza y gobernanza de los datos. La cuarta operacionaliza las variables y documenta los procedimientos de generación y preparación del conjunto sintético. La quinta explica el diseño, entrenamiento y evaluación del modelo predictivo. La sexta presenta la arquitectura funcional, el mecanismo de comparación, las visualizaciones y los componentes de explicabilidad. La séptima analiza los criterios éticos, la reproducibilidad y las limitaciones metodológicas.

La gobernanza de datos se incorpora como una dimensión transversal. Su inclusión responde a que el comportamiento de un modelo no depende exclusivamente del algoritmo seleccionado, sino también de la forma en que se definen, generan, transforman, documentan y utilizan los datos. En este sentido, la trazabilidad debe permitir reconstruir el origen de cada variable, las operaciones

aplicadas, la versión del código, los parámetros del modelo y las condiciones bajo las cuales se produjo cada resultado.

El marco de gestión de riesgos de inteligencia artificial del Instituto Nacional de Estándares y Tecnología plantea que la confiabilidad de estos sistemas requiere atender aspectos como validez, seguridad, transparencia, explicabilidad, privacidad y control de sesgos perjudiciales. Asimismo, organiza la gestión del riesgo mediante cuatro funciones relacionadas: gobernar, contextualizar, medir y gestionar. Estos principios permiten comprender que la evaluación de un prototipo no debe limitarse al cálculo de una métrica predictiva, sino que debe examinar todo el ciclo de vida del sistema (National Institute of Standards and Technology [NIST], 2023).

A partir de esta orientación, el capítulo establece una separación expresa entre demostración tecnológica e implementación institucional. El artefacto permite observar un flujo computacional y analizar sus componentes, pero no proporciona evidencia suficiente para concluir que mejora la eficiencia de la contratación pública, reduce costos, elimina sesgos o identifica correctamente a proveedores idóneos en condiciones reales. Tales afirmaciones necesitarían datos institucionales verificables, evaluación externa, participación de especialistas y pruebas realizadas en contextos de uso definidos.

Enfoque metodológico y alcance del prototipo

La propuesta posee un enfoque aplicado, tecnológico y orientado al diseño de un artefacto informático. Su finalidad consiste en construir una prueba de concepto mediante la integración de datos, algoritmos, visualizaciones y mecanismos explicativos. No corresponde a una investigación experimental destinada a establecer relaciones causales entre características empresariales y adjudicaciones públicas, ni a un estudio estadístico representativo de los proveedores que participan en el Sistema Nacional de Contratación Pública.

El prototipo se desarrolla como una aproximación demostrativa. Esta condición significa que su valor se encuentra principalmente en la representación de una arquitectura funcional y en la documentación de las operaciones necesarias para transformar datos en salidas interpretables. La demostración permite comprobar que los componentes pueden conectarse y ejecutarse, pero no equivale a demostrar que el sistema posee validez externa o utilidad institucional.

La pregunta metodológica que orienta el desarrollo es ¿Cómo puede estructurarse un prototipo experimental que procese datos sintéticos de proveedores tecnológicos, genere puntuaciones mediante un modelo predictivo y presente información explicativa para apoyar la revisión humana de sus resultados?

La formulación evita atribuir al modelo la capacidad de adjudicar contratos o seleccionar automáticamente al proveedor más adecuado. El sistema procesa un conjunto limitado de variables, genera una estimación matemática y presenta una comparación. La decisión administrativa, técnica o jurídica permanece fuera de sus funciones.

Tipo de artefacto desarrollado

El artefacto corresponde a un prototipo analítico de prueba de concepto. Está compuesto por cinco elementos principales:

1. Un procedimiento para generar registros sintéticos de proveedores tecnológicos.
2. Un módulo de preparación y transformación de las variables.
3. Un modelo de clasificación basado en bosques aleatorios.
4. Una función para comparar las puntuaciones de dos registros.
5. Un conjunto de visualizaciones y explicaciones destinadas a facilitar la interpretación.

Estos componentes ya se encuentran presentes en el manuscrito original. El código genera 2000 observaciones, emplea variables numéricas y binarias, configura un clasificador de bosque aleatorio

y divide los datos entre entrenamiento y prueba. También incluye una función que calcula probabilidades y compara dos proveedores.

La reconstrucción editorial no elimina estos recursos. Los reorganiza dentro de una secuencia metodológica coherente y modifica la interpretación de sus resultados. En lugar de afirmar que el modelo calcula una probabilidad real de adjudicación, se utilizará la denominación puntuación predictiva sintética o estimación experimental asociada con la clase positiva.

Naturaleza aplicada de la propuesta

El carácter aplicado se manifiesta en la construcción de un producto computacional dirigido a representar una posible solución tecnológica. Sin embargo, una propuesta aplicada no debe confundirse con una herramienta validada. La aplicabilidad potencial describe la posibilidad de que el diseño sea adaptado y ampliado en el futuro; la validación demuestra su funcionamiento en un contexto real y bajo criterios previamente definidos.

La obra alcanza la etapa de construcción y demostración. No alcanza todavía una validación institucional porque no se documentan:

- una entidad contratante participante;
- una base real de procedimientos;
- criterios oficiales utilizados para construir las variables;
- una muestra de usuarios;
- un instrumento de evaluación de usabilidad;
- un protocolo de comparación con decisiones reales;
- ni una autorización para aplicar la herramienta dentro de un procedimiento público.

Por consiguiente, las referencias a “usuarios reales”, “datos oficiales”, “pruebas exhaustivas” y “recomendaciones precisas” deberán sustituirse por descripciones acordes con una prueba de concepto.

Unidad de análisis

La unidad de análisis está constituida por cada registro sintético de un supuesto proveedor de servicios tecnológicos. Cada fila del conjunto representa una observación hipotética caracterizada mediante atributos relacionados con experiencia, proyectos completados, tiempo de respuesta, costo, satisfacción, certificaciones, personal, cobertura, soporte, especialización y cumplimiento.

La unidad no corresponde a una empresa real, una oferta formal ni un contrato ejecutado. Tampoco representa necesariamente un procedimiento específico. Esta precisión resulta indispensable porque el código actual utiliza nombres empresariales identificables y genera múltiples registros aleatorios asociados con ellos. En la versión final, estos nombres deben sustituirse por identificadores ficticios como Proveedor A, Proveedor B, Proveedor C, Proveedor D, Proveedor E, Proveedor F y Proveedor G.

La utilización de identificadores ficticios evita que las puntuaciones generadas puedan interpretarse como valoraciones reales sobre empresas existentes. También mantiene la coherencia metodológica entre la naturaleza sintética de los datos y su representación dentro de las visualizaciones.

Objeto de la estimación

El código original utiliza la variable ganador_licitacion como resultado que debe predecirse. No obstante, esta etiqueta es generada aleatoriamente mediante una distribución de 70 % para la clase cero y 30 % para la clase uno, sin relación programada con las demás variables.

Esta construcción impide interpretar la salida como una probabilidad válida de ganar una licitación. Si el resultado se asigna aleatoriamente, el modelo no dispone de una relación sustantiva que aprender entre las características del proveedor y la etiqueta.

Para conservar el núcleo del prototipo sin presentar resultados metodológicamente incorrectos, se establecen dos alternativas:

- Alternativa metodológica recomendada: reconstruir la variable objetivo mediante una regla sintética documentada que relacione ciertas características con una puntuación experimental. Esta regla no representaría un criterio oficial, pero permitiría comprobar si el algoritmo recupera relaciones introducidas deliberadamente durante la simulación.
- Alternativa de conservación mínima: mantener la etiqueta aleatoria, pero declarar que el ejercicio se utiliza exclusivamente para demostrar la ejecución del código y no para evaluar capacidad predictiva. Bajo esta alternativa no deberían formularse conclusiones sobre precisión, recomendación o selección.

La primera alternativa es editorialmente más sólida porque permite conservar el entrenamiento, la evaluación y las visualizaciones. No obstante, la regla sintética deberá presentarse de manera transparente y acompañarse de la advertencia de que fue creada con fines didácticos.

Niveles de evaluación

La evaluación del prototipo se organizará en cuatro niveles detallados a continuación

Tabla 7

Niveles de evaluación del prototipo

Nivel	*Propósito*	*Alcance en la obra*
Verificación técnica	Comprobar que el código y sus componentes se ejecutan correctamente	Incluido
Evaluación predictiva	Analizar el comportamiento del clasificador mediante métricas	Incluido con alcance sintético
Evaluación explicativa	Determinar si las explicaciones corresponden con el comportamiento del modelo	Parcial, requiere fortalecimiento
Validación institucional	Comprobar la utilidad en procedimientos y con usuarios reales	No incluida

Nota. *Los niveles delimitan la evidencia disponible y evitan equiparar la ejecución técnica con una validación institucional.*

La verificación técnica comprende la ejecución de las funciones, la generación del conjunto, la preparación de variables y la producción de resultados. La evaluación predictiva examina la matriz

de confusión y las métricas correspondientes, siempre dentro del escenario sintético. La evaluación explicativa requiere comprobar que los factores mostrados por la interfaz influyeron realmente en la estimación. La validación institucional queda planteada como una línea futura.

Alcance metodológico

El alcance comprende:

- Diseño de un flujo computacional reproducible;
- Generación programática de información sintética;
- Operacionalización de variables demostrativas;
- Preparación de los datos;
- Entrenamiento de un clasificador;
- Cálculo de métricas;
- Comparación de alternativas simuladas;
- Presentación gráfica de los resultados;
- Documentación de riesgos y limitaciones.

Quedan fuera del alcance:

- Evaluación de proveedores reales;
- Conexión con registros del sercop;
- Interpretación automática de documentos contractuales;
- Verificación legal de requisitos;
- Adjudicación automatizada;
- Medición del ahorro público;
- Análisis causal;
- Pruebas institucionales;
- Y despliegue en un ambiente productivo.

La delimitación no elimina la utilidad académica de la propuesta. Permite presentar el artefacto como lo que efectivamente es: una estructura demostrativa que integra componentes de aprendizaje automático y explicabilidad, y que puede servir como base para una futura investigación con datos reales, criterios especializados y protocolos de validación más rigurosos.

El enfoque metodológico adoptado permite conservar el corazón tecnológico del manuscrito sin atribuirle resultados que no han sido demostrados. El prototipo mantiene su generador de datos, modelo predictivo, módulo de comparación y recursos visuales, pero se interpreta dentro de un escenario sintético, controlado y reproducible. Sobre esta delimitación, el apartado siguiente reorganiza el uso de Scrum y el análisis de factibilidad como componentes de la estrategia de desarrollo, y no como evidencias de validez científica.

Estrategia de desarrollo y condiciones de factibilidad

La construcción del prototipo se organizó mediante una estrategia iterativa orientada a desarrollar, revisar e integrar progresivamente sus componentes. El proceso comprendió la delimitación inicial del problema, la generación del conjunto sintético, la preparación de las variables, el entrenamiento del modelo, la construcción del módulo de comparación y la elaboración de representaciones visuales para comunicar los resultados.

El manuscrito original identifica esta estrategia como metodología Scrum y estructura el trabajo mediante un product backlog y varios sprints. No obstante, la evidencia disponible no permite demostrar una aplicación completa del marco. La Guía Scrum define responsabilidades específicas, eventos formales y artefactos relacionados con el objetivo del producto, el objetivo del sprint y la definición de terminado. También establece que Scrum funciona como una unidad integrada y que la aplicación aislada de algunos de sus componentes no equivale al uso completo del marco (Schwaber & Sutherland, 2020).

Por esta razón, la presente obra utiliza la expresión estrategia iterativa inspirada en Scrum. Esta denominación permite conservar el núcleo del proceso original, basado en planificación, desarrollo incremental, revisión y ajuste, sin afirmar que se documentaron formalmente todas las responsabilidades, reuniones y compromisos establecidos en la guía oficial.

La estrategia de desarrollo no constituye, por sí misma, la metodología de validación científica del prototipo. Su función se limita a organizar el trabajo técnico. La calidad del modelo depende de procedimientos adicionales relacionados con la definición de la variable objetivo, la preparación de los datos, la separación entre entrenamiento y evaluación, el análisis de las métricas y la fidelidad de las explicaciones.

Organización iterativa del desarrollo

El proceso puede reconstruirse mediante cuatro iteraciones funcionales. Cada una produjo un componente que sirvió como base para la siguiente etapa.

En la primera iteración se definió el alcance del artefacto. Se identificaron como componentes centrales el conjunto de datos, el algoritmo predictivo, la comparación entre registros, las visualizaciones y la presentación de explicaciones. También se estableció que el contexto demostrativo estaría relacionado con proveedores de servicios tecnológicos y escenarios hipotéticos de contratación pública.

En la segunda iteración se desarrolló el flujo de datos. Esta etapa comprendió la generación de observaciones sintéticas, la selección de variables, la codificación de la información categórica y la separación entre predictores y variable objetivo. El producto resultante fue un conjunto estructurado que podía ser procesado mediante bibliotecas de aprendizaje automático.

En la tercera iteración se incorporó el modelo de clasificación. Se configuró un bosque aleatorio, se dividieron los datos en conjuntos de entrenamiento y prueba y se generaron predicciones.

También se plantearon métricas de evaluación y una función destinada a comparar registros simulados.

En la cuarta iteración se añadieron visualizaciones y salidas textuales. Estas representaciones buscaban facilitar la comparación entre alternativas y mostrar los factores considerados. Sin embargo, la revisión editorial determinó que la explicación original no constituye todavía una técnica consolidada de inteligencia artificial explicable, debido a que compara valores promedio sin demostrar que esos factores determinaron efectivamente la predicción. En consecuencia, el módulo deberá fortalecerse mediante un procedimiento como importancia por permutación o valores SHAP.

Tabla 8

Organización iterativa del desarrollo del prototipo

Iteración	*Propósito*	*Actividades recuperadas del manuscrito*	*Producto generado*	*Ajuste editorial requerido*
1. Delimitación	Definir el problema y los componentes	Identificación de requisitos, algoritmo, interfaz, visualizaciones y pruebas	Especificación inicial	Diferenciar prueba de concepto e implementación real
2. Flujo de datos	Construir un conjunto procesable	Generación sintética, codificación, imputación y separación de variables	Base sintética estructurada	Documentar el origen y corregir la variable objetivo
3. Modelo predictivo	Entrenar y evaluar un clasificador	Bosque aleatorio, división 80/20, predicción y métricas	Modelo experimental	Incorporar estratificación, líneas de base y métricas balanceadas
4. Interpretación	Presentar comparaciones y explicaciones	Gráficos, función de comparación y salida textual	Interfaz demostrativa	Sustituir explicaciones aparentes por métodos fieles al modelo

Nota. *Las iteraciones reconstruyen el desarrollo tecnológico descrito en el manuscrito. No representan una validación institucional ni pruebas acreditadas con usuarios finales.*

La Tabla 8 sustituye la descripción extensa de sprints y permite conservar las actividades efectivamente identificables. Los detalles técnicos de cada iteración se desarrollarán en los apartados posteriores, evitando repeticiones.

Factibilidad operativa

La factibilidad operativa examina si el prototipo puede ser ejecutado, comprendido y revisado dentro de las condiciones definidas para una demostración académica. En esta etapa no se evalúa todavía su adopción por una institución pública, porque no se dispone de un escenario organizacional, grupo de usuarios, protocolo de uso ni instrumento de satisfacción documentado.

Desde el punto de vista demostrativo, el prototipo es operativamente viable porque su ejecución se realiza mediante un cuaderno computacional que integra código, resultados y visualizaciones. Esta estructura permite observar las etapas del procesamiento y repetir el procedimiento sin necesidad de desarrollar inicialmente una aplicación independiente.

La viabilidad operativa se encuentra condicionada por la claridad del flujo. El usuario debe poder identificar:

- Qué datos se utilizan;
- Qué representan las variables;
- Cuál es la naturaleza sintética de los registros;
- Qué transformación se aplicó;
- Qué modelo produjo la puntuación;
- Qué significado posee la salida;
- Y qué limitaciones deben considerarse.

La interfaz no debe conducir al usuario a interpretar la puntuación como una recomendación vinculante. Expresiones como “se recomienda elegir” deben sustituirse por formulaciones prudentes, como:

El modelo asignó una puntuación sintética superior a esta alternativa dentro del escenario simulado. El resultado requiere revisión humana y no representa una decisión de contratación.

La factibilidad operativa también depende de que el sistema permita detectar errores. Si un registro contiene valores ausentes, categorías desconocidas o combinaciones fuera de los rangos establecidos, el prototipo debería emitir una advertencia antes de generar una comparación. La producción automática de una puntuación no debe ocultar problemas en los datos de entrada.

El manuscrito afirma que se realizaron pruebas de usabilidad con usuarios reales y que se obtuvo retroalimentación mediante Google Colab. Sin embargo, no presenta el número de participantes, sus perfiles, instrumento, tareas, procedimiento, resultados ni consentimiento. En ausencia de esas evidencias, tales afirmaciones no deben incorporarse a la versión final. La evaluación con usuarios se mantendrá como una fase futura.

Para una implementación institucional, la factibilidad operativa requeriría:

- Identificar los perfiles de usuario;
- Definir permisos y responsabilidades;
- Elaborar protocolos de revisión;
- Capacitar a los responsables;
- Establecer mecanismos de corrección;
- Registrar las decisiones humanas;
- Y evaluar la comprensión de las explicaciones.

Estas condiciones exceden la prueba de concepto, pero deben reconocerse para evitar que la viabilidad técnica sea confundida con preparación operativa.

Factibilidad técnica

La factibilidad técnica se analiza en función de los recursos realmente utilizados y no de una infraestructura hipotética. El prototipo fue desarrollado mediante Python y bibliotecas de análisis de datos, aprendizaje automático y visualización. La ejecución se realizó en Google Colab, un servicio de cuadernos Jupyter alojados que no requiere configuración local inicial y ofrece acceso a recursos computacionales gratuitos, aunque estos recursos no son garantizados, ilimitados ni constantes (Google, s. f.).

El entorno utilizado resulta suficiente para una prueba de concepto con 2000 observaciones y un modelo de bosque aleatorio de complejidad moderada. El tamaño del conjunto no exige servidores de alto rendimiento, infraestructura distribuida ni unidades gráficas especializadas. Por ello, deben eliminarse las afirmaciones según las cuales el proyecto necesita procesar grandes volúmenes de datos o ejecutar algoritmos que requieren equipos de altas prestaciones.

La versión revisada debe diferenciar entre los recursos comprobados y aquellos mencionados solo como posibilidades futuras. Python, pandas, NumPy, scikit-learn y Matplotlib aparecen directamente en el código. Seaborn también se importa para visualización. En cambio, TensorFlow no se utiliza y debe eliminarse de la lista de requerimientos.

El escalado mediante StandardScaler se encuentra implementado, aunque los bosques aleatorios no dependen de la escala de las variables del mismo modo que otros algoritmos. Puede conservarse para mantener un flujo uniforme y facilitar futuras comparaciones con modelos sensibles a la escala, pero no debe afirmarse que sea indispensable para mejorar el rendimiento del bosque.

El prototipo tampoco necesita una base de datos SQL para ejecutar el flujo presentado. Los registros se generan directamente en un DataFrame. MySQL o SQL Server podrían incorporarse en una versión ampliada para almacenar información, pero no deben presentarse como componentes implementados.

Google Cloud AI Platform tampoco forma parte del prototipo mostrado. Su incorporación supondría un escenario de despliegue distinto, con nuevos costos, controles de seguridad y decisiones de arquitectura. Por consiguiente, debe retirarse de la tabla de componentes actuales.

Tabla 9

Recursos tecnológicos y situación dentro del prototipo

Recurso	*Función*	*Situación comprobada*	*Decisión editorial*
Python	Ejecución del código y articulación del flujo	Utilizado	Conservar
pandas	Organización y transformación tabular	Utilizado	Conservar
NumPy	Generación aleatoria y operaciones numéricas	Utilizado	Conservar
scikit-learn	Preparación, clasificación y evaluación	Utilizado	Conservar
Matplotlib	Generación de gráficos	Utilizado	Conservar
Seaborn	Configuración visual complementaria	Importado en el código	Conservar solo si se usa efectivamente
Google Colab	Entorno de ejecución del cuaderno	Utilizado	Conservar, indicando sus límites
TensorFlow	Redes neuronales y aprendizaje profundo	No utilizado	Eliminar
MySQL o SQL Server	Almacenamiento persistente	No implementado	Presentar únicamente como posibilidad futura
Google Cloud AI Platform	Entrenamiento o despliegue administrado	No implementado	Eliminar del prototipo actual
Equipo de alto rendimiento	Procesamiento intensivo	No requerido para 2000 registros	Eliminar como requisito

Nota. *La tabla diferencia los recursos efectivamente utilizados de aquellos mencionados en el manuscrito sin evidencia de implementación.*

La factibilidad técnica de una prueba de concepto no equivale a la de un sistema productivo. Una plataforma institucional necesitaría considerar disponibilidad, continuidad del servicio, autenticación, cifrado, copias de seguridad, control de versiones, monitoreo, integración con fuentes oficiales y respuesta ante incidentes.

Factibilidad económica

El análisis económico original se limita a presentar valores aproximados para una computadora, utilitarios, conexión a internet y software gratuito. Estas cifras no permiten calcular la factibilidad económica porque no incorporan fecha, proveedor, periodicidad, personal, mantenimiento, seguridad, almacenamiento ni soporte.

Para la etapa demostrativa, la inversión tecnológica directa puede mantenerse limitada porque el prototipo utiliza herramientas disponibles sin costo inicial de licencia y se ejecuta en un entorno alojado gratuito. Sin embargo, la ausencia de una tarifa de software no implica costo total cero. El desarrollo requiere tiempo profesional, revisión metodológica, documentación, pruebas, corrección del código y mantenimiento.

Google Colab puede utilizarse gratuitamente, pero sus recursos poseen límites variables y no están garantizados. Por ello, no debería utilizarse como base para estimar los costos de una plataforma institucional sin considerar modalidades pagadas o infraestructura alternativa.

Escenario demostrativo

Comprende:

- Uso de un equipo existente;
- Herramientas de código abierto;
- Ejecución en un cuaderno alojado;
- Almacenamiento limitado;
- Desarrollo y revisión por el equipo autoral.

En este escenario no resulta necesario presentar montos monetarios cuando no existen comprobantes o registros de costos. Es suficiente indicar que se aprovecharon recursos tecnológicos disponibles.

Escenario de implementación futura

Comprendería:

- Personal de ciencia de datos;
- Especialistas en contratación pública;
- Desarrollo de software;
- Infraestructura;
- Almacenamiento y procesamiento;
- Seguridad;
- Auditoría;
- Mantenimiento;
- Capacitación;
- Pruebas con usuarios;
- Y actualización periódica del modelo.

Este segundo escenario no puede valorarse con los datos del manuscrito y debe quedar fuera de cualquier afirmación sobre retorno de inversión. Tampoco puede sostenerse que el sistema generará beneficios económicos sostenibles o reducirá el gasto público, debido a que esos resultados no fueron medidos.

Factibilidad ética y organizacional

La viabilidad del prototipo no depende únicamente del código. Una herramienta relacionada con contratación pública necesita examinar la legitimidad de sus variables, la posibilidad de sesgos, el uso de datos personales, la trazabilidad de los resultados y la responsabilidad de quienes intervienen.

En la etapa sintética, el principal criterio ético consiste en evitar que los ejemplos produzcan afirmaciones sobre empresas reales. Por ello, los nombres de IBM, Tata, Sonda, Akross, Novatech, Kruger y otros proveedores deben sustituirse por identificadores ficticios.

La utilización de nombres reales en un conjunto generado aleatoriamente puede asociar a una empresa con puntuaciones, fortalezas o debilidades que no proceden de información verificable. La anonimización no resulta suficiente porque las entidades no fueron anonimizadas desde una base real; fueron incorporadas como etiquetas decorativas dentro de una simulación. La corrección apropiada consiste en reemplazarlas.

La viabilidad organizacional futura requeriría determinar quién:

- Autoriza el uso del sistema;
- Valida las variables;
- Revisa la calidad de los datos;
- Interpreta la explicación;
- Registra una discrepancia;
- Actualiza el modelo;
- Y responde por la decisión final.

Estas responsabilidades no se resuelven mediante Scrum ni mediante la selección del algoritmo. Deben formar parte de la gobernanza institucional.

Síntesis de factibilidad

La estrategia iterativa permitió desarrollar progresivamente los componentes del prototipo y comprobar su ejecución dentro de un entorno accesible. El análisis de factibilidad muestra que la propuesta puede sostenerse como prueba de concepto, pero no como sistema listo para operar en contratación pública.

Tabla 10

Síntesis de la factibilidad del prototipo

Dimensión	*Situación actual*	*Dictamen*
Operativa	Flujo ejecutable en un cuaderno, sin evaluación documentada con usuarios	Viable como demostración
Técnica	Herramientas disponibles y capacidad suficiente para 2000 observaciones	Viable como prueba de concepto
Económica	Uso de recursos disponibles, sin registro integral de costos	No cuantificable con la información existente
Ética	Datos sintéticos, pero nombres empresariales reales y riesgo de interpretación	Viable después de sustituir identificadores y advertir limitaciones
Institucional	Sin integración con entidades, datos oficiales o procedimientos reales	No validada
Científica	Permite demostración técnica, pero requiere corregir la etiqueta objetivo	Condicionada a la reconstrucción metodológica

Nota. *La viabilidad demostrativa no implica preparación para una implementación productiva o institucional.*

Su consolidación metodológica depende ahora de transparentar la procedencia de los datos, sustituir los nombres empresariales, documentar las reglas de simulación y establecer mecanismos de trazabilidad. Estos elementos se desarrollan en el apartado siguiente, dedicado al origen, naturaleza y gobernanza de los datos.

Origen, naturaleza y gobernanza de los datos

Los datos constituyen el componente sobre el cual se construyen las predicciones, puntuaciones y explicaciones del prototipo. Por esta razón, su procedencia, composición y tratamiento deben describirse con el mismo nivel de precisión que el algoritmo utilizado. La transparencia metodológica no se limita a indicar el número de registros, sino que exige documentar cómo fueron obtenidos, qué representa cada observación, qué relaciones fueron incorporadas durante su generación y cuáles son los usos permitidos y no permitidos.

En la versión revisada de la obra, el conjunto empleado se define expresamente como información sintética generada mediante programación. Los registros no proceden del Portal de Contratación Pública, del Registro Único de Proveedores, de expedientes administrativos, de contratos reales ni

de una entidad participante. Tampoco fueron construidos mediante la anonimización de una base institucional. Se trata de observaciones artificiales creadas para demostrar el flujo técnico del prototipo.

Esta precisión resuelve una contradicción del manuscrito original. En algunas secciones se afirma que fueron integrados registros oficiales de empresas tecnológicas, propuestas presentadas, contratos adjudicados y datos de desempeño histórico. Sin embargo, el código utiliza funciones aleatorias de NumPy para generar las características y asigna de manera independiente la etiqueta que identifica el supuesto resultado de la licitación.

En consecuencia, la única procedencia que puede sostenerse documentalmente es la generación sintética. Esta decisión no obliga a descartar el prototipo. Permite conservarlo como demostración metodológica, siempre que sus resultados no se interpreten como evidencia sobre proveedores reales ni como predicciones válidas de adjudicación.

Naturaleza sintética del conjunto

Un conjunto sintético está formado por observaciones creadas mediante reglas, distribuciones estadísticas o modelos generativos, en lugar de haber sido recogidas directamente de personas, organizaciones o procesos reales. Este tipo de información puede utilizarse para probar código, demostrar arquitecturas, desarrollar ejercicios educativos y explorar procedimientos sin exponer registros institucionales.

No obstante, la información sintética no reproduce automáticamente la complejidad del fenómeno que pretende representar. Las distribuciones, dependencias, ausencias, errores y excepciones presentes en datos reales solo aparecen si fueron introducidas deliberadamente durante la simulación. Por ello, un conjunto sintético puede ser útil para comprobar el funcionamiento técnico y, al mismo tiempo, resultar insuficiente para validar la utilidad práctica del modelo (Jordon et al., 2022).

En el prototipo, cada observación representa un caso hipotético relacionado con un proveedor de servicios tecnológicos. Las variables describen atributos como experiencia, proyectos completados, tiempo de respuesta, costo, satisfacción, certificaciones, personal calificado, cobertura, soporte, especialización y cumplimiento de acuerdos de servicio.

Los rangos definidos en el manuscrito permiten crear diversidad numérica entre los registros, pero no proceden de un estudio de mercado ni de estadísticas oficiales. En consecuencia, deben interpretarse como intervalos demostrativos seleccionados para ejecutar el código y observar el comportamiento de la arquitectura.

Finalidad del conjunto

El conjunto sintético tiene cuatro finalidades delimitadas:

- Comprobar que las funciones de generación y procesamiento se ejecutan correctamente;
- Entrenar y evaluar un clasificador dentro de un entorno controlado;
- Producir ejemplos de comparación entre registros hipotéticos;
- Y demostrar cómo una puntuación puede acompañarse de información explicativa.

No tiene como finalidad:

- Representar el universo de proveedores ecuatorianos;
- Estimar la probabilidad real de adjudicación;
- Identificar empresas más competitivas;
- Establecer criterios oficiales de evaluación;
- Comparar organizaciones existentes;
- Ni apoyar directamente una decisión administrativa.

Esta separación entre usos previstos y usos excluidos forma parte de la gobernanza del conjunto. Gebru et al. (2021) proponen que los conjuntos de datos se acompañen de documentación sobre su

motivación, composición, proceso de creación, usos recomendados, limitaciones y condiciones de mantenimiento. El objetivo es que quienes reutilicen los datos puedan comprender su alcance y determinar si resultan adecuados para una tarea específica.

Unidad de observación

La unidad de observación corresponde a un registro sintético de características de un proveedor tecnológico hipotético. Cada fila representa una combinación artificial de valores y no una empresa, una oferta o un contrato real.

Esta unidad debe diferenciarse de otras posibles unidades de análisis:

- Una empresa registrada;
- Una oferta presentada;
- Un procedimiento de contratación;
- Un contrato adjudicado;
- Una ejecución contractual;
- Una interacción entre proveedor y entidad.

El prototipo no distingue actualmente estas estructuras. Por tanto, la expresión "registro de proveedor" se utiliza con una finalidad exclusivamente demostrativa.

En una futura investigación con datos reales, sería necesario definir si cada fila representa una oferta, un proveedor dentro de un procedimiento, un contrato o un evento de ejecución. Esta decisión modificaría las variables, el modelo y la interpretación de la salida.

Sustitución de nombres empresariales

El conjunto actual utiliza nombres de empresas identificables, entre ellas IBM, Tata, Sonda, Akross, Novatech y Kruger. Los valores asociados con esos nombres no proceden de información comprobada, sino de distribuciones aleatorias.

En la versión publicable, estas denominaciones deben reemplazarse por:

- Proveedor A.
- Proveedor B.
- Proveedor C.
- Proveedor D.
- Proveedor E.
- Proveedor F.
- Proveedor G.

También pueden utilizarse identificadores como PROV_001, PROV_002 y PROV_003.

El identificador del proveedor cumplirá una función de organización y visualización, pero no debe ser incorporado como variable predictora. Codificar el nombre mediante one-hot encoding permitiría que el modelo aprendiera diferencias artificiales asociadas con la etiqueta del proveedor, aunque esa etiqueta no represente una característica técnica.

Por consiguiente, la variable proveedor debe conservarse en la tabla original para identificar los ejemplos, pero excluirse del conjunto de predictores utilizado durante el entrenamiento. Esta decisión evita que el sistema asigne importancia predictiva a una denominación ficticia.

Revisión de la variable objetivo

El problema metodológico más importante del conjunto actual se encuentra en la variable ganador_licitacion. El código la genera mediante una selección aleatoria con una proporción aproximada de 70 % para la clase cero y 30 % para la clase uno. Su valor no depende de la experiencia, del cumplimiento, del costo ni de ninguna otra característica.

En esas condiciones, el modelo recibe una etiqueta sin relación programada con los predictores. El algoritmo puede encontrar asociaciones accidentales dentro de una muestra particular, pero estas

no representan un patrón que pueda generalizarse. Por tanto, la salida no debe denominarse "probabilidad de ganar una licitación".

Para conservar el ejercicio de clasificación, la variable debe sustituirse por:

"clase_compatibilidad_sintetica"

Esta variable representará una categoría creada con fines demostrativos a partir de una regla explícita. La regla combinará variables normalizadas cuyos sentidos hayan sido previamente definidos. Los valores elevados de experiencia, cumplimiento, satisfacción, especialización y capacidad podrán incrementar la puntuación, mientras que valores elevados de costo y tiempo de respuesta podrán reducirla.

La regla no pretende reproducir criterios oficiales de contratación. Su finalidad es introducir una relación controlada entre los predictores y la etiqueta, de manera que pueda evaluarse si el algoritmo recupera patrones que fueron incorporados deliberadamente durante la simulación.

El procedimiento detallado, la dirección de cada variable y la construcción de la clase se presentarán en el apartado 3.4. La etiqueta deberá acompañarse siempre de la expresión "sintética" para evitar que se confunda con un resultado institucional.

Composición del conjunto

El conjunto conservará 2000 observaciones como tamaño demostrativo. Esta cantidad resulta suficiente para ejecutar el flujo de procesamiento y producir particiones de entrenamiento y evaluación, pero no debe describirse como una muestra representativa.

Las variables se organizarán en cuatro grupos:

1. Identificación ficticia: código del proveedor.

2. Capacidad y experiencia: años de experiencia, proyectos completados, personal y certificaciones.
3. Condiciones operativas y económicas: tiempo de respuesta, costo por hora, cobertura y soporte.
4. Desempeño sintético: satisfacción, especialización y cumplimiento de acuerdos de servicio.

La variable objetivo se almacenará por separado y se generará después de calcular la puntuación sintética. Esta separación permite distinguir los datos utilizados para describir el caso de la etiqueta construida para el ejercicio de clasificación.

Procedencia y linaje

El linaje de datos describe la secuencia mediante la cual la información se crea y transforma. En el prototipo deberá registrarse mediante las siguientes etapas:

Código generador → conjunto sintético original → validación de rangos → construcción de la variable objetivo → conjunto procesado → partición de entrenamiento y prueba → predicciones y explicaciones

El conjunto original debe conservarse sin modificaciones después de su generación. Las operaciones de imputación, transformación o escalado se aplicarán sobre una copia o mediante un flujo reproducible. Esta práctica permite comparar la versión inicial con la procesada y reconstruir cada modificación.

La semilla aleatoria debe registrarse junto con la versión del código. El manuscrito utiliza np.random.seed(42), lo que permite repetir la secuencia generada mientras se mantengan las mismas instrucciones, versiones compatibles y orden de ejecución. La semilla favorece la reproducibilidad, pero no demuestra que los datos sean realistas.

Documentación del conjunto

Para acompañar el código, se incorporará una ficha técnica del conjunto sintético. Esta ficha adaptará el principio de las hojas de datos o datasheets, cuyo propósito es documentar la motivación, composición, creación, mantenimiento y usos apropiados de un conjunto (Gebru et al., 2021).

Tabla 11

Ficha de gobernanza del conjunto de datos sintético

Elemento	*Descripción*
Nombre del conjunto	Datos sintéticos de proveedores tecnológicos
Identificador de versión	DSPT-v1.0
Finalidad	Demostrar el flujo de preparación, clasificación, comparación y explicación
Procedencia	Generación programática mediante Python y NumPy
Número de observaciones	2000 registros sintéticos
Unidad de observación	Registro hipotético de características de un proveedor tecnológico
Identificadores	Proveedor A a Proveedor G o códigos ficticios equivalentes
Información real	No contiene ofertas, contratos, empresas evaluadas ni expedientes reales
Datos personales	No se incorporan datos personales
Variable objetivo	Clase de compatibilidad sintética construida mediante una regla documentada
Uso permitido	Docencia, demostración técnica, prueba del código y análisis metodológico
Uso no permitido	Adjudicar contratos, evaluar empresas reales o estimar probabilidades institucionales
Semilla de generación	42
Responsable de documentación	Equipo autoral
Control de versiones	Registro de cambios en código, datos, variables y parámetros
Limitaciones principales	Ausencia de representatividad, relaciones construidas y falta de validación externa

Nota. *La ficha documenta la naturaleza y los límites del conjunto. No certifica que los datos reproduzcan el comportamiento de proveedores o procedimientos reales.*

Control de calidad

La calidad del conjunto sintético no debe evaluarse mediante los mismos criterios utilizados para una base recolectada. No existen errores de captura originales, pero pueden presentarse errores de

programación, rangos incoherentes, combinaciones imposibles o distribuciones insuficientemente justificadas.

Antes de entrenar el modelo deben comprobarse:

- Cantidad esperada de registros;
- Ausencia de identificadores empresariales reales;
- Nombres y tipos de variables;
- Valores mínimos y máximos;
- Frecuencia de las variables binarias;
- Distribución de la clase sintética;
- Ausencia de registros duplicados no intencionales;
- Valores faltantes incorporados deliberadamente;
- Y coherencia de las relaciones introducidas.

La imputación no debe aplicarse si la generación no produce valores ausentes. El código actual utiliza SimpleImputer, aunque los valores se crean completos. Este recurso puede conservarse como demostración de un flujo preparado para manejar ausencias, pero deberá aclararse que, en la versión inicial, no se generaron valores faltantes de manera natural.

Una alternativa más coherente consistiría en introducir un porcentaje pequeño y controlado de valores ausentes exclusivamente para demostrar el procedimiento de imputación. Esa decisión deberá documentarse, indicando las variables afectadas, la proporción utilizada y la razón metodológica.

Versionamiento

Cada modificación del conjunto debe producir una nueva versión. El registro mínimo incluirá:

- Nombre y versión;

- Fecha de generación;
- Semilla;
- Número de observaciones;
- Variables incluidas;
- Rangos;
- Fórmula de la clase sintética;
- Transformaciones;
- Y versión del código.

La versión utilizada para producir las tablas y figuras del libro debe mantenerse congelada. Si se modifica la fórmula, el balance de clases o los rangos, deberán recalcularse las métricas y visualizaciones.

El versionamiento evita que diferentes ejecuciones produzcan resultados que luego se presenten como pertenecientes a un mismo modelo. También permite identificar qué conjunto originó cada matriz de confusión, puntuación o explicación.

Acceso y conservación

El conjunto puede incorporarse al repositorio complementario del libro junto con el código necesario para reproducirlo. Debido a que no contiene información personal ni registros institucionales, su publicación no presenta los mismos riesgos de confidencialidad que una base real.

No obstante, deberá incluirse un archivo de advertencia que declare:

- Que los datos son sintéticos;
- Que los proveedores son ficticios;
- Que las variables no son criterios oficiales;
- Que las puntuaciones no representan adjudicaciones;

- Y que el material no debe utilizarse para tomar decisiones públicas.

La conservación deberá incluir dos versiones:

- El conjunto sintético original;
- Y el conjunto procesado utilizado por el modelo.

También deben conservarse el cuaderno ejecutable, los requisitos de software y una descripción de los pasos necesarios para reproducir los resultados.

Responsabilidades de gobernanza

La gobernanza exige definir responsabilidades, aunque el prototipo sea académico. Para esta obra se reconocen tres funciones:

Tabla 12

Funciones de la Responsabilidad de Gobernanza

Función	*Responsabilidad*
Responsable del conjunto	Documentar la generación, composición, versiones y limitaciones
Responsable del modelado	Verificar transformaciones, particiones, parámetros y métricas
Revisor metodológico	Comprobar coherencia entre datos, resultados, explicaciones y afirmaciones publicadas

Nota. *Estas funciones pueden ser asumidas por integrantes del equipo autoral, pero deben mantenerse diferenciadas conceptualmente.*

La misma persona que desarrolla el código puede revisar su funcionamiento, aunque resulta recomendable que otra persona examine las decisiones metodológicas y los resultados antes de su publicación.

El NIST AI RMF establece que la gobernanza debe integrar políticas, responsabilidades, documentación y mecanismos de supervisión durante todo el ciclo de vida del sistema. La gestión

no debe iniciarse únicamente después de entrenar el modelo, porque numerosas fuentes de riesgo se originan durante la definición del problema y la construcción de los datos (NIST, 2023).

Limitaciones derivadas de los datos

La primera limitación es la ausencia de representatividad. Los registros no permiten describir a los proveedores tecnológicos de Ecuador ni las condiciones de las licitaciones públicas.

La segunda corresponde a la construcción artificial de las relaciones. El comportamiento predictivo dependerá de la fórmula utilizada para crear la clase sintética. Si el modelo alcanza un rendimiento elevado, esto indicará que pudo recuperar una estructura introducida por los autores, no que haya descubierto reglas reales de adjudicación.

La tercera se relaciona con la simplificación. Las variables numéricas no representan documentos, requisitos habilitantes, especificaciones técnicas, condiciones legales ni particularidades de cada procedimiento.

La cuarta limitación es la ausencia de variación temporal. El conjunto no incorpora cambios normativos, evolución del mercado, modificaciones empresariales o transformaciones en las necesidades institucionales.

La quinta se encuentra en la falta de validación externa. El conjunto no ha sido comparado con distribuciones provenientes de fuentes oficiales ni revisado por especialistas para determinar el realismo de sus rangos.

Estas limitaciones deberán mantenerse visibles cuando se presenten los resultados. La gobernanza no pretende eliminar las restricciones del conjunto sintético, sino documentarlas para evitar interpretaciones improcedentes.

La reconstrucción del origen y la gobernanza de los datos permite conservar el ejercicio computacional sin confundirlo con una investigación empírica. El conjunto se reconoce como

sintético, reproducible y limitado a la demostración del prototipo. Los nombres empresariales se reemplazan por identificadores ficticios, la identidad del proveedor deja de utilizarse como predictor y la variable aleatoria de adjudicación se sustituye por una clase sintética construida mediante una regla transparente.

Variables, generación y preparación del conjunto sintético

La construcción del conjunto sintético requiere definir con precisión qué representa cada variable, cuál es su escala, qué dirección adopta dentro del ejercicio y qué función cumple en el modelo. Esta operacionalización permite evitar que los atributos sean tratados como números aislados y facilita la reproducción de las decisiones incorporadas durante la simulación.

Las variables seleccionadas conservan el núcleo del manuscrito original. Representan características hipotéticas relacionadas con experiencia, capacidad operativa, condiciones económicas, calidad percibida y cumplimiento. Sin embargo, ninguna de ellas debe interpretarse como criterio oficial de evaluación o adjudicación. Su inclusión responde exclusivamente a la necesidad de construir un escenario tabular que permita demostrar el funcionamiento del prototipo.

El conjunto se compone de 2000 observaciones generadas programáticamente. Cada observación representa un caso hipotético y se identifica mediante un código ficticio. La generación utiliza una semilla aleatoria fija para favorecer la repetición del ejercicio, aunque esta reproducibilidad no convierte los registros en representaciones realistas de proveedores o licitaciones.

Operacionalización de las variables

Tabla 13

Operacionalización de las variables del conjunto sintético

Variable	*Definición operacional dentro del prototipo*	*Tipo y rango sintético*	*Dirección dentro de la regla*	*Peso propuesto*

experiencia_anios	Años hipotéticos de actividad o experiencia acumulada	Entera: 1 a 30	Mayor valor incrementa la puntuación	0,12
proyectos_completados	Cantidad hipotética de proyectos concluidos	Entera: 5 a 200	Mayor valor incrementa la puntuación	0,10
tiempo_respuesta_horas	Horas estimadas para atender un requerimiento	Continua: 1 a 72	Menor valor incrementa la puntuación	0,08
costo_por_hora	Valor económico hipotético por hora de servicio	Continua: USD 40 a 250	Menor valor incrementa la puntuación	0,12
satisfaccion_cliente	Valoración sintética de satisfacción	Continua: 2 a 5	Mayor valor incrementa la puntuación	0,12
certificaciones	Cantidad hipotética de certificaciones relacionadas	Entera: 1 a 15	Mayor valor incrementa la puntuación	0,08
personal_calificado	Cantidad hipotética de personas calificadas	Entera: 10 a 500	Mayor valor incrementa la puntuación	0,08
cobertura_nacional	Disponibilidad hipotética de cobertura nacional	Binaria: 0 = no; 1 = sí	La presencia incrementa la puntuación	0,06
soporte_24_7	Disponibilidad hipotética de soporte continuo	Binaria: 0 = no; 1 = sí	La presencia incrementa la puntuación	0,06
especializacion_sector_publico	Grado sintético de especialización en el sector público	Continua: 0 a 1	Mayor valor incrementa la puntuación	0,08
cumplimiento_sla	Proporción sintética de cumplimiento de acuerdos de servicio	Continua: 0,70 a 1,00	Mayor valor incrementa la puntuación	0,10
clase_compatibilidad_sintetica	Categoría construida a partir de una regla explícita	Binaria: 0 = compatibilidad ordinaria; 1 = compatibilidad alta	Variable objetivo	No aplica

Nota. *Los rangos, sentidos y pesos fueron definidos con fines demostrativos. No proceden de criterios oficiales ni deben aplicarse a procedimientos reales de contratación.*

Los pesos suman uno y expresan la contribución definida por los autores dentro de la regla de simulación. No fueron obtenidos mediante análisis de contratos reales, consulta a especialistas o estimación estadística. Su publicación tiene como finalidad transparentar la forma en que se construyó la variable objetivo y permitir que otros lectores modifiquen o cuestionen la regla.

La dirección de cada variable debe establecerse antes de construir la puntuación. El código original compara los valores suponiendo que una magnitud mayor siempre es favorable. Esta interpretación resulta inadecuada para el costo y el tiempo de respuesta, porque sus valores elevados no deben presentarse automáticamente como fortalezas.

Normalización de las variables

Las variables se expresan mediante escalas diferentes. La experiencia se mide en años; el costo, en unidades monetarias; la satisfacción, en una escala de cinco puntos; y el cumplimiento, mediante una proporción. Para combinarlas en una regla sintética deben transformarse a una escala común comprendida entre cero y uno.

En las variables cuya dirección positiva se asocia con valores mayores se utiliza la normalización mínimo-máximo:

$$z_j = \frac{x_{j-\min(x_j)}}{max(x_j) - \min(x_j)}$$

donde x_j representa el valor original de la variable y z_j su versión normalizada. Para el costo y el tiempo de respuesta se invierte la escala:

$$z_j^{inv} = 1 - \frac{x_{j-\min(x_j)}}{max(x_j) - \min(x_j)}$$

De esta manera, un costo o tiempo menor produce un valor normalizado más alto dentro de la regla. La transformación expresa únicamente la lógica diseñada para la simulación y no establece que el menor precio deba prevalecer sobre otros criterios en una contratación real.

Las variables binarias se mantienen con valores cero y uno. La especialización ya se genera dentro de ese intervalo. El cumplimiento del acuerdo de servicio se transforma desde su rango de 0,70 a 1,00 hasta una escala de cero a uno.

Construcción de la puntuación sintética

La etiqueta original “ganador_licitacion debe eliminarse. En su lugar se propone una puntuación de compatibilidad sintética calculada mediante la siguiente expresión:

$$P= 0{,}12E+0{,}10Pr+0{,}08T+0{,}12C+0{,}12S+0{,}08Ce+0{,}08Pe+0{,}06Co+0{,}06So+0{,}08Es+0{,}10Cu+\varepsilon$$

donde:

- E representa la experiencia normalizada;
- Pr, los proyectos completados;
- T, el tiempo de respuesta con dirección invertida;
- C, el costo por hora con dirección invertida;
- S, la satisfacción;
- Ce, las certificaciones;
- Pe, el personal calificado;
- Co, la cobertura nacional;
- So, el soporte continuo;
- Es, la especialización en el sector público;
- Cu, el cumplimiento del acuerdo de servicio;
- y ε, un componente aleatorio de baja magnitud.

El componente aleatorio puede generarse mediante una distribución normal con media cero y desviación estándar de 0,03. Su finalidad es evitar que la etiqueta sea una reproducción completamente determinista de la suma ponderada. Después de incorporarlo, la puntuación debe limitarse al intervalo entre cero y uno.

Secuencia de generación

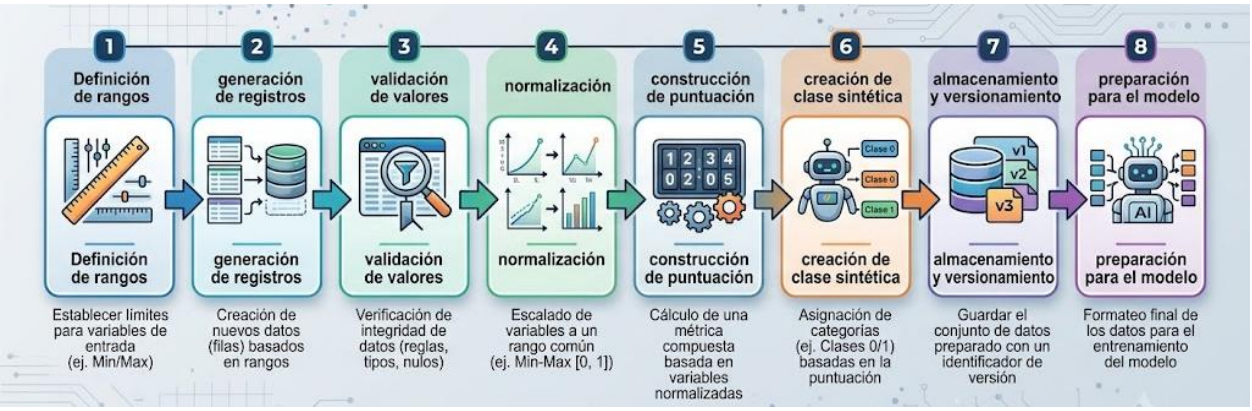
La generación del conjunto se organiza mediante la siguiente secuencia:

- Establecer la semilla aleatoria en 42.
- Crear 2000 identificadores sintéticos.
- Generar los valores dentro de los rangos establecidos.
- Comprobar los tipos, límites y frecuencias.
- Normalizar las variables para construir la regla.
- Invertir la dirección del costo y el tiempo.
- Calcular la puntuación ponderada.
- Incorporar el componente aleatorio controlado.
- Determinar el percentil 70.
- Crear la clase de compatibilidad sintética.
- Eliminar las variables auxiliares utilizadas para construir la etiqueta.
- Guardar una versión original e inalterada del conjunto.

Este procedimiento permite que la variable objetivo sea reproducible y metodológicamente comprensible. El rendimiento posterior del modelo indicará si el clasificador pudo recuperar la estructura introducida durante la simulación. No demostrará que el algoritmo haya descubierto relaciones reales de contratación.

Figura 8

Flujo de generación y preparación del conjunto de datos sintético



Nota. *Elaboración propia. El flujo representa un procedimiento demostrativo y no la extracción o transformación de registros oficiales.*

Validación inicial del conjunto

Después de la generación deben verificarse, como mínimo, los siguientes elementos:

- Que existan exactamente 2000 observaciones;
- Que cada variable conserve el tipo esperado;
- Que los valores permanezcan dentro de sus rangos;
- Que no aparezcan nombres de empresas reales;
- Que la clase positiva represente aproximadamente el 30 %;
- Que la puntuación sintética no permanezca entre los predictores;
- Que no existan duplicados completos producidos involuntariamente;
- Y que la semilla permita repetir el conjunto.

La distribución de la clase debe informarse mediante frecuencias absolutas y porcentajes. También conviene presentar estadísticas descriptivas de las variables numéricas, incluyendo media, desviación estándar, mínimo, mediana y máximo.

Tratamiento de valores ausentes

La generación original produce registros completos, por lo que la imputación no interviene realmente en el conjunto inicial. El uso de SimpleImputer puede mantenerse dentro del flujo como una medida preventiva, pero debe aclararse que no modifica los valores cuando no existen ausencias.

Para las variables continuas se recomienda la mediana en lugar de la media, porque reduce la influencia de valores extremos. Para las variables binarias puede emplearse la categoría más frecuente. La documentación oficial de scikit-learn establece que SimpleImputer permite

completar valores ausentes mediante estadísticas simples calculadas a partir de los datos disponibles.

Las estadísticas de imputación deben calcularse utilizando únicamente el conjunto de entrenamiento. Calcularlas con la base completa antes de separar los datos permitiría que información del conjunto de prueba influyera en la preparación del modelo.

No es recomendable introducir valores ausentes artificialmente solo para justificar el uso del imputador, salvo que esa decisión forme parte explícita de un experimento de robustez. En la versión principal, el componente puede conservarse como salvaguarda del flujo sin afirmar que existían datos incompletos.

Tratamiento de variables categóricas

La variable que identifica al proveedor no debe codificarse mediante one-hot encoding, porque su función es exclusivamente descriptiva. Permitir que el modelo utilice el identificador ficticio introduciría diferencias carentes de contenido técnico.

Las variables cobertura_nacional y soporte_24_7 ya están representadas mediante cero y uno, por lo que no necesitan una codificación adicional.

Si en una versión futura se incorporan categorías con significado, como tipo de servicio o región, podrán transformarse mediante OneHotEncoder. Esta herramienta genera columnas binarias para las categorías observadas y puede configurarse para manejar categorías desconocidas en nuevos datos.

Escalado de las características

El bosque aleatorio utiliza reglas de división construidas a partir del orden de los valores y no requiere que todas las variables tengan la misma escala. Por tanto, StandardScaler no debe presentarse como una condición necesaria para mejorar su rendimiento.

El escalado sí puede conservarse en la línea de base de regresión logística, porque este modelo se beneficia de variables expresadas en escalas comparables. StandardScaler transforma las características restando la media y dividiendo por la desviación estándar calculadas en el conjunto utilizado para su ajuste.

La versión revisada utilizará dos flujos:

- Bosque aleatorio: imputación preventiva y clasificador, sin escalado obligatorio.
- Regresión logística de referencia: imputación, escalado y clasificador.

Esta separación permite comparar modelos sin aplicar automáticamente la misma transformación cuando no resulta necesaria.

Uso de un flujo de procesamiento

Las transformaciones y el clasificador deben integrarse mediante un Pipeline. Este recurso aplica secuencialmente las operaciones y reduce el riesgo de que la preparación se realice con información perteneciente al conjunto de prueba. La documentación oficial de scikit-learn define Pipeline como una secuencia de transformadores seguida por un predictor final.

Cuando existan grupos de variables que requieran tratamientos diferentes, puede emplearse ColumnTransformer, que permite aplicar transformaciones específicas a subconjuntos de columnas y unir posteriormente sus resultados.

Scikit-learn ofrece una interfaz integrada para preparar datos, entrenar modelos y evaluar resultados, y ha sido ampliamente utilizado para construir flujos reproducibles de aprendizaje automático en Python (Pedregosa et al., 2011).

En el prototipo actual, la estructura puede simplificarse porque las variables predictoras son numéricas o binarias. No obstante, encapsular el procesamiento dentro de un flujo favorece la reproducibilidad y facilita futuras ampliaciones.

Registro de transformaciones

Cada operación debe quedar registrada en una ficha de procesamiento:

Tabla 14

Registro de preparación del conjunto sintético

Operación	*Variables afectadas*	*Criterio*	*Momento de aplicación*
Validación de rangos	Todas	Comprobar límites y tipos definidos	Antes de la partición
Exclusión del identificador	proveedor_id	No posee significado predictivo	Antes del entrenamiento
Imputación preventiva	Variables continuas y binarias	Mediana o categoría más frecuente	Ajustada solo con entrenamiento
Escalado	Variables continuas de la regresión logística	Media cero y desviación estándar unitaria	Ajustado solo con entrenamiento
Construcción de la clase	Variables normalizadas	Regla ponderada transparente	Durante la generación sintética
Eliminación de auxiliares	Puntuación y variables normalizadas temporales	Evitar fuga de información	Antes del entrenamiento
Versionamiento	Conjunto original y procesado	Conservar trazabilidad	En cada modificación

Nota. *El registro permite reconstruir las transformaciones y distinguir las operaciones de generación de aquellas aplicadas durante el entrenamiento.*

Código y reproducibilidad

El cuerpo del libro no debe contener seis páginas consecutivas de capturas del código. Se recomienda:

- Presentar únicamente el flujo metodológico;
- Incluir fragmentos breves cuando sean indispensables;
- Trasladar el código completo a un apéndice;
- Y publicar el cuaderno ejecutable en un repositorio asociado con la obra.

El repositorio debe incluir:

- Archivo de requisitos;
- Versión de python;

- Versiones de las bibliotecas;
- Código generador;
- Conjunto sintético congelado;
- Modelo de procesamiento;
- Instrucciones de ejecución;
- Advertencia de uso;
- Y registro de cambios.

La documentación de scikit-learn permite identificar los parámetros y comportamientos de las funciones utilizadas, pero la reproducibilidad del libro requiere además congelar las versiones efectivamente empleadas. Las interfaces y valores predeterminados pueden cambiar entre versiones, por lo que no basta con indicar únicamente "Python 3.x".

La operacionalización y reconstrucción del conjunto sintético permiten corregir el problema central del código original sin eliminar su lógica fundamental. Se conservan las variables, los rangos, la generación programática y la estructura de clasificación, pero la etiqueta aleatoria se sustituye por una clase construida mediante una regla pública y reproducible.

La preparación separa la información descriptiva de los predictores, evita utilizar el nombre del proveedor como característica, aplica las transformaciones únicamente dentro del flujo de entrenamiento y conserva un registro de cada operación. Sobre este conjunto metodológicamente delimitado, el apartado siguiente presenta el diseño del bosque aleatorio, la línea de base, la partición de los datos y los criterios utilizados para evaluar el comportamiento predictivo.

Diseño, entrenamiento y evaluación del modelo predictivo

El componente predictivo del prototipo se estructura como un problema de clasificación supervisada binaria. Las variables descritas en el apartado anterior constituyen los predictores, mientras que la clase_compatibilidad_sintetica funciona como variable objetivo. El propósito del

modelo consiste en identificar la relación programada entre las características de los registros y la categoría construida mediante la regla sintética.

Esta formulación representa un ejercicio controlado. El clasificador no aprende criterios reales de adjudicación, debido a que la variable objetivo no procede de decisiones institucionales. El desempeño alcanzado permite determinar si el algoritmo recupera una estructura incorporada deliberadamente durante la generación de los datos, pero no demuestra que pueda generalizar hacia proveedores, ofertas o procedimientos reales.

La evaluación se diseña para responder a cuatro preguntas:

1. ¿El modelo supera una estrategia que ignora las variables?
2. ¿Su comportamiento es superior al de una alternativa más sencilla?
3. ¿Identifica de manera adecuada tanto la clase mayoritaria como la minoritaria?
4. ¿Sus resultados se mantienen relativamente estables entre diferentes particiones de los datos?

Estas preguntas permiten evitar que una única cifra de exactitud sea interpretada como evidencia suficiente de calidad.

Modelos considerados

El bosque aleatorio se conserva como modelo principal porque constituye el algoritmo implementado en el manuscrito original. Este método combina múltiples árboles construidos mediante mecanismos de aleatorización y agrega sus predicciones para producir una salida conjunta. Breiman (2001) explica que el error de generalización de un bosque depende tanto de la capacidad de sus árboles individuales como de la correlación existente entre ellos.

Sin embargo, el bosque aleatorio no debe evaluarse de manera aislada. Para interpretar su desempeño se incorporan dos referencias:

- Un clasificador ingenuo o dummy.
- Una regresión logística.

El clasificador ingenuo produce predicciones mediante una regla simple, por ejemplo, asignando siempre la categoría más frecuente. No utiliza las relaciones entre las variables y la clase. Su función es establecer un rendimiento mínimo que cualquier modelo útil debería superar. La documentación de scikit-learn define DummyClassifier como un estimador que genera predicciones mediante reglas sencillas e ignora las características de entrada, por lo que resulta apropiado como línea de base.

La regresión logística se utiliza como modelo de referencia interpretable. Su inclusión permite comprobar si la estructura introducida durante la generación sintética puede recuperarse mediante una relación relativamente sencilla. Cuando el bosque aleatorio y la regresión logística presentan resultados semejantes, debe analizarse si la complejidad adicional del primero aporta una ventaja suficiente.

El árbol de decisión individual puede incluirse como recurso complementario para observar reglas, pero no es indispensable mantenerlo como tercer modelo comparativo dentro del cuerpo principal. Su utilización puede documentarse en el apéndice técnico.

Tabla 15

Modelos incluidos en la evaluación del prototipo

Modelo	*Función metodológica*	*Procesamiento requerido*	*Resultado esperado*
Clasificador ingenuo	Establecer una línea mínima de comparación	No requiere escalado	Rendimiento asociado con la distribución de las clases
Regresión logística	Proporcionar una referencia sencilla e interpretable	Imputación y escalado	Estimación lineal de la probabilidad
Bosque aleatorio	Actuar como modelo principal del prototipo	Imputación preventiva; no requiere escalado obligatorio	Representación de relaciones no lineales e interacciones

Nota. *La comparación no busca seleccionar un algoritmo para una implementación institucional, sino evaluar el comportamiento del prototipo dentro del conjunto sintético.*

Configuración del bosque aleatorio

La configuración inicial se mantiene próxima a la propuesta original:

- n_estimators = 200
- max_depth = 10
- min_samples_split = 5
- random_state = 42

El parámetro n_estimators establece la cantidad de árboles incluidos. La profundidad máxima limita el crecimiento de cada árbol y reduce la posibilidad de generar estructuras excesivamente específicas. min_samples_split define la cantidad mínima de observaciones requerida para crear una nueva división. La semilla facilita la reproducción de la secuencia aleatoria.

Estos valores deben presentarse como una configuración inicial de diseño, no como parámetros óptimos. El manuscrito no documenta un procedimiento mediante el cual hayan sido seleccionados. La documentación del RandomForestClassifier señala que este estimador ajusta múltiples árboles sobre submuestras de los datos y combina sus resultados para mejorar la capacidad predictiva y controlar el sobreajuste.

Tabla 16

Configuración inicial del bosque aleatorio

Parámetro	*Valor inicial*	*Función*
Número de árboles	200	Establecer la cantidad de clasificadores del conjunto
Profundidad máxima	10	Limitar la complejidad de los árboles
Muestras mínimas para dividir	5	Evitar divisiones sustentadas en pocos registros
Criterio de división	Gini	Medir la impureza de los nodos
Semilla aleatoria	42	Favorecer la reproducibilidad
Uso de identificador del proveedor	No	Evitar que una etiqueta ficticia funcione como predictor

Ponderación de clases	Sin ponderación en el modelo principal	Mantener la distribución sintética original

Nota. *Los parámetros constituyen una configuración inicial y deben registrarse junto con la versión del código y de los datos.*

La ponderación automática de clases puede analizarse como prueba de sensibilidad, pero no debe incorporarse sin comparar sus efectos. Una proporción aproximada de 70/30 exige métricas adecuadas, aunque no representa necesariamente un desbalance extremo que obligue a modificar el entrenamiento.

División de los datos

El conjunto se divide en dos subconjuntos:

- 80 % para entrenamiento.
- 20 % para prueba.

La partición conserva la proporción de la variable objetivo mediante estratificación. En scikit-learn, train_test_split permite utilizar el argumento stratify para dividir las observaciones de acuerdo con las etiquetas de clase. Esta opción reduce el riesgo de que una partición contenga una proporción muy diferente de casos positivos y negativos.

La operación se realizará con:

- Test_size = 0.20
- Random_state = 42
- Stratify = y

El conjunto de prueba debe permanecer separado durante la selección de parámetros y la comparación preliminar de modelos. Su función consiste en proporcionar una evaluación final sobre registros que no participaron en el ajuste.

Las operaciones que aprenden información de los datos, como imputación, escalado o selección de variables, deben ajustarse únicamente con el conjunto de entrenamiento. Aplicarlas sobre la base completa antes de la división permitiría que información del conjunto de prueba influyera indirectamente en el modelo.

La partición 80/20 permite conservar el criterio original del manuscrito, pero se fortalece mediante estratificación y encapsulación de las transformaciones en flujos de procesamiento.

Validación cruzada

Una sola división puede producir un resultado dependiente de los registros que fueron asignados a cada subconjunto. Para reducir esta dependencia, el conjunto de entrenamiento se evalúa mediante validación cruzada estratificada de cinco particiones.

En este procedimiento, los datos de entrenamiento se distribuyen en cinco grupos. En cada repetición, cuatro grupos se utilizan para ajustar el modelo y el restante para evaluarlo. El proceso continúa hasta que cada grupo haya funcionado como conjunto de validación.

StratifiedKFold conserva aproximadamente la proporción de las clases en cada partición, mientras que cross_validate permite calcular varias métricas dentro del mismo procedimiento.

La validación cruzada se utiliza para:

- Comparar el bosque aleatorio con las líneas de base;
- Estimar la variación de las métricas;
- Identificar señales de inestabilidad;
- Y revisar configuraciones alternativas.

La selección del modelo debe realizarse con los resultados promedio y su dispersión, no con la mejor partición individual.

La versión final deberá informar, al menos:

- Media de cada métrica;
- Desviación estándar;
- Cantidad de particiones;
- Semilla utilizada;
- Y procedimiento de estratificación.

La validación cruzada se aplica únicamente al conjunto de entrenamiento. Una vez seleccionada la configuración, el modelo se reajusta con todo ese subconjunto y se evalúa una sola vez sobre el conjunto de prueba reservado.

Ajuste de parámetros

El ajuste debe ser limitado y coherente con el carácter demostrativo del prototipo. No resulta necesario evaluar centenares de combinaciones, porque la variable objetivo fue construida mediante una regla conocida y el propósito no consiste en competir por el mayor rendimiento posible.

Los parámetros que pueden examinarse son:

- Número de árboles: 100, 200 y 300;
- Profundidad máxima: 5, 10 y sin límite;
- Muestras mínimas para dividir: 2, 5 y 10;
- Cantidad de variables consideradas en cada división;
- y ponderación de clases como análisis adicional.

El ajuste se realizará mediante las particiones internas de validación cruzada. El conjunto de prueba no se utilizará para seleccionar la combinación.

No debe presentarse el modelo con mayor puntuación como automáticamente superior. También deben considerarse:

- Diferencia frente a la regresión logística;
- Estabilidad entre particiones;
- Complejidad;
- Posibilidad de explicación;
- Y consistencia con la regla sintética.

Cuando varias configuraciones obtengan resultados semejantes, se priorizará la más sencilla y estable.

Matriz de confusión

La matriz de confusión compara las clases reales del conjunto sintético con las categorías predichas. Sus cuatro resultados son:

- Verdaderos positivos.
- Verdaderos negativos.
- Falsos positivos.
- Falsos negativos.

Dentro de este ejercicio, un verdadero positivo corresponde a un registro de compatibilidad sintética alta que fue reconocido como tal. Un falso negativo representa un caso de compatibilidad alta clasificado dentro de la categoría ordinaria. Estas denominaciones se refieren exclusivamente a la etiqueta artificial y no deben interpretarse como decisiones correctas o incorrectas sobre proveedores reales.

La matriz permite observar si el modelo concentra sus aciertos en la clase mayoritaria. Un sistema que predice siempre la categoría cero puede alcanzar una exactitud cercana al 70 %, pero no identificar ningún caso de la clase positiva.

Por esta razón, la matriz de confusión debe presentarse antes o junto con las métricas globales.

Métricas de evaluación

La evaluación utilizará un conjunto reducido de métricas complementarias:

- Exactitud.
- Exactitud balanceada.
- Precisión.
- Sensibilidad o recall.
- F1.
- ROC-AUC.
- Promedio de precisión o average precision.

La exactitud representa la proporción total de predicciones correctas. Se conserva porque facilita una lectura general, pero no se utilizará como único indicador.

La exactitud balanceada calcula el promedio del nivel de acierto obtenido en cada clase. Resulta útil cuando las categorías poseen frecuencias diferentes, ya que evita que la mayoritaria determine de manera desproporcionada el resultado. Scikit-learn la define como el promedio de la sensibilidad obtenida en cada clase.

La precisión responde a la proporción de predicciones positivas que fueron correctas. La sensibilidad muestra qué proporción de casos positivos fue identificada. F1 integra ambas medidas mediante una media armónica.

La curva ROC representa la relación entre la tasa de verdaderos positivos y la tasa de falsos positivos para diferentes umbrales. El área bajo la curva resume la capacidad de discriminación del clasificador. Fawcett (2006) señala que el análisis ROC permite visualizar y comparar el comportamiento de los clasificadores, aunque debe interpretarse con conocimiento de sus limitaciones.

Debido a que la clase positiva representa aproximadamente el 30 %, también se informará el promedio de precisión, que resume el comportamiento de la relación precisión-sensibilidad. La documentación oficial establece que average_precision_score sintetiza la curva de precisión y sensibilidad mediante una suma ponderada de las precisiones alcanzadas en distintos umbrales.

Saito y Rehmsmeier (2015) demostraron que las curvas de precisión-sensibilidad pueden proporcionar una interpretación más informativa que las curvas ROC cuando existe una fuerte diferencia entre la cantidad de observaciones negativas y positivas.

Tabla 17

Métricas seleccionadas para evaluar el modelo

Métrica	*Función*	*Criterio de interpretación*
Exactitud	Resume la proporción total de aciertos	No debe utilizarse de forma aislada
Exactitud balanceada	Promedia el desempeño en ambas clases	Indicador principal ante clases desiguales
Precisión	Evalúa la confiabilidad de las predicciones positivas	Disminuye cuando existen falsos positivos
Sensibilidad	Evalúa la detección de la clase positiva	Disminuye cuando existen falsos negativos
F1	Combina precisión y sensibilidad	Útil como medida de equilibrio
ROC-AUC	Examina discriminación en distintos umbrales	Debe acompañarse de métricas por clase
Promedio de precisión	Resume la relación precisión-sensibilidad	Relevante para la clase positiva

Nota. *Las métricas evalúan la recuperación de una regla sintética y no la capacidad de predecir adjudicaciones reales.*

Criterio de comparación entre modelos

El bosque aleatorio debe superar claramente al clasificador ingenuo. Si presenta resultados semejantes, no existiría evidencia de que las variables estén aportando información útil.

La comparación con la regresión logística permite valorar si las relaciones introducidas requieren realmente un modelo de conjunto. Pueden presentarse tres escenarios:

1. El bosque supera ampliamente a la regresión logística: la estructura podría contener relaciones no lineales o interacciones que favorecen al método de árboles.
2. Ambos modelos presentan resultados semejantes: la regresión logística podría ofrecer una solución más sencilla e interpretable.
3. La regresión logística supera al bosque: la regla sintética podría ser principalmente lineal o la configuración del bosque necesitar revisión.

No se seleccionará un modelo basándose únicamente en diferencias mínimas. La comparación incorporará rendimiento, estabilidad e interpretabilidad.

Tabla 18

Plantilla para reportar la validación cruzada

Modelo	*Exactitud balanceada*	*Precisión*	*Sensibilidad*	*F1*	*ROC-AUC*	*Promedio de precisión*
Clasificador ingenuo	Por calcular	Por calcular	Por calcular	Por calcular	Por calcular	Por calcular
Regresión logística	Por calcular	Por calcular	Por calcular	Por calcular	Por calcular	Por calcular
Bosque aleatorio	Por calcular	Por calcular	Por calcular	Por calcular	Por calcular	Por calcular

Nota. *Los resultados deberán expresarse mediante media y desviación estándar en las cinco particiones. La tabla no debe conservar la frase "por calcular" en la versión publicada.*

La tabla se incluye como guía de trabajo. No debe pasar a la maquetación definitiva hasta ejecutar el código corregido.

Evaluación final sobre el conjunto de prueba

Después de seleccionar la configuración mediante validación cruzada, el modelo se entrena con el 80 % reservado para aprendizaje. Posteriormente se generan las predicciones y puntuaciones sobre el 20 % de prueba.

La evaluación final deberá incluir:

- Distribución de las clases;
- Matriz de confusión;
- Exactitud;
- Exactitud balanceada;
- Precisión;
- Sensibilidad;
- F1;
- ROC-AUC;
- Promedio de precisión;
- Y curva de precisión-sensibilidad.

Los resultados deben presentarse con un lenguaje delimitado. Se recomienda utilizar “El modelo recuperó con determinado nivel de desempeño la estructura introducida en el conjunto sintético.”

No debe afirmarse “El modelo predice correctamente qué proveedor ganará una licitación.”

La primera formulación corresponde con el diseño metodológico. La segunda atribuye validez externa a una demostración artificial.

Umbral de clasificación

La salida probabilística puede transformarse en una clase mediante un umbral de 0,50. Este valor se utilizará inicialmente por simplicidad, pero no debe interpretarse como una frontera institucional.

Modificar el umbral cambia la relación entre falsos positivos y falsos negativos. Reducirlo puede incrementar la sensibilidad y aumentar simultáneamente los falsos positivos. Elevarlo puede incrementar la precisión de la clase positiva y omitir más casos.

El prototipo no dispone de evidencia para asignar costos institucionales a cada error. Por ello, el umbral principal se mantendrá en 0,50 y cualquier análisis alternativo se presentará únicamente como prueba de sensibilidad.

La interfaz debe mostrar la puntuación continua además de la categoría. Esta práctica evita ocultar diferencias entre registros que se encuentran cerca o lejos del punto de corte.

Calibración e interpretación de las puntuaciones

Las puntuaciones generadas por predict_proba no deben denominarse probabilidades reales de adjudicación. En este prototipo representan la proporción de árboles que favorece una clase dentro del modelo y del conjunto sintético.

Una puntuación de 0,75 indica que el modelo asigna mayor apoyo a la categoría positiva dentro de la estructura artificial. No significa que exista un 75 % de probabilidad de que una empresa obtenga un contrato real.

Para utilizar probabilidades en un entorno aplicado sería necesario evaluar su calibración, es decir, comprobar si grupos de casos que reciben una determinada probabilidad presentan frecuencias observadas semejantes. Este análisis no puede realizarse de manera sustantiva con la variable sintética para inferir comportamiento institucional.

Prevención del sobreajuste y fuga de información

El sobreajuste ocurre cuando el modelo reproduce particularidades del conjunto de entrenamiento y presenta un rendimiento inferior ante observaciones nuevas. La limitación de profundidad, el número mínimo de casos por división y la validación cruzada contribuyen a examinar este riesgo.

La fuga de información se produce cuando el modelo recibe datos que revelan directa o indirectamente la variable objetivo. Para evitarla:

- No se incorpora la puntuación utilizada para crear la clase;
- No se utilizan variables normalizadas auxiliares;
- No se incluye el identificador del proveedor;
- Las transformaciones se ajustan solo con entrenamiento;
- Y el conjunto de prueba permanece aislado durante la selección del modelo.

La prevención de la fuga resulta especialmente importante porque la clase fue creada mediante una regla conocida. Si cualquiera de sus resultados intermedios permanece dentro de los predictores, el desempeño sería artificialmente elevado.

Reproducibilidad del entrenamiento

El entrenamiento debe acompañarse de un registro mínimo:

- Versión del conjunto;
- Versión de python;
- Versión de scikit-learn;
- Variables utilizadas;
- Semilla;
- Proporción de prueba;
- Procedimiento de estratificación;
- Número de particiones;

- Parámetros;
- Métricas;
- Y fecha de ejecución.

La semilla fija reduce la variación asociada con la partición y la construcción del bosque, pero no sustituye el registro de las versiones del entorno.

El informe final debe conservar tanto las predicciones como las puntuaciones del conjunto de prueba, de manera que las tablas, curvas y explicaciones puedan reconstruirse.

Presentación de los resultados

Los resultados no deben colocarse como capturas de la consola de Google Colab. Se recomienda presentarlos mediante:

- 1 tabla comparativa de modelos;
- 1 matriz de confusión rediseñada;
- 1una curva de precisión-sensibilidad;
- y 1 síntesis interpretativa.

El informe de clasificación completo puede trasladarse al apéndice. En el cuerpo se incluirán únicamente los indicadores necesarios para sostener el análisis. No deben emplearse expresiones como:

- modelo robusto;
- alta precisión;
- resultado confiable;
- recomendación correcta;
- desempeño satisfactorio;

El diseño de evaluación permite conservar el bosque aleatorio, la partición 80/20 y las métricas inicialmente planteadas, pero incorpora condiciones necesarias para interpretar sus resultados con mayor rigor. La estratificación conserva la distribución de las clases, la validación cruzada permite examinar la estabilidad, y las líneas de base muestran si el modelo utiliza realmente la información de los predictores.

La exactitud deja de ser el único indicador y se complementa con exactitud balanceada, precisión, sensibilidad, F1, ROC-AUC y promedio de precisión. Los resultados se entienden como evidencia sobre la recuperación de una regla sintética, no como demostración de que el sistema pueda anticipar adjudicaciones reales.

El apartado siguiente reconstruye la arquitectura funcional del prototipo, la función de comparación, las visualizaciones y el mecanismo explicativo. En esa etapa se sustituirá la comparación simple de promedios por una explicación vinculada con el comportamiento efectivo del modelo.

Arquitectura funcional, explicabilidad e interfaz

La arquitectura funcional del prototipo organiza los componentes necesarios para transformar un conjunto sintético de características en una puntuación experimental acompañada de información explicativa. Su diseño comprende la entrada de datos, la validación de las variables, el procesamiento, la ejecución del modelo, la generación de explicaciones, la presentación de resultados y la revisión humana.

El prototipo no corresponde a una aplicación web implementada en un entorno productivo. Su funcionamiento se desarrolla dentro de un cuaderno computacional en Google Colab, donde el código, las tablas, los gráficos y las descripciones pueden ejecutarse de forma secuencial. Este entorno facilita la demostración y la revisión académica, pero no incorpora por sí mismo

mecanismos de autenticación, control de acceso, almacenamiento institucional, auditoría continua o integración con plataformas oficiales.

La arquitectura revisada conserva los componentes tecnológicos presentes en el manuscrito, pero modifica la relación entre ellos. La salida principal deja de denominarse "probabilidad de ganar una licitación" y pasa a presentarse como puntuación estimada de compatibilidad sintética. Del mismo modo, la expresión "se recomienda elegir" se sustituye por una formulación que comunica el resultado sin convertirlo en una decisión.

La arquitectura se organiza en seis capas:

1. Capa de datos sintéticos.
2. Capa de validación y procesamiento.
3. Capa predictiva.
4. Capa de explicación.
5. Capa de presentación e interacción.
6. Capa de revisión humana y registro.

Tabla 19

Componentes de la arquitectura funcional del prototipo

Capa	*Función principal*	*Entrada*	*Salida*
Datos sintéticos	Almacenar los registros y sus variables	Conjunto DSPT-v1.0	Registros identificados y documentados
Validación y procesamiento	Comprobar rangos, tipos, ausencias y transformaciones	Registros sintéticos	Matriz preparada para el modelo
Modelo predictivo	Estimar la pertenencia a la clase sintética	Variables predictoras	Puntuación y clase estimada
Explicación	Identificar factores relacionados con la salida	Modelo, datos y predicción	Importancias globales y contribuciones locales
Presentación	Mostrar puntuaciones, perfiles y explicaciones	Resultados del modelo	Tablas, gráficos y advertencias
Revisión humana	Examinar, aceptar o cuestionar la información	Salidas y documentación	Registro de revisión y observaciones

Nota. *La arquitectura corresponde a una prueba de concepto. No representa una plataforma habilitada para intervenir en procedimientos reales de contratación.*

Flujo funcional del prototipo

El flujo comienza con la selección de uno o varios registros sintéticos. Antes de realizar la predicción, el sistema verifica que los valores pertenezcan a los tipos y rangos definidos en la Tabla 13. Si encuentra datos ausentes, valores imposibles o campos no reconocidos, debe advertir al usuario antes de continuar.

Una vez validado el registro, las variables se organizan en el mismo orden utilizado durante el entrenamiento. Las transformaciones se aplican mediante el flujo de procesamiento previamente ajustado. Esta condición evita diferencias entre el tratamiento aplicado durante el entrenamiento y el utilizado para nuevos casos.

El modelo genera dos salidas relacionadas:

- La puntuación estimada para la clase de compatibilidad sintética alta.
- La categoría resultante según el umbral definido.

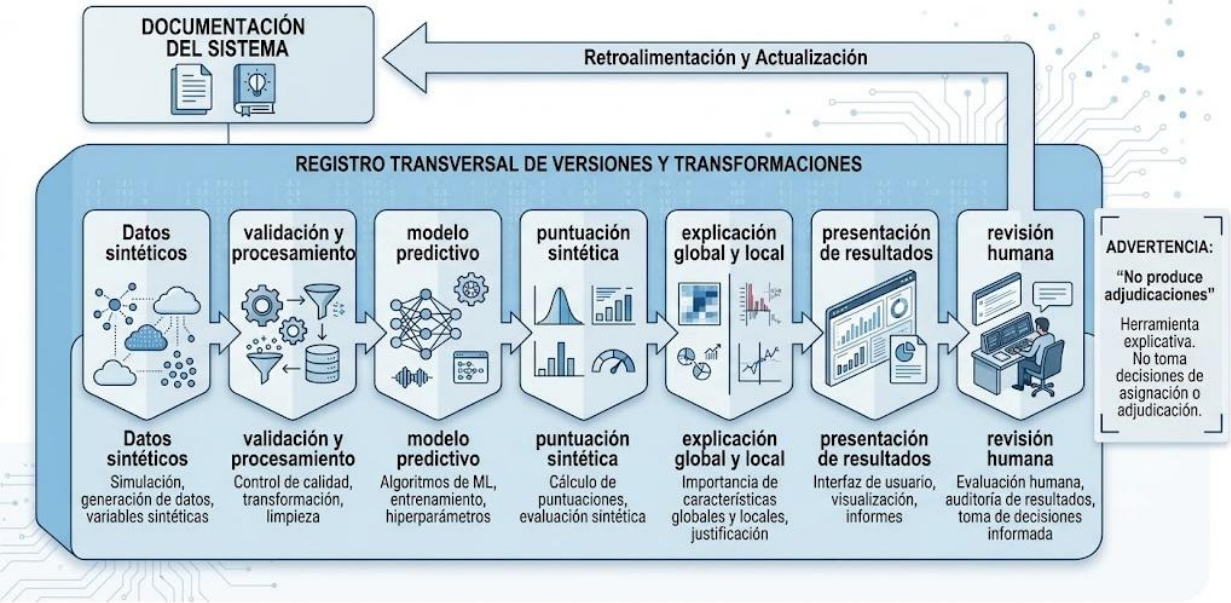
La puntuación debe mostrarse como un valor comprendido entre cero y uno, pero acompañada de una advertencia que impida interpretarla como probabilidad real de adjudicación.

La salida recomendada es "*El modelo asignó al registro una puntuación de compatibilidad sintética de [valor]. Esta estimación corresponde a un escenario artificial y no representa una probabilidad real de contratación, una evaluación empresarial ni una decisión administrativa.*"

Después de calcular la puntuación, la capa explicativa identifica las características que influyeron en el resultado. La explicación se presenta antes de cualquier comparación entre alternativas, de manera que el usuario pueda examinar cada caso individualmente.

Figura 9

Arquitectura funcional del prototipo explicable



Nota. *Elaboración propia. La decisión permanece bajo responsabilidad humana y se encuentra fuera del funcionamiento automatizado del prototipo.*

Módulo de entrada y validación

El módulo de entrada permite seleccionar registros existentes dentro del conjunto sintético o introducir manualmente un perfil hipotético. En ambos casos, la interfaz debe identificar claramente que la información corresponde a una simulación.

Cada variable debe estar acompañada por:

- Nombre comprensible.
- Definición.
- Unidad.
- Rango permitido.
- Dirección esperada dentro de la regla.

- Advertencia sobre su carácter demostrativo.

La validación debe impedir:

- Valores negativos.
- Valores fuera de los intervalos.
- Categorías distintas de cero y uno en variables binarias.
- Campos numéricos ingresados como texto.
- Registros sin las variables necesarias.
- Identificadores empresariales reales.

La interfaz no debe completar automáticamente valores inválidos sin informar al usuario. Cuando sea necesario aplicar una imputación, deberá mostrarse una advertencia indicando qué campo fue completado y mediante qué estadístico.

Revisión del módulo de comparación

La función original comparar_proveedores(proveedor1, proveedor2) agrupa las observaciones según el nombre empresarial y calcula el promedio de todas las variables. Este procedimiento presenta dos problemas.

En primer lugar, cada nombre aparece asociado con numerosos registros generados aleatoriamente. Por tanto, los promedios no describen empresas, ofertas ni perfiles reales. En segundo lugar, el uso de nombres empresariales identificables produce el riesgo de atribuir características artificiales a organizaciones existentes.

La función debe sustituirse por:

"comparar_perfiles_sinteticos(registro_a, registro_b)"

La nueva función comparará:

- Dos registros individuales seleccionados del conjunto.
- dos perfiles hipotéticos introducidos manualmente.

Cada perfil recibirá un identificador neutro:

- Perfil A.
- Perfil B.

El procedimiento seguirá esta secuencia:

1. Validar los dos perfiles.
2. Aplicar el flujo de procesamiento.
3. Calcular sus puntuaciones.
4. Obtener las explicaciones locales.
5. Presentar las diferencias relevantes.
6. Comunicar las limitaciones.
7. Registrar la revisión humana.

La comparación no debe indicar que un perfil es “mejor” de manera general. La salida deberá utilizar la formulación:

- Dentro de la regla y del modelo sintético utilizados, el Perfil A obtuvo una puntuación superior al Perfil B. La diferencia se relacionó principalmente con las características que se muestran a continuación.

Cuando las puntuaciones sean próximas, el sistema deberá advertir:

- La diferencia entre los perfiles es reducida y no permite establecer una separación sustantiva dentro del escenario simulado.

Explicación global del modelo

La explicación global busca mostrar qué variables utiliza el modelo para discriminar entre las dos clases sintéticas. Para este propósito se recomienda emplear la importancia por permutación sobre el conjunto de prueba.

La importancia por permutación modifica aleatoriamente una variable y mide cuánto disminuye el desempeño del modelo. Cuando la alteración produce una reducción elevada, se interpreta que el clasificador depende de esa característica para mantener su rendimiento. El procedimiento puede aplicarse a diferentes estimadores y repetirse para estimar la variación de la importancia.

La importancia debe calcularse únicamente después de comprobar que el modelo posee un rendimiento predictivo superior a las líneas de base. Una variable puede aparecer como poco importante debido a que el modelo completo funciona deficientemente. Además, la importancia describe la dependencia de un modelo específico y no el valor intrínseco, causal o institucional de la característica.

La configuración recomendada es:

- Datos: conjunto de prueba.
- Repeticiones por variable: 30.
- Semilla: 42.
- Métrica: exactitud balanceada.
- Resultado: promedio y desviación estándar de la disminución.

La visualización deberá ordenar las variables desde la mayor hasta la menor importancia y mostrar la variabilidad obtenida entre repeticiones.

La interpretación debe expresarse de esta forma:

La alteración de la variable cumplimiento de SLA produjo una de las mayores disminuciones del desempeño, lo que indica que el modelo depende de esta característica para recuperar la estructura de la clase sintética.

No deberá afirmarse:

El cumplimiento de SLA es la causa principal de una adjudicación.

La primera afirmación describe al modelo. La segunda atribuye causalidad y validez institucional que no fueron demostradas.

Explicación local mediante SHAP

La explicación local debe identificar cómo las variables contribuyeron a la puntuación de un registro específico. Para ello se propone incorporar SHAP como método principal.

SHAP asigna un valor de contribución a cada característica de una predicción individual. El resultado permite observar qué variables desplazaron la salida desde un valor de referencia hacia la puntuación obtenida. Lundberg y Lee (2017) presentaron SHAP como un marco unificado de explicaciones aditivas y establecieron que cada característica recibe un valor de importancia para una predicción concreta.

Para cada perfil deberán mostrarse:

- Valor de referencia del modelo.
- Puntuación obtenida.
- Cinco características con mayor contribución positiva.
- Cinco características con mayor contribución negativa.
- Magnitud y dirección de las contribuciones.
- Advertencia sobre la naturaleza sintética del resultado.

La explicación no se limitará a indicar cuál perfil posee el valor bruto más elevado. Una variable puede presentar una magnitud mayor y, sin embargo, reducir la puntuación. También puede ocurrir que una característica tenga un valor favorable, pero poca influencia en una predicción particular.

La salida textual puede adoptar esta forma:

- La puntuación del Perfil A se situó por encima del valor de referencia. Las principales contribuciones positivas correspondieron al cumplimiento de SLA, la satisfacción sintética y la experiencia. El costo por hora y el tiempo de respuesta ejercieron contribuciones negativas. Estas relaciones describen el funcionamiento del modelo sintético y no constituyen criterios oficiales de evaluación.

Los resultados concretos de SHAP no deben escribirse en el manuscrito hasta ejecutar el código corregido. La metodología puede describirse ahora, pero las variables identificadas y sus valores deberán proceder de la ejecución final.

Diferencia entre descripción y explicación

La versión original presenta como explicación una comparación de promedios. Esta operación constituye una descripción de las diferencias entre registros, pero no demuestra por qué el modelo produjo una determinada puntuación. Esta diferenciación es fundamental para que la interfaz no presente una comparación descriptiva como si fuera una explicación fiel del algoritmo.

Rediseño de las visualizaciones

La versión original incorpora un gráfico de barras, un gráfico radar y una comparación de probabilidades. El gráfico radar utiliza variables expresadas en escalas incompatibles, por lo que debe eliminarse. La interfaz revisada utilizará tres recursos:

- Gráfico de perfil normalizado.

- Gráfico de puntuaciones sintéticas.
- Gráfico de contribuciones locales.

Gráfico de perfil normalizado

El gráfico compara los valores transformados entre cero y uno. Antes de mostrar el costo y el tiempo de respuesta, su dirección debe invertirse para que los valores altos representen siempre una condición favorable dentro de la regla sintética.

Interfaz explicable

La interfaz se entiende como el conjunto de elementos mediante los cuales el usuario introduce información, obtiene resultados y examina las explicaciones. En la prueba de concepto, esta interfaz se implementa mediante celdas, controles y salidas dentro del cuaderno computacional.

Una interfaz explicable debería mostrar los elementos en el siguiente orden:

- Declaración de alcance.
- Selección o ingreso del perfil.
- Validación de datos.
- Puntuación sintética.
- Categoría estimada.
- Explicación local.
- Comparación descriptiva.
- Importancia global.
- Limitaciones.
- Registro de revisión humana.

La interfaz debe evitar elementos que induzcan a aceptar automáticamente la alternativa con mayor puntuación. No se utilizarán botones denominados “Elegir ganador”, “Seleccionar proveedor” o “Adjudicar”. Pueden emplearse:

- Comparar perfiles.
- Generar puntuación.
- Mostrar explicación.
- Registrar observaciones.
- Revisar limitaciones.

Explicación adaptada al usuario

La información puede organizarse en dos niveles. El nivel general está dirigido a lectores y usuarios no especializados. Debe mostrar:

- Puntuación.
- Principales factores.
- Dirección de las contribuciones.
- Advertencias.
- Límites del resultado.

El nivel técnico debe incluir:

- Versión del conjunto.
- Versión del modelo.
- Fecha de ejecución.
- Variables procesadas.
- Umbral.
- Valores SHAP.
- Importancia por permutación.
- Métricas.
- Registro de transformaciones.

Esta organización evita sobrecargar al usuario general y conserva la información necesaria para la auditoría técnica.

Registro de la revisión humana

Después de mostrar los resultados, la interfaz debe permitir registrar una observación. Las opciones pueden ser:

- Resultado comprendido.
- Resultado requiere revisión.
- Datos de entrada insuficientes.
- Explicación inconsistente.
- Caso fuera del alcance del prototipo.
- Otra observación.

El registro no significa que el usuario apruebe una adjudicación. Documenta únicamente su valoración sobre la claridad y coherencia de la salida.

La información mínima del registro comprende:

- identificador de ejecución;
- fecha y hora;
- versión del conjunto;
- versión del modelo;
- perfiles comparados;
- puntuaciones;
- explicación mostrada;
- observación del revisor;
- y estado de la revisión.

El NIST AI RMF destaca la necesidad de integrar la gestión de riesgos, la documentación y la asignación de responsabilidades durante el ciclo de vida del sistema. La trazabilidad de las salidas y de la intervención humana forma parte de esta gestión y no debe considerarse una actividad posterior o separada del diseño.

Verificación funcional de la arquitectura

La verificación debe comprobar que cada componente produce el comportamiento esperado.

Tabla 20

Verificación funcional

Prueba	*Procedimiento*	*Resultado esperado*
Generación reproducible	Ejecutar dos veces con la misma semilla	Obtener el mismo conjunto
Validación de rangos	Ingresar un valor fuera del límite	Mostrar advertencia y detener el cálculo
Exclusión del identificador	Revisar las variables del modelo	Confirmar que el nombre no se utiliza como predictor
Producción de puntuación	Procesar un perfil válido	Obtener un valor entre cero y uno
Comparación	Procesar dos perfiles	Mostrar puntuaciones y diferencias descriptivas
Explicación local	Calcular SHAP para un registro	Mostrar contribuciones positivas y negativas
Explicación global	Permutar las variables	Obtener importancia promedio y variación
Advertencia de alcance	Generar cualquier resultado	Mantener visible la naturaleza sintética
Registro humano	Ingresar una observación	Conservar la revisión junto con la ejecución

Nota. *Las pruebas verifican el funcionamiento del artefacto y no su validez para procedimientos reales.*

Condiciones para declarar el prototipo funcional

La expresión “prototipo funcional” solo deberá utilizarse cuando se hayan comprobado los siguientes aspectos:

- el código se ejecuta sin errores;
- la variable objetivo corregida sustituye a la etiqueta aleatoria;
- los nombres reales fueron eliminados;

- las transformaciones se encuentran encapsuladas;
- los modelos fueron evaluados;
- la salida se denomina puntuación sintética;
- la comparación se realiza entre perfiles;
- la importancia por permutación fue calculada;
- SHAP fue implementado para explicaciones locales;
- los gráficos utilizan escalas adecuadas;
- las advertencias son visibles;
- y los resultados pueden reproducirse.

Hasta completar estas condiciones, es más preciso utilizar *"Versión demostrativa en proceso de consolidación metodológica."*

La arquitectura revisada conserva la lógica central del artefacto original: datos, modelo, comparación, visualización e interacción. Sin embargo, sustituye la recomendación automática por una presentación analítica, separa las diferencias descriptivas de las explicaciones algorítmicas y mantiene la decisión fuera del sistema.

La importancia por permutación permite examinar globalmente las variables de las que depende el desempeño, mientras que SHAP permite analizar la contribución de cada característica en una predicción específica. La interfaz comunica estos resultados mediante puntuaciones, gráficos y advertencias, y registra la revisión humana como parte de la trazabilidad.

Prototipo funcional

Una versión operativa del sistema de recomendación, que permite a los usuarios explorar y entender las recomendaciones generadas a través de una interfaz clara y accesible.

```
# Importación de bibliotecas necesarias
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
```

```
from sklearn.preprocessing import StandardScaler
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score, classification_report, confusion_matrix
from sklearn.impute import SimpleImputer
import matplotlib.pyplot as plt
import seaborn as sns

# Configuración para Google Colab
%matplotlib inline
plt.style.use('seaborn')

# Función para generar datos simulados de licitaciones
def generar_datos_licitaciones(n_samples=2000):
    """
    Esta función genera datos simulados para licitaciones de proveedores de soporte IT en
Ecuador.
    Incluye características como experiencia, proyectos completados, tiempo de respuesta,
etc.
    """
    np.random.seed(42)  # Asegura reproducibilidad de los resultados
    proveedores = ['IT Software One', 'Akross', 'Sonda', 'Tata', 'IBM', 'Novatech',
'Kruger']
    data = {
        'proveedor': np.random.choice(proveedores, n_samples),
        'experiencia_anios': np.random.randint(1, 30, n_samples),
        'proyectos_completados': np.random.randint(5, 200, n_samples),
        'tiempo_respuesta_horas': np.random.uniform(1, 72, n_samples),
        'costo_por_hora': np.random.uniform(40, 250, n_samples),
        'satisfaccion_cliente': np.random.uniform(2, 5, n_samples),
        'certificaciones': np.random.randint(1, 15, n_samples),
        'personal_calificado': np.random.randint(10, 500, n_samples),
        'cobertura_nacional': np.random.choice([0, 1], n_samples),
        'soporte_24_7': np.random.choice([0, 1], n_samples),
        'especializacion_sector_publico': np.random.uniform(0, 1, n_samples),
        'cumplimiento_sla': np.random.uniform(0.7, 1, n_samples),
        'ganador_licitacion': np.random.choice([0, 1], n_samples, p=[0.7, 0.3])
    }
    return pd.DataFrame(data)

# Generamos los datos simulados
df = generar_datos_licitaciones()

# Función para preprocesar los datos
def preprocesar_datos(df):
    """
    Esta función preprocesa los datos para el modelo de machine learning:
    1. Maneja los valores faltantes reemplazándolos con la media.
    2. Codifica las variables categóricas (en este caso, 'proveedor').
    3. Separa las características (X) de la variable objetivo (y).
    """
    imputer = SimpleImputer(strategy='mean')
    df_numeric = df.select_dtypes(include=[np.number])
    df[df_numeric.columns] = imputer.fit_transform(df_numeric)

    df_encoded = pd.get_dummies(df, columns=['proveedor'], drop_first=True)

    X = df_encoded.drop('ganador_licitacion', axis=1)
    y = df_encoded['ganador_licitacion']

    return X, y
```

```
# Preprocesamos los datos
X, y = preprocesar_datos(df)

# Dividimos los datos en conjuntos de entrenamiento y prueba
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Escalamos las características para mejorar el rendimiento del modelo
scaler = StandardScaler()
X_train_scaled = scaler.fit_transform(X_train)
X_test_scaled = scaler.transform(X_test)

# Creamos y entrenamos el modelo de Random Forest
model = RandomForestClassifier(n_estimators=200, max_depth=10, min_samples_split=5,
random_state=42)
model.fit(X_train_scaled, y_train)

# Evaluamos el modelo
y_pred = model.predict(X_test_scaled)
accuracy = accuracy_score(y_test, y_pred)
print(f"Precisión del modelo: {accuracy:.2f}")
print("\nInforme de clasificación:")
print(classification_report(y_test, y_pred))

# Función para comparar proveedores
def comparar_proveedores(proveedor1, proveedor2):
    """
    Esta función compara dos proveedores y recomienda cuál elegir basándose en el modelo
entrenado.
    También proporciona visualizaciones para ayudar en la toma de decisiones.
    """
    if proveedor1 not in df['proveedor'].unique() or proveedor2 not in
df['proveedor'].unique():
        print(f"Error: Uno o ambos proveedores no existen en los datos.")
        return

    # Obtenemos los datos promedio de cada proveedor
    columnas_numericas =
df.select_dtypes(include=[np.number]).columns.drop('ganador_licitacion')
    datos_proveedor1 = df[df['proveedor'] == proveedor1][columnas_numericas].mean()
    datos_proveedor2 = df[df['proveedor'] == proveedor2][columnas_numericas].mean()

    # Preparamos los datos para el modelo
    datos_modelo1 = pd.DataFrame(index=[0], columns=X.columns)
    datos_modelo2 = pd.DataFrame(index=[0], columns=X.columns)

    for col in X.columns:
        if col.startswith('proveedor_'):
            datos_modelo1[col] = 1 if col == f'proveedor_{proveedor1}' else 0
            datos_modelo2[col] = 1 if col == f'proveedor_{proveedor2}' else 0
        elif col in columnas_numericas:
            datos_modelo1[col] = datos_proveedor1[col]
            datos_modelo2[col] = datos_proveedor2[col]

    # Escalamos los datos y obtenemos las probabilidades
    datos_escalados1 = scaler.transform(datos_modelo1)
    datos_escalados2 = scaler.transform(datos_modelo2)

    prob1 = model.predict_proba(datos_escalados1)[0][1]
    prob2 = model.predict_proba(datos_escalados2)[0][1]
```

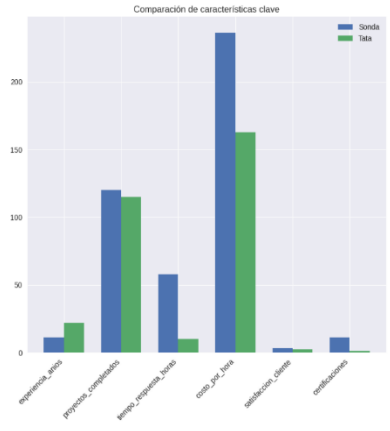
```
    print(f"Probabilidad de ganar licitación para {proveedor1}: {prob1:.2f}")
    print(f"Probabilidad de ganar licitación para {proveedor2}: {prob2:.2f}")

    recomendado = proveedor1 if prob1 > prob2 else proveedor2
    print(f"\nSe recomienda elegir a: {recomendado}")

    # Visualización mejorada
    fig, (ax1, ax2, ax3) = plt.subplots(1, 3, figsize=(24, 8))

    # Gráfico de barras para comparar características clave
    caracteristicas = ['experiencia_anios', 'proyectos_completados',
'tiempo_respuesta_horas',
                       'costo_por_hora', 'satisfaccion_cliente', 'certificaciones']
    x = np.arange(len(caracteristicas))
    width = 0.35

    valores1 = datos_proveedor1[caracteristicas]
    valores2 = datos_proveedor2[caracteristicas]

    ax1.bar(x - width/2, valores1, width, label=proveedor1)
    ax1.bar(x + width/2, valores2, width, label=proveedor2)

    ax1.set_xticks(x)
    ax1.set_xticklabels(caracteristicas, rotation=45, ha='right')
    ax1.legend()
    ax1.set_title("Comparación de características clave")

    # Gráfico de radar para visualizar fortalezas relativas
    caracteristicas_radar = caracteristicas + ['personal_calificado',
'cobertura_nacional',
                                               'soporte_24_7',
'especializacion_sector_publico', 'cumplimiento_sla']
    valores1 = datos_proveedor1[caracteristicas_radar]
    valores2 = datos_proveedor2[caracteristicas_radar]

    angles = np.linspace(0, 2*np.pi, len(caracteristicas_radar), endpoint=False)
    valores1 = np.concatenate((valores1, [valores1[0]]))
    valores2 = np.concatenate((valores2, [valores2[0]]))
    angles = np.concatenate((angles, [angles[0]]))

    ax2 = plt.subplot(132, polar=True)
    ax2.plot(angles, valores1, 'o-', linewidth=2, label=proveedor1)
    ax2.plot(angles, valores2, 'o-', linewidth=2, label=proveedor2)
    ax2.set_xticks(angles[:-1])
    ax2.set_xticklabels(caracteristicas_radar)
    ax2.legend(loc='upper right', bbox_to_anchor=(0.1, 0.1))
    ax2.set_title("Fortalezas relativas de los proveedores")

    # Gráfico de probabilidades de ganar la licitación
    probs = [prob1, prob2]
    ax3.bar([proveedor1, proveedor2], probs)
    ax3.set_ylim(0, 1)
    ax3.set_title("Probabilidad de ganar la licitación")
    ax3.set_ylabel("Probabilidad")

    for i, v in enumerate(probs):
        ax3.text(i, v, f'{v:.2f}', ha='center', va='bottom')

    plt.tight_layout()
```

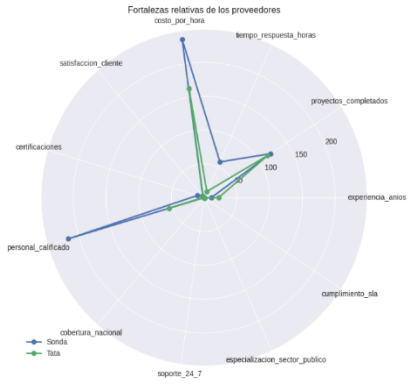
```
    plt.show()

    # Explicación detallada de la recomendación
    print("\nExplicación de la recomendación:")
    print(f"Se recomienda elegir a {recomendado} por las siguientes razones:")

    if recomendado == proveedor1:
        mejor = datos_proveedor1
        peor = datos_proveedor2
    else:
        mejor = datos_proveedor2
        peor = datos_proveedor1

    for caract in caracteristicas_radar:
        if mejor[caract] > peor[caract]:
            print(f"- {caract}: {recomendado} tiene un valor de {mejor[caract]:.2f}
comparado con {peor[caract]:.2f} del otro proveedor.")

# Ejemplo de uso: Comparación de proveedores
```

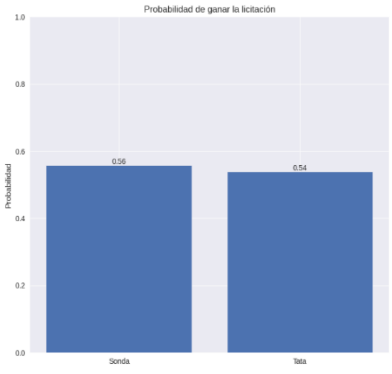
Informe de resultados

Un análisis exhaustivo de los datos procesados, las decisiones tomadas por el modelo y las razones detrás de estas, con el objetivo de garantizar la transparencia y comprensibilidad del sistema

```
--- RESUMEN EJECUTIVO ---
Basado en nuestro análisis, recomendamos seleccionar a Sonda para la licitación de servicios de IT.
Probabilidad de ganar la licitación:
- Sonda: 0.56
- Tata: 0.54

Factores clave que respaldan esta recomendación:
- experiencia_previa_Proyectos exitosos en ministerios de economía y educación <= -0.48
- proveedor_IBM <= -0.33
- costo_por_hora <= -0.83
```







Criterios éticos, reproducibilidad y limitaciones

La construcción de un prototipo explicable relacionado con contratación pública exige reconocer que la corrección técnica del código no constituye evidencia suficiente de legitimidad, utilidad institucional o confiabilidad. El tratamiento ético debe acompañar todas las etapas del proceso, desde la definición del problema y la generación de los datos hasta la presentación de las puntuaciones y la interpretación de las explicaciones.

El primer criterio ético corresponde a la transparencia sobre la naturaleza de la información. El conjunto utilizado en la propuesta es sintético, no procede de expedientes contractuales ni representa el comportamiento de empresas reales. Esta condición debe comunicarse de manera visible en el código, las tablas, las figuras, la interfaz y las conclusiones. Ocultar o minimizar el carácter artificial de los registros podría inducir al lector a atribuir validez empírica a resultados que únicamente demuestran el funcionamiento de una arquitectura computacional.

El segundo criterio se relaciona con la protección de la reputación de terceros. La asociación de nombres empresariales identificables con valores creados aleatoriamente puede producir interpretaciones incorrectas sobre su experiencia, costo, capacidad o posibilidad de adjudicación. Por esta razón, la versión final utilizará únicamente identificadores ficticios y eliminará cualquier resultado que permita relacionar una puntuación simulada con una organización existente.

El tercer criterio corresponde a la finalidad del sistema. El prototipo no debe emplearse para evaluar, excluir, jerarquizar o seleccionar proveedores reales. Tampoco debe integrarse directamente a un procedimiento administrativo sin un proceso previo de reconstrucción metodológica, validación jurídica, evaluación técnica y revisión institucional. La salida del modelo constituye una puntuación experimental y no una decisión, dictamen o recomendación vinculante.

La supervisión humana debe mantenerse durante todo el proceso. Su finalidad no consiste en confirmar automáticamente la alternativa con mayor puntuación, sino en revisar la calidad de los

datos, examinar las explicaciones, detectar posibles inconsistencias y rechazar el resultado cuando este no sea pertinente. La responsabilidad de una decisión pública no puede trasladarse al algoritmo, al desarrollador o a la interfaz.

El prototipo también debe evitar la automatización excesiva. Las expresiones, colores, gráficos y botones de la interfaz pueden influir en la percepción del usuario. Presentar una alternativa como "ganadora" o "recomendada" podría generar una confianza superior a la evidencia disponible. En su lugar, la interfaz debe mostrar las puntuaciones, sus factores asociados, las advertencias y las limitaciones, dejando fuera del sistema cualquier acción de adjudicación.

La reproducibilidad se garantiza mediante la documentación de las versiones del conjunto de datos, el código, las bibliotecas, la semilla aleatoria, las variables, las transformaciones y los parámetros del modelo. El repositorio asociado con la obra deberá contener el cuaderno ejecutable, el conjunto sintético congelado, el archivo de requisitos y las instrucciones necesarias para repetir el procedimiento.

La semilla aleatoria facilita la repetición de la simulación, pero no demuestra que sus resultados sean válidos fuera del entorno artificial. Reproducir una ejecución significa obtener nuevamente el mismo resultado bajo condiciones equivalentes. No significa que el modelo represente adecuadamente la contratación pública ni que sus puntuaciones puedan generalizarse hacia casos reales.

Las principales limitaciones metodológicas se derivan de seis condiciones:

1. Los datos son sintéticos y no representativos.
2. La variable objetivo se construye mediante una regla definida por los autores.
3. Los rangos y pesos no proceden de criterios oficiales.
4. El prototipo simplifica la contratación mediante variables tabulares.
5. No se realizaron pruebas documentadas con usuarios institucionales.

6. No se efectuó una validación externa con procedimientos reales.

Estas limitaciones no invalidan la prueba de concepto, pero delimitan el tipo de conclusiones que pueden formularse. El prototipo permite demostrar un flujo de generación, procesamiento, clasificación y explicación. No permite concluir que se reducen costos, se mejora la adjudicación, se eliminan sesgos o se incrementa la eficiencia institucional.

La futura incorporación de datos reales requeriría definir con precisión la unidad de análisis, el período, las fuentes, los criterios de inclusión, la calidad de los registros, las bases jurídicas para su tratamiento y las medidas de protección aplicables. También sería necesario evitar que adjudicaciones pasadas se utilicen automáticamente como representación de decisiones ideales, puesto que los modelos pueden reproducir regularidades históricas sin determinar si estas son técnicamente pertinentes o socialmente aceptables.

El marco de gestión de riesgos de inteligencia artificial del NIST establece que la gobernanza, la contextualización, la medición y la gestión deben integrarse durante todo el ciclo de vida del sistema. En correspondencia con esta orientación, el prototipo debe documentarse como una herramienta experimental, sometida a revisión y limitada a los usos expresamente declarados (NIST, 2023).

En consecuencia, el capítulo metodológico no presenta el artefacto como un producto terminado. Su aporte consiste en reconstruir de manera transparente las decisiones técnicas, identificar los riesgos del diseño original y establecer las condiciones necesarias para una futura evaluación. Sobre esta base, el capítulo siguiente discute los aportes, las limitaciones y la proyección del prototipo, y formula conclusiones ajustadas a la evidencia disponible.

CAPITULO IV: DISCUSIÓN, CONCLUSIONES Y PROYECCIÓN DEL PROTOTIPO

CHAPTER IV: DISCUSSION, CONCLUSIONS, AND PROTOTYPE OUTLOOK

CAPITULO IV: DISCUSIÓN, CONCLUSIONES Y PROYECCIÓN DEL PROTOTIPO

El presente capítulo integra los principales argumentos desarrollados en la obra y examina críticamente el alcance del prototipo de inteligencia artificial explicable. La discusión no se orienta a demostrar una aplicación institucional concluida, sino a valorar la coherencia de la arquitectura propuesta, las decisiones metodológicas adoptadas, las posibilidades de desarrollo y las restricciones que deben reconocerse antes de cualquier uso relacionado con contratación pública.

La reconstrucción del manuscrito permitió establecer una diferencia fundamental entre el propósito tecnológico de la propuesta y la evidencia disponible. El prototipo articula componentes pertinentes para una prueba de concepto: datos tabulares, preparación de variables, un modelo de clasificación, puntuaciones sintéticas, comparaciones visuales y mecanismos de explicación. Sin embargo, estos componentes fueron desarrollados sobre información generada mediante programación y no sobre procedimientos, ofertas o adjudicaciones reales.

Esta condición modifica la interpretación de los resultados. Las puntuaciones generadas por el modelo no constituyen probabilidades reales de que un proveedor obtenga una adjudicación. Representan la respuesta de un clasificador ante una estructura artificial incorporada durante la generación del conjunto. Del mismo modo, las explicaciones describen la manera en que el modelo utiliza las variables sintéticas, pero no demuestran relaciones causales ni criterios jurídicamente aplicables a la contratación pública.

El valor de la obra se encuentra, por tanto, en la documentación y análisis de una arquitectura explicable, en la identificación de los riesgos asociados con su diseño y en la formulación de condiciones metodológicas para una evolución futura. Esta delimitación permite conservar el núcleo innovador del proyecto sin atribuirle resultados que no fueron comprobados.

Discusión de la propuesta metodológica

La propuesta parte de una preocupación pertinente: los modelos predictivos pueden procesar numerosas variables y generar puntuaciones difíciles de comprender para quienes deben utilizar sus resultados. En contextos institucionales, esta opacidad adquiere relevancia porque las decisiones requieren justificación, trazabilidad y posibilidad de revisión. La incorporación de mecanismos explicativos responde a la necesidad de evitar que una salida algorítmica sea presentada como un resultado incuestionable.

El prototipo demuestra que es técnicamente posible organizar un flujo compuesto por generación de datos, preparación, clasificación, comparación y representación de resultados. El uso de Python, scikit-learn y un entorno de cuaderno computacional permite integrar las etapas dentro de un recurso reproducible. Esta configuración resulta apropiada para una prueba de concepto porque facilita observar el código, repetir la ejecución y examinar las transformaciones aplicadas.

No obstante, la ejecución correcta del flujo no equivale a la validación del modelo. La calidad de un sistema predictivo depende, en primer lugar, de la definición de la variable objetivo. En el código original, la etiqueta ganador_licitacion se generaba aleatoriamente y no mantenía una relación programada con las características del proveedor. Bajo esa estructura, el clasificador no podía aprender un patrón sustantivo que permitiera interpretar sus resultados como predicciones de adjudicación.

La reconstrucción metodológica sustituye esta etiqueta por una clase de compatibilidad sintética. Esta decisión fortalece el ejercicio porque permite establecer una relación transparente entre las variables y el resultado que se desea recuperar. Sin embargo, también debe reconocerse que el rendimiento del modelo dependerá de la regla utilizada para crear la clase. El algoritmo no descubrirá criterios reales, sino que intentará reconstruir una estructura introducida previamente por los autores.

Esta característica convierte al conjunto en un entorno controlado para verificar la lógica del prototipo. Si el modelo supera las líneas de base y recupera la clase con un desempeño estable, podrá afirmarse que el flujo de clasificación funciona dentro de la simulación. No podrá concluirse que el mismo comportamiento se mantendrá con información real.

La incorporación de una regresión logística y un clasificador ingenuo como referencias permite interpretar mejor el bosque aleatorio. En el manuscrito original, el algoritmo se presentaba como una elección directa, sin demostrar por qué era preferible frente a alternativas más sencillas. La comparación revisada permite establecer si la complejidad del bosque produce una mejora relevante.

Esta discusión es importante porque los sistemas vinculados con decisiones de alto impacto no deben adoptar modelos complejos únicamente por su popularidad o capacidad técnica. Rudin (2019) sostiene que, cuando existen modelos interpretables con un rendimiento adecuado, debe analizarse críticamente la necesidad de recurrir a cajas negras acompañadas por explicaciones aproximadas. En el prototipo, esta consideración implica que el bosque aleatorio solo se justifica si ofrece una ventaja comprobable frente a una alternativa más transparente.

La evaluación tampoco puede limitarse a la exactitud. La distribución aproximada de 70 % y 30 % entre las clases permite que un clasificador que prediga siempre la categoría mayoritaria obtenga un porcentaje aparentemente elevado de aciertos. Por ello, la exactitud balanceada, la precisión, la sensibilidad, F1 y las curvas asociadas con distintos umbrales ofrecen una visión más completa.

La matriz de confusión cumple una función central porque permite observar si el modelo reconoce ambas categorías o concentra sus predicciones en una sola. Esta información evita que un resultado numérico global oculte la incapacidad para identificar la clase minoritaria.

En el momento de la edición no deben incorporarse cifras definitivas que no hayan sido obtenidas mediante la ejecución del código corregido. Los resultados del modelo original no son compatibles

con la nueva variable objetivo, el nuevo procesamiento ni la exclusión del nombre del proveedor. Por consiguiente, las métricas deberán recalcularse antes de cerrar la versión publicada.

La decisión de eliminar la identidad del proveedor como predictor también fortalece la propuesta. En el código original, el nombre empresarial se transformaba mediante codificación binaria y podía influir en las estimaciones. Dado que los nombres se asignaban a registros aleatorios, cualquier relación aprendida sería artificial. En un entorno real, utilizar directamente la identidad de una empresa podría además favorecer la reproducción de patrones históricos y generar efectos desiguales sobre nuevos participantes.

La separación entre información descriptiva y predictiva permite conservar un identificador para presentar los casos sin permitir que ese identificador influya en el modelo. Esta práctica resulta coherente con la finalidad de comparar características y no reputaciones nominales.

Otro aporte metodológico es la diferenciación entre descripción y explicación. El manuscrito original presentaba como explicación una comparación entre los valores promedio de dos proveedores. Sin embargo, observar que un registro posee más experiencia, mayor personal o un costo superior no demuestra que esas variables hayan determinado la predicción.

La nueva arquitectura propone utilizar importancia por permutación para examinar el comportamiento global del modelo y SHAP para analizar las contribuciones locales. La primera permite observar cuánto disminuye el rendimiento cuando se altera una variable; la segunda permite identificar cómo las características desplazan una predicción desde un valor de referencia.

Estos métodos tampoco deben interpretarse como pruebas causales. Una variable puede influir en la puntuación porque el modelo aprendió una relación estadística dentro del conjunto, sin que ello signifique que modificarla produzca necesariamente un resultado real. La explicación describe al algoritmo y no sustituye la evaluación técnica o jurídica.

La interfaz revisada refuerza esta delimitación al sustituir expresiones como “ganador” o “proveedor recomendado” por “perfil con mayor puntuación sintética”. Esta modificación no representa un cambio meramente terminológico. Reduce el riesgo de que la salida sea interpretada como una decisión y recuerda que el modelo opera dentro de un escenario construido.

La eliminación del gráfico radar también responde a una revisión metodológica. La visualización original combinaba variables expresadas en años, horas, dólares, cantidades, valores binarios y proporciones. Sin una escala común, las variables de mayor magnitud dominaban la figura y podían producir una impresión visual equivocada. La comparación mediante perfiles normalizados permite representar las diferencias con mayor coherencia.

La propuesta de gobernanza de datos amplía el alcance del capítulo metodológico. El conjunto sintético deja de considerarse un recurso auxiliar y pasa a documentarse mediante una ficha que identifica su finalidad, versión, composición, usos permitidos y limitaciones. Esta práctica responde a la necesidad de que los datos sean evaluados antes de reutilizarse.

Gebru et al. (2021) plantean que la documentación de los conjuntos debe incluir información sobre motivación, composición, proceso de creación, mantenimiento y usos previstos. En el caso del prototipo, esta documentación resulta especialmente importante porque la semejanza temática con la contratación pública podría llevar a utilizar la base fuera de su propósito demostrativo.

La estrategia iterativa inspirada en Scrum también se conserva, pero se reclasifica. Su utilidad se encuentra en la organización del desarrollo, no en la validación científica. Dividir las actividades en iteraciones permite reconstruir el proceso seguido por los autores y conservar el backlog, los sprints y las revisiones como antecedentes del trabajo técnico. Sin embargo, no demuestra que el modelo sea preciso o que la interfaz haya sido evaluada por usuarios.

La factibilidad se interpreta bajo esta misma lógica. El prototipo es técnicamente viable como demostración porque se ejecuta con herramientas accesibles y un conjunto de tamaño moderado.

Esta viabilidad no permite afirmar que una plataforma institucional pueda desarrollarse sin costos o desplegarse con la misma infraestructura.

Una implementación real necesitaría especialistas, seguridad, almacenamiento, mantenimiento, monitoreo, integración con sistemas oficiales y mecanismos de auditoría. También requeriría evaluar las consecuencias de los errores y establecer responsabilidades. Por esta razón, la factibilidad económica no puede reducirse al costo de una computadora o a la disponibilidad gratuita de Python y Google Colab.

En conjunto, la discusión metodológica muestra que la principal fortaleza de la propuesta no se encuentra en haber construido un sistema listo para decidir sobre licitaciones, sino en haber desarrollado una base tecnológica susceptible de revisión y mejora. La reconstrucción editorial convierte esa base en una prueba de concepto más transparente, reproducible y coherente con la evidencia disponible.

Los resultados numéricos definitivos deberán incorporarse después de ejecutar la versión corregida del código. Hasta entonces, la discusión debe mantenerse centrada en el diseño, la coherencia de las operaciones, la gestión de los riesgos y las condiciones que una futura investigación deberá satisfacer.

Aportes y utilidad potencial del prototipo

El principal aporte del prototipo consiste en representar una arquitectura mediante la cual una puntuación generada por un modelo de aprendizaje automático puede acompañarse de información destinada a facilitar su comprensión. Esta integración responde a una necesidad relevante: los resultados algorítmicos no deberían presentarse de manera aislada cuando pueden influir en el análisis de alternativas dentro de entornos institucionales.

La propuesta no se limita al entrenamiento de un clasificador. Integra la generación y documentación de datos, el procesamiento de variables, la evaluación predictiva, la explicación de

resultados, la visualización y la revisión humana. Esta combinación permite abordar el prototipo como un sistema sociotécnico en el que el algoritmo constituye solo uno de los componentes. La calidad de la salida depende también de las decisiones adoptadas al formular el problema, seleccionar las variables, construir la etiqueta, preparar los datos y comunicar las limitaciones.

Desde una perspectiva tecnológica, la obra demuestra la posibilidad de articular herramientas disponibles en Python para construir un flujo reproducible. El uso de bibliotecas especializadas permite generar registros sintéticos, preparar variables, entrenar modelos, calcular métricas y representar resultados. La ejecución en un cuaderno computacional facilita que el lector observe las operaciones y repita el procedimiento, siempre que se documenten las versiones del entorno y del conjunto de datos.

Este aporte debe interpretarse dentro de los límites de una prueba de concepto. La ejecución del código demuestra que los componentes pueden conectarse, pero no que la herramienta se encuentre preparada para operar dentro de una entidad pública. Tampoco demuestra que el modelo pueda sustituir los procedimientos de evaluación establecidos en la contratación pública.

Desde una perspectiva metodológica, la reconstrucción del conjunto sintético permite mostrar la importancia de definir correctamente la variable objetivo. El código inicial generaba aleatoriamente el resultado que se pretendía predecir. La sustitución de esa etiqueta por una clase de compatibilidad sintética construida mediante una regla documentada transforma el ejercicio en un entorno controlado, donde puede evaluarse si el modelo recupera una relación introducida deliberadamente.

Este procedimiento posee valor didáctico porque permite observar cómo las decisiones tomadas durante la generación de los datos influyen posteriormente en las métricas y explicaciones. También evidencia que un algoritmo no descubre por sí mismo el significado de una categoría. Aprende las relaciones contenidas en los ejemplos proporcionados y reproduce, con distintos niveles de precisión, la estructura que encuentra en ellos.

La comparación con un clasificador ingenuo y una regresión logística constituye otro aporte metodológico. Esta estrategia permite establecer si el bosque aleatorio utiliza efectivamente las características del conjunto y si su complejidad aporta una mejora frente a alternativas más sencillas. En aplicaciones donde la comprensión y la rendición de cuentas poseen especial importancia, la selección de un modelo debe considerar no solo el rendimiento, sino también la posibilidad de examinar su funcionamiento. Rudin (2019) sostiene que, en decisiones de alto impacto, debe valorarse prioritariamente el uso de modelos interpretables cuando estos alcanzan un desempeño adecuado.

Desde la perspectiva de la inteligencia artificial explicable, la obra diferencia entre descripción, importancia global y explicación local. La comparación de valores entre perfiles permite observar diferencias en sus características, pero no demuestra cuáles de ellas determinaron la salida del algoritmo. La importancia por permutación examina las variables de las que depende el rendimiento general, mientras que los valores SHAP permiten representar la contribución de cada característica a una predicción específica (Lundberg & Lee, 2017).

Esta distinción contribuye a evitar explicaciones aparentes. Una gráfica puede presentar información relevante y, sin embargo, no explicar la predicción. Para que una explicación sea metodológicamente adecuada debe guardar correspondencia con el comportamiento efectivo del modelo y comunicar sus resultados de manera comprensible para el destinatario. Phillips et al. (2021) señalan que un sistema explicable debe proporcionar razones, asegurar que estas sean comprensibles, representar correctamente el proceso que generó la salida y reconocer sus límites de funcionamiento.

La propuesta también incorpora la gobernanza de datos como componente central. El conjunto sintético se acompaña de información sobre su procedencia, composición, versión, finalidad, usos permitidos y restricciones. Esta documentación resulta necesaria para evitar que la base sea reutilizada como si contuviera información real. Gebru et al. (2021) proponen documentar los

conjuntos mediante elementos que permitan comprender su motivación, creación, composición, mantenimiento y posibles usos.

El versionamiento del conjunto, del código y del modelo permite relacionar cada resultado con las condiciones específicas bajo las cuales fue producido. Esta trazabilidad es indispensable cuando las transformaciones, parámetros o fórmulas pueden cambiar entre ejecuciones. Sin un registro de versiones, dos puntuaciones diferentes podrían presentarse como resultados del mismo sistema, aunque hayan sido generadas mediante configuraciones distintas.

Otro aporte se encuentra en la separación entre puntuación y decisión. La interfaz revisada no utiliza expresiones como “elegir ganador” o “seleccionar proveedor”. Presenta una puntuación sintética, sus factores relacionados y una advertencia sobre su alcance. La decisión permanece fuera del sistema y bajo responsabilidad humana.

Esta delimitación favorece una confianza calibrada. La finalidad de la explicación no consiste en convencer al usuario de aceptar la salida, sino en proporcionarle elementos para examinarla. Una explicación responsable debe permitir tanto comprender como cuestionar el resultado. Cuando la información es insuficiente, se encuentra fuera de rango o genera una predicción inestable, el sistema debe comunicar esas condiciones.

Desde el punto de vista editorial y formativo, la obra puede servir como recurso para analizar los desafíos que surgen al trasladar conceptos de aprendizaje automático hacia problemas institucionales. El prototipo permite discutir temas como calidad de datos, variable objetivo, desbalance de clases, explicabilidad, trazabilidad, supervisión humana y límites de generalización.

La utilidad potencial de la propuesta no se encuentra, por tanto, en haber resuelto la selección de proveedores, sino en proporcionar una estructura para estudiar cómo podría diseñarse responsablemente una herramienta de apoyo. Esta estructura puede orientar investigaciones

posteriores que incorporen fuentes reales, criterios especializados, evaluación con usuarios y mecanismos de auditoría.

Tabla 21

Aportes del prototipo y límites de la evidencia disponible

Dimensión	*Aporte principal*	*Evidencia disponible*	*Límite*
Tecnológica	Integración de datos, modelo, visualización y explicación	Código ejecutable en un entorno de cuaderno	No corresponde a una plataforma productiva
Metodológica	Construcción transparente de una clase sintética	Regla, variables y pesos documentados	No representa criterios oficiales
Predictiva	Comparación entre líneas de base y bosque aleatorio	Métricas obtenidas sobre datos sintéticos	No permite predecir adjudicaciones reales
Explicativa	Diferenciación entre importancia global y contribución local	Permutación y SHAP propuestos para el flujo	La explicación describe al modelo, no relaciones causales
Gobernanza	Documentación, versionamiento y usos permitidos	Ficha del conjunto y registro de transformaciones	No sustituye una política institucional
Ética	Separación entre puntuación y decisión	Advertencias y supervisión humana	No elimina todos los riesgos de uso indebido
Formativa	Ejemplo para analizar límites de la IA aplicada	Arquitectura y discusión metodológica	Requiere conocimientos complementarios para su implementación

Nota. *Los aportes corresponden a una prueba de concepto desarrollada con datos sintéticos y no demuestran impacto sobre procedimientos reales.*

Riesgo de automatización excesiva

Uno de los principales riesgos consiste en que los usuarios atribuyan a la puntuación un nivel de autoridad superior al que posee. La presencia de gráficos, porcentajes y explicaciones puede generar una percepción de objetividad, aunque el resultado dependa de decisiones humanas adoptadas durante la construcción del conjunto y del modelo.

La interfaz debe evitar diseños que destaquen una alternativa como ganadora. También debe comunicar la incertidumbre, los límites de la simulación y las variables que no fueron consideradas. La supervisión humana no debe convertirse en una confirmación rutinaria de la salida algorítmica.

Riesgo de reproducción de patrones históricos

En una futura versión con datos reales, utilizar adjudicaciones pasadas como variable objetivo podría reproducir relaciones anteriores. Las empresas con mayor participación histórica podrían recibir puntuaciones superiores, mientras los nuevos proveedores dispondrían de menos registros.

La ausencia de una adjudicación anterior no significa necesariamente falta de capacidad. Puede responder a condiciones de participación, requisitos, disponibilidad de información o características del mercado. El modelo no puede distinguir automáticamente entre una regularidad legítima y una desigualdad histórica.

Riesgo de variables sustitutas

Determinadas variables aparentemente neutrales pueden funcionar como aproximaciones de otras características. La ubicación, el tamaño de la empresa, la cantidad de personal o la frecuencia de contratos pueden relacionarse con condiciones económicas y estructurales que produzcan efectos desiguales.

La exclusión de variables sensibles no garantiza ausencia de sesgo cuando otras características permiten inferirlas indirectamente. Por ello, una implementación real requeriría evaluar los resultados entre distintos grupos y justificar la pertinencia de cada predictor.

Riesgo de explicaciones persuasivas

Una explicación puede estar diseñada para aumentar la aceptación del sistema sin describir fielmente su funcionamiento. Mostrar frases como “se recomienda debido a su experiencia” puede resultar convincente, aunque esa variable haya tenido poca influencia en la predicción.

Los métodos explicativos deben estar vinculados con las salidas reales del modelo. También deben comunicar cuando una predicción no posee una explicación estable o cuando las variables más influyentes no resultan coherentes con el propósito institucional.

Riesgo de uso fuera del alcance

El conjunto, el código o las puntuaciones podrían reutilizarse fuera de su finalidad educativa. Para reducir esta posibilidad, el repositorio debe contener una advertencia explícita y una descripción de usos prohibidos.

El sistema no debe emplearse para:

- evaluar empresas identificables;
- establecer listas de proveedores;
- adjudicar contratos;
- justificar exclusiones;
- medir desempeño institucional;
- ni estimar ahorro público.
- Condiciones mínimas para una futura aplicabilidad

Una implementación real requeriría cumplir, al menos, las siguientes condiciones:

- Definir el usuario, la tarea y el objeto de la recomendación.
- Determinar si el sistema recomienda oportunidades a proveedores o apoya el análisis de ofertas.
- Identificar la base jurídica para el tratamiento de los datos.
- Utilizar información oficial documentada y suficientemente representativa.
- Definir la unidad de análisis y evitar mezclar proveedores, ofertas y contratos.
- Validar las variables con especialistas técnicos, jurídicos y administrativos.
- Separar requisitos obligatorios de criterios susceptibles de puntuación.
- Comparar modelos interpretables y complejos.
- Evaluar desempeño, calibración, estabilidad y posibles efectos desiguales.
- Diseñar explicaciones adaptadas a los perfiles de usuario.

- Incorporar mecanismos de auditoría y registro de versiones.
- Realizar pruebas con usuarios mediante protocolos documentados.
- Establecer quién puede cuestionar, corregir o detener el sistema.
- Ejecutar un piloto controlado antes de cualquier despliegue.
- Mantener la decisión final bajo responsabilidad institucional.

El NIST (2023) establece que la gestión de riesgos debe integrar gobernanza, contextualización, medición y respuesta durante todo el ciclo de vida. Esta orientación implica que la evaluación ética no puede añadirse únicamente después del entrenamiento. Debe comenzar al definir el propósito y continuar durante la recopilación de datos, el diseño, las pruebas, el despliegue y el monitoreo.

Conclusiones

La contratación pública constituye un entorno de decisión complejo en el que convergen requisitos jurídicos, técnicos, económicos y documentales. La incorporación de inteligencia artificial puede contribuir al procesamiento y organización de información, pero no elimina la obligación de justificar las decisiones, mantener su trazabilidad y conservar la responsabilidad humana.

La revisión del manuscrito permitió establecer que la propuesta desarrollada corresponde a una prueba de concepto y no a un sistema institucional validado. El prototipo integra elementos funcionales de generación de datos, clasificación, comparación, visualización y explicación, pero opera exclusivamente sobre registros sintéticos. Por consiguiente, sus resultados no representan el comportamiento de empresas reales ni permiten estimar probabilidades de adjudicación.

El bosque aleatorio puede conservarse como algoritmo principal del prototipo debido a su capacidad para representar relaciones no lineales y combinar múltiples árboles. Sin embargo, su utilización debe justificarse mediante la comparación con un clasificador ingenuo y una regresión logística. Cuando un modelo sencillo alcanza resultados semejantes, su mayor interpretabilidad puede convertirlo en una alternativa preferible.

La evaluación no debe limitarse a la exactitud global. La matriz de confusión, la exactitud balanceada, la precisión, la sensibilidad, F1, ROC-AUC y las medidas relacionadas con precisión-sensibilidad permiten examinar con mayor detalle el comportamiento del modelo ante ambas clases. Las métricas definitivas deberán calcularse mediante la versión corregida del código antes de cerrar la publicación.

La explicabilidad representa un componente necesario, pero no suficiente, para el uso responsable de modelos predictivos. La comparación visual de características describe diferencias entre perfiles, mientras la importancia por permutación y SHAP permiten examinar la dependencia global y las contribuciones locales del modelo. Estas técnicas explican el comportamiento algorítmico, pero no demuestran causalidad ni legitimidad institucional.

La gobernanza de datos constituye uno de los aportes centrales de la reconstrucción. Documentar la procedencia, finalidad, composición, versión, transformaciones y restricciones del conjunto permite evitar usos incompatibles con su naturaleza sintética. La trazabilidad debe extenderse al código, los parámetros, las métricas, las explicaciones y la intervención humana.

La sustitución de nombres empresariales por identificadores ficticios resulta obligatoria para evitar que puntuaciones artificiales sean asociadas con organizaciones reales. Del mismo modo, el nombre del proveedor no debe incorporarse como predictor, porque no representa una característica técnica y podría generar relaciones nominales sin fundamento.

La interfaz debe comunicar la salida como una puntuación sintética y no como probabilidad real de ganar una licitación. También debe excluir expresiones que sugieran selección o adjudicación automática. La presentación de advertencias, contribuciones de variables y limitaciones favorece una revisión más crítica del resultado.

La supervisión humana no debe limitarse a confirmar la alternativa priorizada. Debe permitir cuestionar los datos, examinar la explicación, identificar casos fuera del alcance y rechazar la salida

cuando existan razones justificadas. La responsabilidad de una decisión pública permanece en las personas y autoridades competentes.

La factibilidad demostrativa del prototipo se encuentra sustentada por la disponibilidad de herramientas de código abierto y por la posibilidad de ejecutar el flujo con recursos moderados. No obstante, la factibilidad de una plataforma institucional no puede inferirse a partir de esta prueba. Una implementación real necesitaría infraestructura, seguridad, integración, mantenimiento, personal especializado, evaluación jurídica y monitoreo continuo.

La obra aporta una arquitectura conceptual y metodológica susceptible de ampliación. Su valor reside en mostrar cómo pueden combinarse un modelo predictivo, mecanismos explicativos, documentación de datos y revisión humana, además de evidenciar las condiciones que deben cumplirse antes de trasladar una demostración tecnológica hacia un entorno institucional.

En consecuencia, el prototipo debe publicarse como una experiencia académica de diseño y análisis crítico. Su presentación transparente, acompañada de código reproducible y datos sintéticos claramente identificados, ofrece una base para futuras investigaciones sin afirmar resultados que excedan la evidencia disponible.

Recomendaciones y líneas futuras

La primera recomendación consiste en ejecutar nuevamente todo el flujo después de incorporar las correcciones metodológicas. La nueva ejecución deberá utilizar identificadores ficticios, la clase de compatibilidad sintética, la exclusión del identificador como predictor, las particiones estratificadas y los modelos de referencia. Las tablas y figuras definitivas deberán generarse únicamente a partir de esa versión.

También debe verificarse que la puntuación utilizada para crear la clase no permanezca entre los predictores. La presencia de esa variable o de sus transformaciones produciría una fuga de información y generaría métricas artificialmente elevadas.

La segunda recomendación corresponde a la consolidación de la explicabilidad. Se propone calcular importancia por permutación sobre el conjunto de prueba y utilizar SHAP para examinar predicciones individuales. Las explicaciones deberán compararse con la regla sintética para comprobar que el modelo utiliza variables coherentes con la estructura incorporada.

No resulta necesario implementar simultáneamente una gran cantidad de métodos XAI. La combinación de una explicación global y una local es suficiente para la prueba de concepto. Incorporar técnicas adicionales sin una pregunta definida aumentaría la complejidad sin garantizar mayor comprensión.

La tercera recomendación consiste en diseñar una evaluación específica de las explicaciones. La calidad no debe medirse únicamente por la apariencia de los gráficos. Es necesario examinar su fidelidad, estabilidad, comprensibilidad y utilidad. Una evaluación inicial podría comparar si los usuarios identifican correctamente las variables que aumentan y reducen la puntuación y si comprenden que el resultado no equivale a una decisión.

La cuarta recomendación es elaborar un repositorio complementario. Este debe contener:

- el código corregido;
- el conjunto sintético original;
- la versión procesada;
- el archivo de requisitos;
- las instrucciones de ejecución;
- la ficha de gobernanza;
- el historial de versiones;
- las métricas;
- y las advertencias sobre usos no permitidos.

La quinta recomendación consiste en desarrollar una línea de investigación basada en datos reales, pero únicamente después de definir con precisión la finalidad del sistema. Existen al menos dos posibles orientaciones.

La primera podría recomendar oportunidades contractuales a proveedores a partir de la compatibilidad entre sus perfiles y los requisitos publicados. En esta configuración, el proveedor funcionaría como usuario y el procedimiento como objeto recomendado.

La segunda podría apoyar a una entidad en la organización de información sobre ofertas. Esta modalidad implicaría mayores exigencias jurídicas y éticas, porque la herramienta podría influir en una decisión administrativa. En ese caso, el sistema debería operar únicamente como apoyo y mantener fuera del modelo la adjudicación final.

La sexta recomendación corresponde a la obtención de datos reales. La investigación deberá documentar:

- fuente;
- período;
- unidad de análisis;
- variables;
- criterios de inclusión;
- valores faltantes;
- duplicados;
- actualizaciones;
- permisos;
- y restricciones de uso.

La séptima recomendación consiste en validar las variables con especialistas. Profesionales de contratación pública, informática, análisis de datos, auditoría y derecho deberían examinar la

pertinencia de cada característica. Las variables que representan requisitos obligatorios deberán tratarse mediante reglas o restricciones, no como elementos compensables dentro de una suma.

La octava recomendación es evaluar modelos interpretables antes de adoptar algoritmos complejos. La regresión logística, los árboles pequeños, los sistemas de puntuación y las reglas pueden proporcionar resultados suficientes para determinadas tareas. La complejidad solo debería incrementarse cuando exista una mejora demostrable y cuando los métodos explicativos mantengan fidelidad.

La novena recomendación corresponde a la evaluación de sesgos. Una versión con datos reales deberá examinar si el modelo favorece sistemáticamente a empresas con mayor historial, tamaño, ubicación o capacidad administrativa. También deberá comprobar si las variables actúan como sustitutos de condiciones que no deberían influir en el resultado.

Tabla 22

Ruta propuesta para la maduración del prototipo

Fase	*Propósito*	*Actividades principales*	*Evidencia requerida*
1. Consolidación sintética	Corregir y reproducir la prueba de concepto	Reconstruir etiqueta, excluir identificadores, evaluar modelos e implementar XAI	Código ejecutable, métricas y explicaciones
2. Revisión especializada	Validar la lógica conceptual	Examinar variables, restricciones, finalidad y riesgos	Informes técnicos y jurídicos
3. Preparación de datos reales	Construir un conjunto documentado	Definir fuentes, unidad de análisis, calidad, permisos y gobernanza	Ficha del conjunto y protocolo de tratamiento
4. Validación predictiva externa	Evaluar el modelo en información no utilizada	Comparar líneas de base, calibración, estabilidad y sesgos	Informe de validación independiente
5. Evaluación con usuarios	Analizar comprensión y utilidad	Pruebas de tareas, entrevistas y revisión de explicaciones	Instrumentos, muestra, resultados y consentimiento
6. Piloto controlado	Observar el funcionamiento en un entorno limitado	Integración parcial, monitoreo y supervisión reforzada	Registro de incidentes, métricas y decisiones humanas
7. Decisión de implementación	Determinar si el sistema debe desplegarse	Evaluación técnica, ética, jurídica, económica y organizacional	Dictamen institucional integral

Nota. *El avance entre fases debe depender de la evidencia obtenida y no únicamente de la disponibilidad técnica del sistema.*

La proyección futura del prototipo deberá mantener como principio que la inteligencia artificial actúe como apoyo y no como sustituto de las competencias institucionales. El sistema puede organizar información, producir estimaciones y mostrar relaciones, pero la interpretación de la necesidad pública, la aplicación de la normativa y la responsabilidad de la decisión corresponden a las personas autorizadas.

El desarrollo posterior también deberá conservar la capacidad de detener o retirar el modelo. Una herramienta que presenta deterioro en sus datos, métricas, explicaciones o efectos no debería continuar utilizándose únicamente porque ya fue integrada a una plataforma.

La propuesta desarrollada evidencia que la construcción de sistemas explicables no depende únicamente de seleccionar un algoritmo o añadir gráficos a una interfaz. Requiere definir correctamente el problema, documentar los datos, evaluar el modelo, verificar la fidelidad de las explicaciones y establecer límites claros para la intervención automatizada. El prototipo reconstruido ofrece una base académica para continuar investigando estas relaciones, pero su aplicación futura deberá estar condicionada por evidencia, gobernanza, revisión especializada y supervisión humana.

BIBLIOGRAFÍA

Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE Transactions on Knowledge and Data Engineering, 17(6), 734–749. https://doi.org/10.1109/TKDE.2005.99

Adomavicius, G., & Tuzhilin, A. (2011). Context-aware recommender systems. En F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.), Recommender systems handbook (pp. 217–253). Springer. https://doi.org/10.1007/978-0-387-85820-3_7

Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 58, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). Classification and regression trees. Wadsworth & Brooks/Cole Advanced Books & Software.

Burke, R. (2002). Hybrid recommender systems: Survey and experiments. User Modeling and User-Adapted Interaction, 12(4), 331–370. https://doi.org/10.1023/A:1021240730564

Condori. (2021). "SISTEMAS DE RECOMENDACIÓN PARA LA COMPRA Y VENTA DE ACCIONES EN LA BOLSA DE VALORES USANDO MACHINE LEARNING.

Couñago. (2021). Sistemas de recomendación basado en contenido.

Cuji, Gavilanes, & Sanchez. (2017). EL MODELO PREDICTIVO DE DESERCIÓN ESTUDIANTIL BASADO EN ARBOLES DE DECISIONES.

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv. https://doi.org/10.48550/arXiv.1702.08608

Ecuador. Asamblea Constituyente. (2008). Ley Orgánica del Sistema Nacional de Contratación Pública (Registro Oficial Suplemento 395, 4 de agosto de 2008; última reforma 26 de junio de 2025). Servicio Nacional de Contratación Pública. https://portal.compraspublicas.gob.ec/sercop/wp-content/uploads/2025/07/LEY-ORGANICA-DEL-SISTEMA-NACIONAL-DE-CONTRATACION-PUBLICA..pdf

Fawcett, T. (2006). An introduction to ROC analysis. Pattern Recognition Letters, 27(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

Fonseca, & Cornelio. (2022). Sistemas de Recomendación Híbridos.

Fonseca, B. B. (2021). Sistemas de Recomendación.

Gebru, T., Morgenstern, J., Vecchione, B., Wortman Vaughan, J., Wallach, H., Daumé III, H., & Crawford, K. (2021). Datasheets for datasets. Communications of the ACM, 64(12), 86–92. https://doi.org/10.1145/3458723

Goldstein, A., Kapelner, A., Bleich, J., & Pitkin, E. (2015). Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation. Journal of Computational and Graphical Statistics, 24(1), 44–65. https://doi.org/10.1080/10618600.2014.907095

Gonzalo. (2022). EVALUACIÓN DE MODELOS DE INTELIGENCIA ARTIFICIAL EXPLICABLE.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press. https://www.deeplearningbook.org/

Google. (s. f.). Google Colaboratory: Frequently asked questions. Recuperado el 13 de julio de 2026, de https://research.google.com/colaboratory/faq.html

Hardy. (2021). Inteligencia artificial.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction (2.ª ed.). Springer. https://doi.org/10.1007/978-0-387-84858-7

Herlocker, J. L., Konstan, J. A., Terveen, L. G., & Riedl, J. T. (2004). Evaluating collaborative filtering recommender systems. ACM Transactions on Information Systems, 22(1), 5–53. https://doi.org/10.1145/963770.963772

Hernández, N. B. (2020). licitaciones públicas.

Hernández, N. B. (2020). Transparencia y equidad en las recomendaciones basadas en IA en procesos de licitación pública.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. MIS Quarterly, 28(1), 75–105. https://www.jstor.org/stable/25148625

Hu, Y., Koren, Y., & Volinsky, C. (2008). Collaborative filtering for implicit feedback datasets. En Proceedings of the 2008 Eighth IEEE International Conference on Data Mining (pp. 263–272). IEEE. https://doi.org/10.1109/ICDM.2008.22

Janiesch, C. (2021). Explanación Base Learning" de Gerald Dejong.

Jordon, J., Szpruch, Ł., Houssiau, F., Bottarelli, M., Cherubin, G., Maple, C., Cohen, S. N., & Weller, A. (2022). Synthetic data—What, why and how? arXiv. https://doi.org/10.48550/arXiv.2205.03257

Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix factorization techniques for recommender systems. Computer, 42(8), 30–37. https://doi.org/10.1109/MC.2009.263

Lipton, Z. C. (2018). The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery. Queue, 16(3), 31–57. https://doi.org/10.1145/3236386.3241340

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. En I. Guyon, U. von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), Advances in neural information processing systems 30 (pp. 4765–4774). Curran Associates. https://proceedings.neurips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (1955). A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955. Dartmouth College. https://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf

Mimbrera, E. (2021). Inteligencia Artificial Explicable.

Moreno, V. (2023). Sistemas de Recomendación Basados en Memoria (Nearest Neighbour).

Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R., & Yu, B. (2019). Interpretable machine learning: Definitions, methods, and applications. Proceedings of the National Academy of Sciences of the United States of America, 116(44), 22071–22080. https://doi.org/10.1073/pnas.1900654116

National Institute of Standards and Technology. (2023). Artificial intelligence risk management framework (AI RMF 1.0) (NIST AI 100-1). https://doi.org/10.6028/NIST.AI.100-1

Organisation for Economic Co-operation and Development. (2024). Explanatory memorandum on the updated OECD definition of an AI system (OECD Artificial Intelligence Papers No. 8). OECD Publishing. https://doi.org/10.1787/623da898-en

Pathak. (2024). La XAI .

Pazzani, M. J., & Billsus, D. (2007). Content-based recommendation systems. En P. Brusilovsky, A. Kobsa, & W. Nejdl (Eds.), The adaptive web: Methods and strategies of web personalization (Lecture Notes in Computer Science, Vol. 4321, pp. 325–341). Springer. https://doi.org/10.1007/978-3-540-72079-9_10

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., VanderPlas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825–2830. https://www.jmlr.org/papers/v12/pedregosa11a.html

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. Journal of Management Information Systems, 24(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302

Phillips, P. J., Hahn, C. A., Fontana, P. C., Yates, A. N., Greene, K., Broniatowski, D. A., & Przybocki, M. A. (2021). Four principles of explainable artificial intelligence (NISTIR 8312). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.IR.8312

Portal Compras públicas. (2023). Obtenido de https://www.compraspublicas.gob.ec/ProcesoContratacion/compras/

Ramirez. (2018). Deep Learning.

Resnick, P., & Varian, H. R. (1997). Recommender systems. Communications of the ACM, 40(3), 56–58. https://doi.org/10.1145/245108.245121

Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., & Riedl, J. (1994). GroupLens: An open architecture for collaborative filtering of netnews. En Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work (pp. 175–186). Association for Computing Machinery. https://doi.org/10.1145/192844.192905

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. En Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 1135–1144). Association for Computing Machinery. https://doi.org/10.1145/2939672.2939778

Ricci, F., Rokach, L., & Shapira, B. (Eds.). (2022). Recommender systems handbook (3.ª ed.). Springer. https://doi.org/10.1007/978-1-0716-2197-4

Robledo. (2023). LOS SISTEMAS DE RECOMENDACIÓN EN UN MARKTEPLACE DE APIS.

Rojo. (2019). MODELO PREDICTIVO DE ANÁLISIS DE RIESGO CREDITICIO USANDO MACHINE LEARNING EN UNA ENTIDAD DEL SECTOR MICROFINANCIERO.

Rubio, G., Charry, P., & Perez, C. (2022). El Machine Learning.

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1(5), 206–215. https://doi.org/10.1038/s42256-019-0048-x

Russell, S. J., & Norvig, P. (2021). Artificial intelligence: A modern approach (4.ª ed.). Pearson.

Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLOS ONE, 10(3), Article e0118432. https://doi.org/10.1371/journal.pone.0118432

Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001). Item-based collaborative filtering recommendation algorithms. En Proceedings of the 10th International Conference on World Wide Web (pp. 285–295). Association for Computing Machinery. https://doi.org/10.1145/371920.372071

Schwaber, K., & Sutherland, J. (2020). The Scrum guide: The definitive guide to Scrum—The rules of the game. Scrum Guides. https://scrumguides.org/docs/scrumguide/v2020/2020-Scrum-Guide-US.pdf

Scikit-learn developers. (s. f.). Scikit-learn user guide and API reference. Recuperado el 13 de julio de 2026, de https://scikit-learn.org/stable/

SERCORP. (2020). Obtenido de https://portal.compraspublicas.gob.ec/sercop/cat_normativas/oficios-2020

SERCORP. (2022).

SERCORP. (2024). Sistema Nacional de Contratación Pública.

Sistema Oficial de Contratación Pública. (2021). Obtenido de https://www.compraspublicas.gob.ec/ProcesoContratacion/compras/

Strobl, C., Boulesteix, A.-L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. BMC Bioinformatics, 8, Article 25. https://doi.org/10.1186/1471-2105-8-25

Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction (2.ª ed.). MIT Press. http://incompleteideas.net/book/the-book-2nd.html

Valenzuela, Correa, C., & Rendon, D. (2022). Sistemas de Recomendación Basado en Filtrados Colaborativos.

Vaquero. (2020). artificial, Inteligencia.

Xia. (2023). DESARROLLO DE UN SISTEMA DE RECOMENDACIÓN PARA UNA EMPRESA DE SERVICIOS ONLINE.

Zhang, Y., & Chen, X. (2020). Explainable recommendation: A survey and new perspectives. Foundations and Trends in Information Retrieval, 14(1), 1–102. https://doi.org/10.1561/1500000066

Acerca de la obra:

La presente obra analiza la convergencia entre inteligencia artificial explicable, modelos predictivos, sistemas de recomendación y toma de decisiones en el contexto de la contratación pública. Su propósito es fundamentar y documentar el diseño de un prototipo experimental de apoyo analítico capaz de procesar variables asociadas con perfiles sintéticos de proveedores tecnológicos, generar puntuaciones mediante modelos de clasificación y presentar información explicativa que facilite la revisión humana de los resultados.

Sobre los autores:

Miguel Giovanny Molina Villacís
https://orcid.org/0000-0002-7080-2354
Universidad de Guayaquil

Ingeniero en Electrónica y Telecomunicaciones, máster en Telecomunicaciones y en Administración de Empresas. Docente y director de proyectos de investigación de la Universidad de Guayaquil, con experiencia en tecnologías de la información, inteligencia artificial y aprendizaje automático.

María Fernanda Molina Miranda
https://orcid.org/0000-0002-4237-4364
Universidad de Guayaquil

Ingeniera en Telemática y máster en Redes y Servicios Telemáticos. Docente e investigadora de la Universidad de Guayaquil, cuya producción científica se orienta al desarrollo de la inteligencia artificial explicable y su aplicación en contextos tecnológicos.

Angel Eduardo Cuenca Ortega
https://orcid.org/0000-0001-7798-611X
Universidad de Guayaquil

Ingeniero de Sistemas, máster en Ingeniería del Software y doctor en Informática. Docente investigador de la Universidad de Guayaquil, especializado en lógica de reescritura, razonamiento simbólico y verificación formal de sistemas complejos.

Luis Arturo Espin Pazmiño
https://orcid.org/0000-0002-1663-2489
Universidad de Guayaquil

Doctor en Ciencias Informáticas, magíster en Administración de Empresas con mención en Telecomunicaciones e ingeniero en Telecomunicaciones. Se desempeña como docente investigador de la Universidad de Guayaquil, con experiencia académica en informática, telecomunicaciones y diseño curricular.

Renzo Rogelio Padilla Gómez
https://orcid.org/0000-0003-4301-1335
Universidad de Guayaquil

Ingeniero en Sistemas Computacionales y magíster en Docencia y Gerencia en Educación Superior y en Educación Informática. Docente e investigador de la Universidad de Guayaquil, con trayectoria en programación, inteligencia artificial, robótica y tecnologías de la información.

EDUINCA
EDITORIAL DE EDUCACIÓN, INVESTIGACIÓN Y CULTURA ACADÉMICA

ISBN: 978-9907-829-07-5