

# Prototipo Funcional Basado en Inteligencia Artificial para Evaluar el Riesgo Crediticio que asumen los Ahorrantes en Entidades Financieras Costarricenses Reguladas por la SUGEF

José Andrés Lizano-Gallegos<sup>1,2\*</sup>, Adolfo Jara Cordero<sup>2,3†</sup> and  
Luis G. Leon-Vega<sup>1,2†</sup>

<sup>1\*</sup>Universidad Cenfotec, Costa Rica.

<sup>2</sup>Universidad Cenfotec, Costa Rica.

<sup>3</sup>Universidad Cenfotec, Costa Rica.

\*Corresponding author(s). E-mail(s): [jlizanog@ucenfotec.ac.cr](mailto:jlizanog@ucenfotec.ac.cr);  
Contributing authors: [ajara@ucenfotec.ac.cr](mailto:ajara@ucenfotec.ac.cr); [leon@ucenfotec.ac.cr](mailto:leon@ucenfotec.ac.cr);

<sup>†</sup>These authors contributed equally to this work.

## Abstract

Every day, savers and investors place their money in financial institutions hoping their capital will be safe and will grow over time. Yet behind each decision lies a complex reality: understanding the true financial health of these institutions is difficult, and the information needed—although publicly available is often overwhelming, technical, and hard to interpret. In Costa Rica, the Superintendence of Financial Institutions (SUGEF) publishes a vast set of indicators meant to inform the public, but the sheer volume and multidimensional nature of these data create a gap between what is available and what people can actually use to make confident financial decisions. To solve this gap, this work introduces WISE Investors; it transforms dispersed financial data into clear, accessible insights through a combination of machine learning, interactive visual tools, and a conversational agent powered by Large Language Models that combines RAG (semantic vector search) with AI-generated SQL. The analytical methodology integrates Isolation Forest to detect unusual behaviors, K-Means to group institutions with similar financial profiles, and synthetic indicators that allow users to compare entities month by month. This work also proposes a prudential, percentile-based ranking approach that helps highlight meaningful trends while avoiding misleading short-term fluctuations. Experimental evaluation with 50 financial-domain

queries shows that the conversational agent achieves 100% fidelity, a maximum semantic similarity of 28.1%, and response coverage of up to 86% depending on the model. In terms of efficiency, the recommended configuration (OpenAI gpt-5.4-mini) delivers answers in 7.3 seconds on average with a consumption of 4955 tokens per query, offering the best balance between accuracy, cost, and scalability. The study contributes three key advances: a method to convert public financial indicators into an interpretable measure of institutional performance; a Risk-aware Ranking that incorporates historical context to provide more stable and realistic evaluations; and an integrated platform where users can explore results visually and ask questions in natural language. This prototype aims to reduce information asymmetry and empower people to make more informed, confident financial decisions.

**Keywords:** Riesgo crediticio, Inteligencia artificial aplicada, Análisis de datos financieros, entidades financieras, SUGEF, HybridRAG, Agent

## 1 Introduction

La intermediación financiera desempeña un papel esencial en el funcionamiento de las economías modernas al canalizar el ahorro hacia actividades productivas mediante la captación de recursos y la colocación de créditos [1]. En este proceso, los ahorrantes e inversionistas depositan sus recursos en entidades financieras con la expectativa de preservar su capital y obtener una rentabilidad acorde con el riesgo asumido. Sin embargo, esta decisión implica inevitablemente la exposición al riesgo de crédito de la entidad financiera seleccionada, cuya evaluación suele requerir conocimientos especializados sobre indicadores financieros, solvencia patrimonial, calidad de cartera, eficiencia operativa y otros elementos que, aunque se encuentran disponibles públicamente, resultan de difícil acceso o interpretación para la mayoría de los usuarios.

En Costa Rica, la Superintendencia General de Entidades Financieras (SUGEF) [2] publica periódicamente un amplio conjunto de indicadores financieros correspondientes a las entidades supervisadas, constituyendo una valiosa fuente de información para evaluar su situación financiera. No obstante, la cantidad de variables disponibles, la naturaleza multidimensional del problema y la ausencia de herramientas orientadas al usuario final dificultan la realización de análisis comparativos que permitan valorar, de manera objetiva, el riesgo de crédito asumido al mantener recursos en una determinada entidad financiera. En la práctica, muchas decisiones de ahorro e inversión continúan realizándose sin un análisis sistemático e integrado de la información financiera disponible, lo que genera una importante asimetría entre la información publicada por el supervisor y la capacidad de los usuarios para convertirla en conocimiento para la toma de decisiones.

La naturaleza multidimensional de la información financiera hace que la comparación simultánea entre entidades exceda las capacidades de los enfoques tradicionales basados en el análisis aislado de indicadores individuales. En este contexto, el aprendizaje automático ha demostrado ser una herramienta eficaz para el análisis

de grandes volúmenes de información y la identificación de patrones complejos. Particularmente, los métodos no supervisados permiten descubrir estructuras latentes en los datos sin requerir etiquetas previas, lo que resulta útil cuando no existe una clasificación objetiva de las observaciones [3]. En el ámbito financiero, estas técnicas han sido empleadas para la detección de anomalías, la segmentación de entidades con características similares y el análisis exploratorio de información multidimensional. Sin embargo, gran parte de estas aplicaciones se orienta al análisis interno de instituciones financieras o a procesos de supervisión prudencial, mientras que existen menos propuestas dirigidas a transformar dichos resultados en herramientas accesibles para ahorrantes e inversionistas.

Con el propósito de contribuir a reducir esta brecha, este trabajo propone *WISE Investors*: un sistema inteligente de apoyo a la decisión que transforma información financiera pública, dispersa y técnicamente compleja en herramientas objetivas para apoyar decisiones de ahorro e inversión, como se ilustra en Fig. 1. El sistema integra una metodología analítica basada en aprendizaje automático, una plataforma interactiva para la exploración visual de resultados y un agente conversacional fundamentado en modelos de lenguaje de gran escala (*Large Language Models, LL*), capaz de responder consultas en lenguaje natural sobre las entidades financieras supervisadas. La metodología combina detección de observaciones atípicas mediante *Isolation Forest*, segmentación de perfiles financieros con *K-Means* y la construcción de indicadores sintéticos para generar rankings comparativos desde una perspectiva mensual y bajo un enfoque temporal prudencial (*Risk-aware*) basado en percentiles.

Las principales contribuciones científicas y tecnológicas de la investigación son tres. En primer lugar, se desarrolla una metodología para transformar cuatro indicadores financieros públicos en una medida relativa e interpretable del desempeño de las entidades, apoyada en una representación vectorial, técnicas de aprendizaje no supervisado y la proyección sobre un perfil financiero ideal. En segundo lugar, se propone el *Risk-aware Ranking*, que incorpora una ventana histórica y una regla de agregación prudencial para reducir la influencia de mejoras transitorias sin ocultar deterioros recientes. En tercer lugar, se integran el motor analítico, las visualizaciones interactivas y una interfaz conversacional dentro de *WISE Investors*, de modo que los resultados puedan ser explorados y consultados por usuarios con distintos niveles de conocimiento técnico.

El resto del artículo se estructura de la siguiente manera. La Sección 2 presenta el estado del arte y analiza las principales fuentes de información, soluciones existentes y aproximaciones académicas relacionadas con el análisis financiero, identificando el vacío que motiva esta investigación. La Sección 3 describe los métodos utilizados, incluyendo los datos, el fundamento matemático y la construcción de la metodología *WISE*, la arquitectura del agente conversacional y la plataforma desarrollada. La Sección 4 presenta los principales resultados obtenidos, tanto de la metodología y la plataforma como de la evaluación experimental del agente conversacional. Posteriormente, la Sección 5 discute los resultados, alcances y limitaciones de la propuesta. Finalmente, la Sección 6 presenta las principales conclusiones y líneas futuras de investigación.

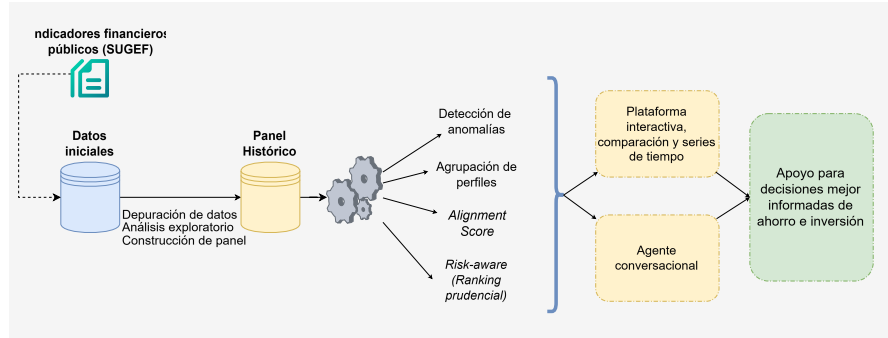


**Fig. 1** Flujo integral de WISE Investors. Los datos públicos de la SUGEF alimentan el panel histórico y la metodología WISE; sus resultados son utilizados por la plataforma interactiva y el agente conversacional para apoyar decisiones de ahorro e inversión mejor informadas.

## 2 Estado del arte

La SUGEF pone a disposición del público un amplio conjunto de indicadores financieros correspondientes a las entidades de intermediación financiera supervisadas. Esta información constituye la principal fuente pública para evaluar el comportamiento financiero de bancos, cooperativas de ahorro y crédito, mutuales y financieras en Costa Rica. Entre los indicadores publicados se incluyen métricas relacionadas con la calidad de la cartera crediticia, suficiencia patrimonial, rentabilidad, liquidez, eficiencia operativa y estructura financiera [4], las cuales permiten realizar análisis relativamente completos sobre la situación de cada entidad.

No obstante, la disponibilidad de información no implica necesariamente que esta sea accesible para la toma de decisiones. La cantidad de indicadores publicados, su evolución temporal, las diferencias entre modelos de negocio y la necesidad de interpretar simultáneamente múltiples variables convierten el análisis financiero en una tarea que requiere conocimientos técnicos y herramientas especializadas. Como consecuencia, existe una brecha importante entre la información que el supervisor pone a disposición del público y la capacidad de la mayoría de los ahorrantes e inversionistas para transformarla en conocimiento útil.

A partir de esta necesidad han surgido diversas organizaciones que desarrollan productos de análisis financiero dirigidos al sector especializado. Empresas como CEFSA, Perfil Financiero, FCS Capital, Academia de Centroamérica, la Cámara de Bancos e Instituciones Financieras de Costa Rica y Economía Hoy elaboran informes técnicos, estudios sectoriales, presentaciones ejecutivas y, en algunos casos, plataformas de visualización que facilitan el análisis del sistema financiero [5] [6] [7] [8]. De manera complementaria, agencias calificadoras internacionales como Moody's, Fitch Ratings y S&P Global Ratings generan evaluaciones independientes sobre la fortaleza financiera de determinadas entidades o emisiones [9] [10] [11], constituyendo una referencia ampliamente utilizada por inversionistas institucionales y participantes del mercado.

En términos generales, estas soluciones comparten una característica común: están orientadas principalmente a bancos, cooperativas, mutuales, reguladores, inversionistas institucionales o clientes corporativos. Sus productos suelen comercializarse mediante suscripciones, consultorías o servicios especializados y responden a necesidades de supervisión, gestión del riesgo, inteligencia de negocios o análisis de inversiones. En consecuencia, el pequeño ahorrante o inversionista individual generalmente no dispone de herramientas que integren sistemáticamente la información pública de la SUGEF para facilitar su interpretación, dependiendo en gran medida de publicaciones ocasionales realizadas por estas organizaciones o por economistas especializados en medios de comunicación y redes sociales, donde normalmente se analizan uno o algunos indicadores específicos, pero no se ofrece una evaluación integral y sistemática del comportamiento financiero de las entidades.

Desde la perspectiva académica, la literatura documenta múltiples aplicaciones de técnicas de aprendizaje automático en el ámbito financiero, entre ellas la detección de anomalías, la identificación de patrones de riesgo y la segmentación de entidades financieras. En particular, los métodos no supervisados permiten identificar estructuras latentes sin requerir etiquetas previas y han sido ampliamente utilizados para detectar observaciones atípicas y agrupar observaciones con características similares [3, 12]. En el ámbito bancario, Mercadier et al. [13] emplean técnicas de aprendizaje no supervisado sobre un conjunto multidimensional de indicadores financieros para agrupar bancos según sus perfiles de riesgo y desarrollar una herramienta de monitoreo orientada a identificar fragilidades en el sistema financiero. Estas aplicaciones evidencian la utilidad de los métodos no supervisados en contextos donde el objetivo no consiste necesariamente en predecir un evento de incumplimiento, sino en identificar patrones de comportamiento, anomalías y grupos relativamente homogéneos a partir de información multidimensional.

Otra línea relacionada corresponde a la construcción de indicadores compuestos, utilizados para sintetizar múltiples dimensiones de información en una medida que facilite la comparación y el ordenamiento de las observaciones. Nardo et al. [14] y OECD [15] desarrollan un marco metodológico para su construcción que contempla etapas como la selección de variables, el análisis multivariado, la normalización, la ponderación y la agregación. Entre las alternativas consideradas se encuentran esquemas de agregación lineal, geométrica y multicriterio, así como diferentes métodos para determinar las ponderaciones de los indicadores. Este tipo de aproximaciones permite construir índices sintéticos que condensan información multidimensional en puntuaciones comparables y rankings. Sin embargo, tanto estas metodologías como buena parte de las aplicaciones académicas de aprendizaje automático se concentran principalmente en la construcción y evaluación del modelo o indicador, sin materializarse necesariamente en herramientas dirigidas al usuario final que permitan explorar, interpretar y consultar sus resultados de forma intuitiva.

En conjunto, las soluciones disponibles evidencian que existe una amplia oferta de información financiera y de herramientas especializadas para el análisis del sistema financiero. Sin embargo, también muestran una brecha importante entre la disponibilidad de dicha información y la existencia de herramientas orientadas al pequeño ahorrante o inversionista individual que integren los distintos indicadores financieros,

faciliten su interpretación y permitan realizar comparaciones sistemáticas entre entidades. Asimismo, la revisión realizada no permitió identificar soluciones que incorporen agentes conversacionales basados en modelos de lenguaje de gran escala (LLM) para asistir a los usuarios en la exploración e interpretación de información financiera pública mediante consultas en lenguaje natural. En consecuencia, no se identificó una herramienta, a nivel nacional, dirigida específicamente al pequeño ahorrante costarricense que procese los indicadores públicos publicados por la SUGEF, integre técnicas de aprendizaje automático para caracterizar el comportamiento financiero de las entidades y permita a los usuarios explorar e interpretar los resultados mediante un agente conversacional basado en inteligencia artificial generativa. Este vacío constituye la principal motivación de la presente investigación.

### 3 Métodos

Esta sección describe los métodos utilizados para el desarrollo de WISE Investors. En primer lugar, se presentan la fuente de información y los indicadores financieros utilizados en la investigación. Posteriormente, se desarrolla el fundamento matemático de la metodología WISE y se describen las etapas que conforman su construcción, incluyendo la preparación de los datos, la identificación de observaciones atípicas y perfiles financieros, el cálculo del *Alignment Score* y la generación de los rankings. Finalmente, se presentan la arquitectura y el funcionamiento del agente conversacional, así como la estructura general de la plataforma WISE Investors.

#### 3.1 Datos

La metodología propuesta utiliza información pública proveniente de la SUGEF, organismo responsable de la supervisión prudencial del sistema financiero costarricense. La SUGEF publica periódicamente indicadores financieros correspondientes a las entidades supervisadas, los cuales constituyen la principal fuente de información utilizada en esta investigación y pueden consultarse directamente en el sitio web del supervisor [4].

A partir de esta fuente se desarrolla en Python un proceso automatizado para la adquisición, consolidación, limpieza y transformación de los datos. Para ello se construyen módulos y *pipelines* que descargan los indicadores financieros publicados por la SUGEF, integran los archivos históricos y organizan la información en una estructura homogénea para su posterior procesamiento. El código utilizado para ejecutar este proceso y reproducir la construcción del conjunto de datos se encuentra disponible en el repositorio de GitHub desarrollado como parte de esta investigación [16].

El procesamiento de los datos comprende distintas etapas de control y depuración. En primer lugar, la información descargada se consolida en series históricas de frecuencia mensual, cuya estructura temporal y consistencia se validan antes de continuar con las etapas posteriores del procesamiento. Posteriormente, se excluyen las entidades supervisadas que no corresponden a intermediarios financieros y que, por tanto, se encuentran fuera del alcance de la investigación, como las casas de cambio. Asimismo, se estandariza la identificación de las entidades para mantener su consistencia a través del tiempo. Mediante el análisis exploratorio se identifican

discontinuidades, valores faltantes e inconsistencias históricas en determinadas series. Los casos identificados se revisan individualmente y, cuando corresponde, se aplican procedimientos de interpolación y otras correcciones para preservar la consistencia de las series. Asimismo, se consideran las fechas de corte correspondientes a entidades que dejan de encontrarse activas dentro del período analizado. El detalle del análisis exploratorio, incluyendo las inconsistencias identificadas, las interpolaciones realizadas y las demás correcciones aplicadas a las series, se documenta en el material suplementario [I\\_Dashboard\\_Sugef\\_Exploratorio\\_Univariado.html](#).

A partir de las series de tiempo depuradas se construye un panel histórico de frecuencia mensual organizado por entidad y fecha, que integra la evolución de las entidades financieras supervisadas y sus principales indicadores para el período comprendido entre octubre de 2015 y abril de 2026. De esta manera, el panel utilizado por WISE no corresponde simplemente a una descarga directa de los datos publicados por la SUGEF, sino al resultado de un proceso reproducible de adquisición, consolidación, validación, depuración y transformación implementado mediante los módulos y *pipelines* desarrollados para esta investigación. La Fig. 2 sintetiza las principales etapas de este proceso.

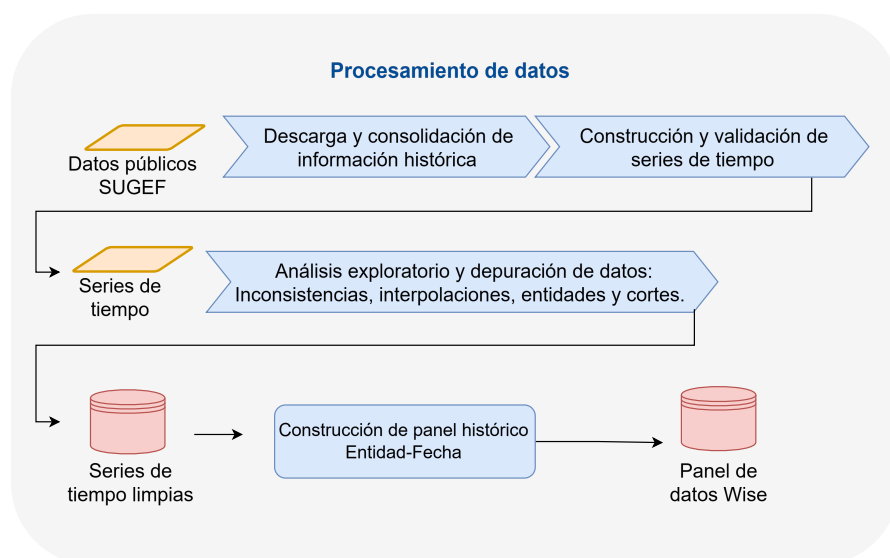


**Fig. 2** Proceso de adquisición, consolidación, limpieza y procesamiento de los datos, desde la información pública obtenida de la SUGEF hasta la construcción del panel histórico utilizado por la metodología WISE. Fuente: Elaboración propia.

La metodología WISE se basa en cuatro indicadores financieros seleccionados a partir de un conjunto inicial de 15 indicadores calculados y publicados por la SUGEF. La selección combina criterios de relevancia financiera, complementariedad, direccionalidad e interpretabilidad, apoyados por un análisis exploratorio univariado y multivariado de los indicadores disponibles. Asimismo, se procura mantener una representación parsimoniosa del espacio financiero, utilizando un número reducido

de variables que facilite tanto el análisis multivariado como la interpretación y visualización de los resultados.

La Tabla 1 presenta el conjunto inicial de indicadores considerados. Estos abarcan distintas dimensiones del comportamiento financiero de las entidades supervisadas y constituyen el punto de partida del proceso de selección. El detalle del análisis exploratorio univariado se presenta en el material suplementario `I.Dashboard_Sugef_Exploratorio_Univariado.html`, mientras que el análisis de las relaciones entre variables y del comportamiento multivariado se documenta en `II.Dashboard_AnalisisMultivariado.html`.

**Table 1** Indicadores financieros considerados en el análisis exploratorio

| Indicador                                                              | Dimensión financiera                   | Seleccionado |
|------------------------------------------------------------------------|----------------------------------------|--------------|
| Activo Productivo / Activo Total                                       | Estructura y calidad de activos        | No           |
| Activo Productivo de Intermediación Financiera / Activo Productivo     | Estructura de activos productivos      | No           |
| Inversiones / Activo Productivo de Intermediación Financiera           | Composición de activos productivos     | No           |
| Cartera al Día / Cartera Total                                         | Calidad de cartera                     | Sí           |
| Cartera al Día / Activo Productivo                                     | Calidad y composición de activos       | No           |
| Mora mayor a 90 días / Cartera Directa                                 | Riesgo crediticio                      | Sí           |
| Estimaciones / Mora mayor a 90 días                                    | Cobertura del riesgo crediticio        | No           |
| Compromiso Patrimonial                                                 | Pérdidas no cubiertas por estimaciones | No           |
| Pasivo con Costo / Pasivo Total                                        | Estructura de fondeo                   | No           |
| Captaciones a Plazo / Pasivo con Costo                                 | Estructura de fondeo                   | No           |
| Obligaciones con Entidades Financieras del País / Pasivo con Costo     | Estructura de fondeo                   | No           |
| Obligaciones con Entidades Financieras del Exterior / Pasivo con Costo | Estructura de fondeo                   | No           |
| Activo Productivo de Intermediación Financiera / Pasivo con Costo      | Intermediación financiera              | No           |
| Rentabilidad / Patrimonio                                              | Rentabilidad                           | Sí           |
| Utilidad Operativa Bruta / Gastos Administrativos                      | Eficiencia operativa                   | Sí           |

Desde el punto de vista financiero, los cuatro indicadores seleccionados representan dimensiones complementarias del desempeño de una entidad de intermediación financiera. La proporción de Cartera al Día sobre Cartera Total aproxima la calidad del principal activo productivo de estas entidades; la Mora mayor a 90 días sobre Cartera Directa refleja la materialización del riesgo crediticio y aspectos relacionados con la gestión de la cartera; la Utilidad Operativa Bruta sobre Gastos Administrativos aproxima la eficiencia operativa y la capacidad de los ingresos generados para cubrir la estructura de gastos; y la Rentabilidad sobre el Patrimonio sintetiza el resultado de la gestión en términos del rendimiento generado sobre los recursos patrimoniales.

Un criterio particularmente relevante para la formulación matemática de WISE es que los indicadores seleccionados presentan una dirección de desempeño claramente

interpretable. Una mayor proporción de cartera al día, una mayor eficiencia operativa y una mayor rentabilidad se consideran condiciones favorables, mientras que una mayor morosidad representa una condición desfavorable. Esta propiedad permite definir de manera consistente la dirección del vector ideal utilizado posteriormente en el cálculo del *Alignment Score*. El análisis exploratorio muestra que varios de los indicadores inicialmente considerados no presentan esta característica. Por ejemplo, una mayor proporción de captaciones a plazo no implica necesariamente un mejor o peor desempeño financiero, sino que puede responder a diferencias en la estructura de fondeo o en el modelo de negocio de cada entidad. En consecuencia, estos indicadores pueden resultar relevantes para caracterizar a las instituciones, pero no necesariamente son adecuados para construir una medida direccional de alineamiento como la propuesta en esta investigación.

Los cuatro indicadores resultantes del proceso de selección y su interpretación financiera se resumen en la Tabla 2.

**Table 2** Indicadores financieros de la metodología WISE

| Indicador                                             | Descripción                                                                                                   |
|-------------------------------------------------------|---------------------------------------------------------------------------------------------------------------|
| Cartera al Día sobre Cartera Total                    | Proporción de créditos vigentes y en cumplimiento frente al total de la cartera crediticia.                   |
| Mora mayor a 90 días sobre Cartera Directa            | Porcentaje de la cartera directa con mora o atrasos de más de 90 días.                                        |
| Utilidad Operativa Bruta sobre Gastos Administrativos | Relación entre la utilidad operativa generada y los gastos administrativos; aproxima la eficiencia operativa. |
| Rentabilidad sobre el Patrimonio                      | Capacidad de la institución para generar beneficios en relación con sus recursos patrimoniales.               |

### 3.2 Fundamento matemático de la metodología WISE

La metodología WISE se fundamenta en una representación vectorial del comportamiento financiero de las entidades supervisadas. Bajo esta formulación, cada observación entidad-fecha se representa mediante un vector cuyas componentes corresponden a los cuatro indicadores financieros seleccionados.

En adelante, se utiliza la convención de representar los vectores mediante símbolos en negrita. Sea

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \in \mathbb{R}^4, \quad (1)$$

donde  $\mathbf{x}$  representa el vector de indicadores financieros estandarizados asociado a una observación entidad-fecha, y  $x_j$ , con  $j = 1, \dots, 4$ , representa la componente correspondiente al  $j$ -ésimo indicador financiero. En particular,  $x_1$  corresponde a Cartera

al Día sobre Cartera Total,  $x_2$  a Mora mayor a 90 días sobre Cartera Directa,  $x_3$  a Utilidad Operativa Bruta sobre Gastos Administrativos y  $x_4$  a Rentabilidad sobre el Patrimonio.

Esta representación permite interpretar cada observación entidad-fecha como un punto dentro de un espacio vectorial de cuatro dimensiones, denominado en esta investigación *espacio financiero*. Sobre dicho espacio es posible definir medidas de similitud, calcular distancias entre observaciones y aplicar algoritmos de aprendizaje automático utilizando una estructura matemática común [17].

Con el propósito de cuantificar el grado de alineamiento entre vectores dentro del espacio financiero se emplea el producto interno [17, 18]. Sean  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^4$  dos vectores de indicadores financieros estandarizados, los cuales pueden corresponder a observaciones distintas o coincidir. El producto interno entre ambos se define como

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y} = \sum_{j=1}^4 x_j y_j, \quad (2)$$

donde  $x_j$  y  $y_j$  representan, respectivamente, la  $j$ -ésima componente de los vectores  $\mathbf{x}$  y  $\mathbf{y}$ .

Asociada al producto interno se define la norma euclidiana

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{\sum_{j=1}^4 x_j^2}, \quad (3)$$

la cual representa la longitud o magnitud del vector dentro del espacio financiero. La norma permite cuantificar el tamaño del vector considerando conjuntamente sus cuatro componentes. Por ejemplo, un vector con componentes de mayor magnitud tendrá, en términos generales, una norma mayor que otro cuyas componentes sean de menor magnitud. De esta forma, la norma proporciona una medida escalar de la magnitud global del vector que representa el comportamiento financiero de una entidad..

La relación entre dirección y magnitud puede observarse a partir de la expresión geométrica del producto interno. Para  $\mathbf{x}, \mathbf{v} \in \mathbb{R}^4$  no nulos,

$$\langle \mathbf{x}, \mathbf{v} \rangle = \|\mathbf{x}\| \|\mathbf{v}\| \cos \theta, \quad (4)$$

donde  $\theta$  representa el ángulo entre ambos vectores. Esta expresión muestra que el producto interno incorpora simultáneamente dos componentes: el grado de alineamiento entre los vectores, representado por  $\cos \theta$ , y sus magnitudes, representadas por las respectivas normas.

Esta característica resulta especialmente conveniente para la metodología WISE. Si se utilizara únicamente el ángulo entre vectores o, de manera equivalente, la similitud del coseno, dos entidades cuyos vectores presenten la misma dirección obtendrían exactamente la misma medida de similitud, independientemente de su magnitud. Por ejemplo, dos entidades pueden mantener la misma proporción entre sus indicadores, pero una de ellas presentar valores estandarizados de mayor magnitud en la dirección favorable. La similitud del coseno consideraría ambos comportamientos equivalentes en términos de alineamiento, mientras que el producto interno permite distinguir la intensidad con la que dicho comportamiento se manifiesta.

En consecuencia, el producto interno incorpora simultáneamente información sobre la dirección y la magnitud del comportamiento financiero. Esta propiedad resulta consistente con el objetivo de la metodología WISE, que busca valorar tanto el grado de alineamiento con una dirección financiera considerada favorable como la intensidad con la que dicho comportamiento se manifiesta.

Definidos el producto interno y la norma euclidiana, es posible interpretar geoméricamente el alineamiento mediante la proyección ortogonal. Sea  $\mathbf{x} \in \mathbb{R}^4$  el vector de indicadores financieros estandarizados correspondiente a una observación entidad-fecha y sea  $\mathbf{v} \in \mathbb{R}^4$ , con  $\mathbf{v} \neq \mathbf{0}$ , un vector que define una dirección de referencia dentro del espacio financiero. La proyección ortogonal de  $\mathbf{x}$  sobre  $\mathbf{v}$  se define como

$$\text{proj}_{\mathbf{v}}(\mathbf{x}) = \frac{\langle \mathbf{x}, \mathbf{v} \rangle}{\|\mathbf{v}\|^2} \mathbf{v}. \quad (5)$$

En este contexto, la proyección representa la componente del comportamiento financiero de la observación  $\mathbf{x}$  que se encuentra orientada en la dirección de referencia definida por  $\mathbf{v}$ . Esta interpretación proporciona la base geométrica para establecer posteriormente una medida escalar de alineamiento.

En la metodología WISE, la dirección de referencia corresponde al vector ideal de desempeño financiero

$$\mathbf{v}_{\text{ideal}} = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix} \in \mathbb{R}^4. \quad (6)$$

Las componentes de este vector indican el sentido en que cada indicador contribuye a un comportamiento financiero más favorable. Los valores positivos corresponden a indicadores para los cuales un mayor valor representa un mejor desempeño: cartera al día, eficiencia operativa y rentabilidad sobre el patrimonio. El valor negativo corresponde al indicador de mora, dado que un menor nivel de morosidad representa una condición financiera más favorable.

Dado que la metodología WISE se desarrolla bajo un enfoque no supervisado, no existen etiquetas que permitan identificar empíricamente una observación como referente de desempeño financiero. Por esta razón, el vector ideal se define como una dirección de referencia balanceada dentro del espacio financiero. Su construcción asigna la misma magnitud a las cuatro dimensiones, de manera que ningún indicador recibe una ponderación mayor *a priori*, mientras que el signo de cada componente establece el sentido considerado favorable para el indicador correspondiente. En consecuencia,  $\mathbf{v}_{\text{ideal}}$  no representa una entidad observada ni un perfil aprendido a partir de etiquetas, sino una referencia metodológica respecto a la cual se compara el posicionamiento de las observaciones.

Desde una perspectiva geométrica, esta construcción puede ilustrarse mediante un caso simplificado de dos dimensiones. Si ambos indicadores presentan una relación positiva con el desempeño financiero y reciben la misma ponderación, la dirección de referencia corresponde al vector  $[1, 1]^T$ , asociado con la recta de  $45^\circ$ , que representa un balance entre ambas dimensiones. La metodología WISE extiende este principio al espacio financiero de cuatro dimensiones, incorporando además el signo negativo de

la morosidad para reflejar que, a diferencia de los demás indicadores seleccionados, valores menores de esta variable representan una condición más favorable.

Dado que el interés de la metodología consiste en obtener una medida escalar que permita comparar el grado de alineamiento relativo de las entidades, y no en reconstruir el vector proyectado, el *Alignment Score* (AS) se define mediante el producto interno entre el vector estandarizado de cada observación y el vector ideal,

$$AS(\mathbf{x}) = \langle \mathbf{x}, \mathbf{v}_{\text{ideal}} \rangle. \quad (7)$$

Utilizando la relación entre el producto interno, las normas y el ángulo entre vectores, el *Alignment Score* también puede expresarse como

$$AS(\mathbf{x}) = \|\mathbf{x}\| \|\mathbf{v}_{\text{ideal}}\| \cos \theta, \quad (8)$$

donde  $\theta$  representa el ángulo entre  $\mathbf{x}$  y  $\mathbf{v}_{\text{ideal}}$ . Dado que

$$\|\mathbf{v}_{\text{ideal}}\| = \sqrt{1^2 + (-1)^2 + 1^2 + 1^2} = 2, \quad (9)$$

se obtiene

$$AS(\mathbf{x}) = 2\|\mathbf{x}\| \cos \theta. \quad (10)$$

Esta expresión permite observar directamente las dos dimensiones que intervienen en la puntuación. Para dos observaciones con igual magnitud, aquella cuyo vector presenta un menor ángulo respecto a la dirección ideal obtiene un mayor *Alignment Score*. De manera análoga, si dos vectores presentan el mismo ángulo respecto a la dirección ideal y  $\cos \theta > 0$ , aquel con mayor norma obtiene una puntuación superior. Matemáticamente, si  $\|\mathbf{x}\| = \|\mathbf{y}\|$  y  $\theta_{\mathbf{x}} < \theta_{\mathbf{y}}$ , para ángulos comprendidos entre 0 y  $\pi$ , entonces

$$\cos \theta_{\mathbf{x}} > \cos \theta_{\mathbf{y}} \implies AS(\mathbf{x}) > AS(\mathbf{y}). \quad (11)$$

Asimismo, para dos vectores con el mismo ángulo  $\theta$  respecto a la dirección ideal, siempre que  $\cos \theta > 0$ ,

$$\|\mathbf{x}\| > \|\mathbf{y}\| \implies AS(\mathbf{x}) > AS(\mathbf{y}). \quad (12)$$

De esta forma, el *Alignment Score* recompensa simultáneamente el alineamiento con la dirección considerada favorable y la magnitud con la que dicho comportamiento se manifiesta. Por ejemplo, si dos entidades presentan la misma proporción relativa entre sus indicadores y, por tanto, la misma dirección en el espacio financiero, pero una exhibe valores estandarizados de mayor magnitud en la dirección favorable, esta última obtiene un *Alignment Score* superior.

La Fig. 3 presenta una representación bidimensional ilustrativa de esta construcción geométrica. En ella se consideran únicamente dos indicadores cuya dirección favorable es positiva, con el propósito de visualizar la relación entre el vector que representa una observación, la dirección ideal y su proyección.

El *Alignment Score* constituye, por tanto, una medida escalar del posicionamiento de una observación respecto a la dirección financiera considerada favorable. Debido a que sus componentes se obtienen a partir de indicadores estandarizados, su valor

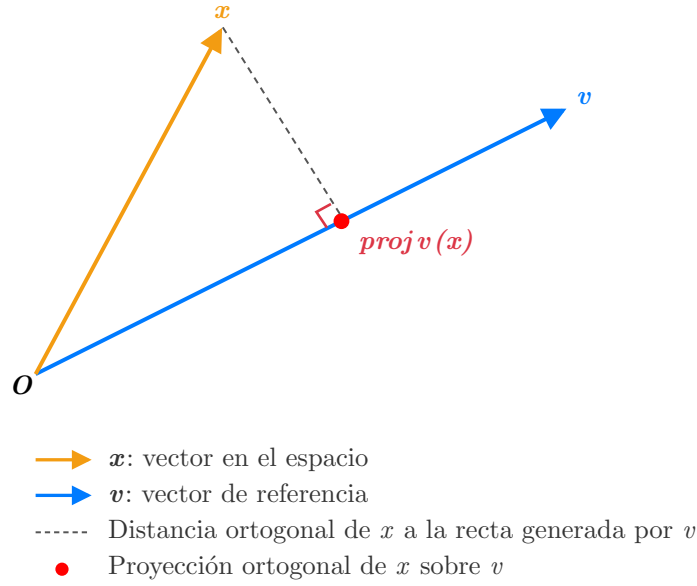


**Fig. 3** Representación geométrica bidimensional del fundamento del *Alignment Score*. El vector  $x$  representa una observación en un espacio financiero simplificado de dos indicadores y  $v$  representa la dirección asociada al perfil ideal. La proyección ortogonal de  $x$  sobre  $v$  ilustra la componente del comportamiento financiero orientada en dicha dirección.

no posee una interpretación absoluta, sino que adquiere significado al comparar observaciones sometidas al mismo proceso de estandarización.

### 3.3 Construcción de la metodología WISE

Una vez definida la representación matemática del problema, la metodología integra herramientas de aprendizaje automático y análisis multivariado para caracterizar el comportamiento financiero de las entidades supervisadas. La Fig. 4 resume las principales etapas del procedimiento metodológico:

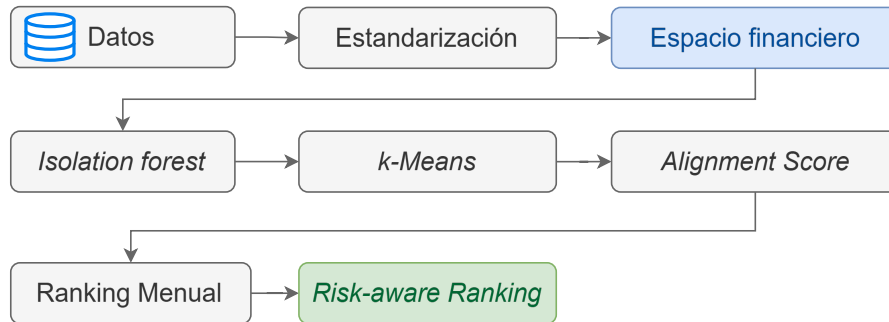


**Fig. 4** Flujo general de la metodología WISE. Muestra en orden, los procesos para generar el ranking

### 3.3.1 Preparación de los datos

Los indicadores financieros son sometidos a procesos de limpieza, tratamiento de valores faltantes y estandarización. Esta última etapa garantiza que todas las variables contribuyan de manera equilibrada tanto al cálculo del *Alignment Score* como a los algoritmos de aprendizaje automático, evitando que las diferencias de escala dominen el análisis.

### 3.3.2 Identificación de observaciones atípicas (*Isolation Forest*)

Para identificar observaciones cuyo comportamiento financiero difiere significativamente del patrón general del sistema se emplea el algoritmo *Isolation Forest* [19]. Esta técnica de aprendizaje no supervisado asigna a cada observación un puntaje de anomalía (*anomaly score*) que permite cuantificar qué tan inusual resulta su comportamiento respecto al conjunto de datos analizado, sin requerir etiquetas previas.

Para establecer el umbral de clasificación entre observaciones regulares e inusuales se utiliza un nivel de contaminación del 1%. Este parámetro indica la proporción esperada de observaciones atípicas y permite determinar el punto de corte sobre la distribución de los puntajes de anomalía. La Fig. 5 ilustra este procedimiento: la línea vertical punteada representa el umbral de clasificación y las observaciones ubicadas en la región definida como anómala corresponden aproximadamente al 1% del conjunto analizado.

La clasificación obtenida constituye una herramienta complementaria para la interpretación del comportamiento financiero de las entidades y no modifica el cálculo del *Alignment Score*. Los aspectos relacionados con la calibración del algoritmo, la selección de hiperparámetros y la evaluación de las observaciones atípicas se documentan con mayor detalle en el material suplementario [II\\_Dashboard\\_AnalisisMultivariado.html](#).

### 3.3.3 Identificación de perfiles financieros (*K-Means*)

Con el propósito de identificar grupos de observaciones con características financieras similares se emplea el algoritmo *K-Means* [20]. Este método de aprendizaje no supervisado particiona las observaciones en conglomerados a partir de su proximidad en el espacio financiero, buscando minimizar la variabilidad dentro de cada grupo. De esta forma, observaciones con combinaciones similares de los cuatro indicadores tienden a quedar agrupadas dentro de un mismo conglomerado.

La determinación del número de conglomerados combina criterios cuantitativos y económico-financieros. Como referencia cuantitativa se utiliza el método del codo, mediante el cual se analiza la reducción de la variabilidad intraclúster a medida que aumenta el número de conglomerados. Este procedimiento permite identificar valores de  $k$  a partir de los cuales la incorporación de conglomerados adicionales produce mejoras progresivamente menores en el ajuste de la segmentación, ver Fig. 6.

No obstante, la selección del número de conglomerados no se realiza exclusivamente a partir de este criterio. Dado que el objetivo de la segmentación consiste en identificar perfiles con una interpretación financiera diferenciada, se incorpora como criterio principal la coherencia económico-financiera de los grupos obtenidos. Para ello, se

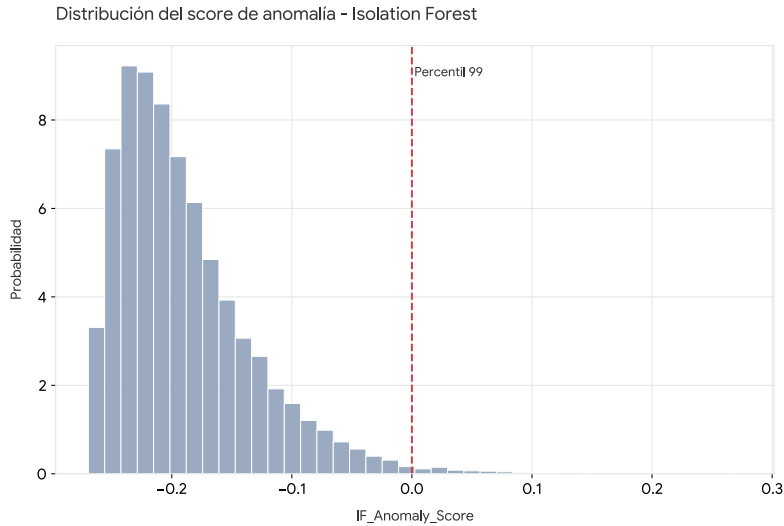


**Fig. 5** Distribución del puntaje de anomalía obtenido mediante *Isolation Forest*. La línea vertical punteada representa el umbral de clasificación determinado utilizando un nivel de contaminación del 1%, que permite distinguir entre observaciones regulares e inusuales. Fuente: Elaboración propia.

desarrolla una herramienta de análisis que permite comparar las segmentaciones generadas para diferentes valores de  $k$  y examinar el comportamiento de los indicadores dentro de cada conglomerado.

Debido a que el espacio financiero está conformado por cuatro indicadores y una representación tridimensional permite visualizar únicamente tres de ellos simultáneamente, para cada valor de  $k$  se generan las cuatro combinaciones posibles de tres indicadores. Estas visualizaciones permiten examinar la separación, el solapamiento y las características de los conglomerados desde diferentes perspectivas del espacio financiero, facilitando la evaluación de si la incorporación de nuevos grupos representa patrones financieros diferenciados o únicamente subdivisiones de perfiles previamente identificados.

Los conglomerados obtenidos mediante *K-Means* no poseen por sí mismos una interpretación económica. Por esta razón, la evaluación de las distintas segmentaciones considera también el comportamiento relativo de los cuatro indicadores dentro de cada grupo, con el propósito de determinar si los conglomerados identificados corresponden a perfiles financieros suficientemente diferenciados e interpretables. De esta manera, la selección del número de conglomerados combina la evidencia proporcionada por el método del codo con la interpretación económico-financiera de las segmentaciones obtenidas.

La configuración finalmente seleccionada y la caracterización de los perfiles financieros resultantes se presentan en la sección de Resultados. El proceso de calibración, la evaluación de diferentes valores de  $k$  y las herramientas de visualización utilizadas para analizar las distintas segmentaciones se documentan con mayor detalle en el material suplementario [II\\_Dashboard\\_AnalisisMultivariado.html](#).

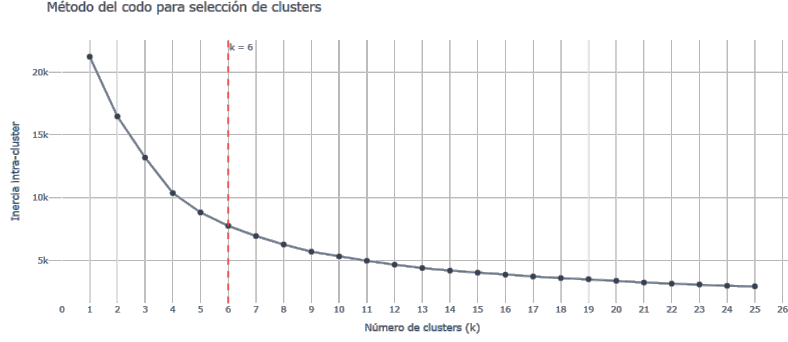


**Fig. 6** Método del codo para la selección del número de conglomerados en *K-Means*. La curva muestra la reducción de la inercia intraclúster a medida que aumenta el número de conglomerados ( $k$ ). La línea vertical punteada señala la configuración seleccionada ( $k = 6$ ), cuya elección se complementa con criterios de interpretación económico-financiera. Fuente: Elaboración propia.

### 3.3.4 Construcción del Ranking Mensual

En cada período, las entidades financieras se ordenan de forma descendente según su *Alignment Score*, de manera que la entidad con mayor puntuación ocupa la primera posición. Además del ranking general del sistema financiero supervisado, la metodología construye rankings específicos por grupo de intermediarios financieros, permitiendo realizar comparaciones entre instituciones con características regulatorias similares, como bancos, cooperativas de ahorro y crédito, mutuales o financieras.

El Ranking Mensual proporciona una fotografía del posicionamiento relativo de las entidades en una fecha determinada y permite identificar cuáles presentan un mayor grado de alineamiento con el perfil financiero ideal dentro del sistema financiero o de su respectivo grupo de referencia.

### 3.3.5 Construcción del *Risk-aware Ranking*

El Ranking Mensual puede verse influenciado por mejoras o deterioros transitorios observados en el último período, mientras que el uso de promedios históricos puede suavizar episodios relevantes del comportamiento financiero. Con el propósito de incorporar una perspectiva prudencial se desarrolló el *Risk-aware Ranking*, el cual resume el comportamiento reciente de cada entidad utilizando una ventana móvil de 36 meses. Para los indicadores donde un mayor valor representa un mejor desempeño, el valor prudencial corresponde al mínimo entre la observación más reciente y el percentil 10 de la ventana histórica. Para aquellos donde un menor valor representa un mejor desempeño, se utiliza el máximo entre la observación más reciente y el percentil 90. Los valores prudenciales obtenidos se estandarizan y posteriormente se procesan mediante el mismo cálculo del *Alignment Score* utilizado para el *Ranking Mensual*.

Esta estrategia permite incorporar información histórica sin perder sensibilidad frente al comportamiento reciente. Por ejemplo, una entidad que haya presentado un desempeño relativamente débil durante varios años podría mostrar una mejora puntual en el último período y alcanzar una posición elevada en el *Ranking Mensual*. En el *Risk-aware Ranking*, dicha mejora no modifica inmediatamente la valoración

prudencial, ya que esta continúa reflejando el comportamiento observado a lo largo de la ventana histórica. De manera análoga, si una entidad ha mantenido un desempeño sólido durante un período prolongado, pero experimenta un deterioro importante en la información más reciente, el *Risk-aware Ranking* incorpora ese deterioro de forma inmediata, evitando que quede diluido por un historial favorable.

Al igual que el *Ranking Mensual*, el *Risk-aware Ranking* se calcula tanto para el conjunto del sistema financiero como para cada grupo de entidades, permitiendo comparar instituciones que comparten modelos de negocio y marcos regulatorios similares bajo una perspectiva prudencial.

Mientras el *Ranking Mensual* ofrece una fotografía del comportamiento financiero en una fecha específica, el *Risk-aware Ranking* proporciona una evaluación más conservadora al reducir la influencia de mejoras transitorias y preservar la sensibilidad frente a deterioros recientes. Ambos rankings constituyen herramientas complementarias para el análisis comparativo del comportamiento financiero de las entidades supervisadas.

### 3.4 Agente conversacional

La arquitectura del agente se concibe como un grafo de estados diseñado para mitigar alucinaciones en la consulta de indicadores financieros. El procesamiento se organiza en cuatro fases principales, abstraídas del diagrama de arquitectura (Fig. 7):

1. **Normalización de dominio y Enrutamiento adaptativo.** En esta fase, un módulo de *Normalización de consultas* traduce términos coloquiales a una terminología financiera reconocible. Luego, el *Enrutador dinámico* decide si la consulta debe resolverse mediante SQL, recuperación semántica (RAG) o reformulación.
2. **Procesamiento Híbrido (SQL y RAG).** El flujo se bifurca en dos ramas: la *Ruta SQL*, que ejecuta consultas cuantitativas sobre indicadores financieros, y la *Ruta RAG*, que recupera evidencia documental desde la base vectorial. Ambas convergen en el módulo de *Generación de respuestas*, garantizando coherencia entre evidencia numérica y textual.
3. **Discriminación, Retroalimentación y Auto-Corrección.** Un *Discriminador de respuestas* evalúa si la salida satisface la intención de la pregunta. Si detecta inconsistencias, activa el módulo de *Reformulación*, solicitando aclaraciones al usuario y cerrando el ciclo de retroalimentación.
4. **Métricas de Evaluación.** El desempeño del agente se valida mediante métricas externas y registros experimentales almacenados en la base de *Logs*, que documenta el uso del sistema.

#### 3.4.1 Normalización de dominio y Enrutamiento adaptativo

Con el fin de facilitar la visualización, se introduce la Fig. 8. En los pasos de la parte superior se ilustra el proceso de normalización y enrutamiento como *Normaliza términos* y *Enrutador dinámico*, donde la consulta inicial es transformada en términos financieros estandarizados y posteriormente evaluada por el enrutador. Este mecanismo permite decidir de manera flexible si la pregunta debe resolverse mediante la

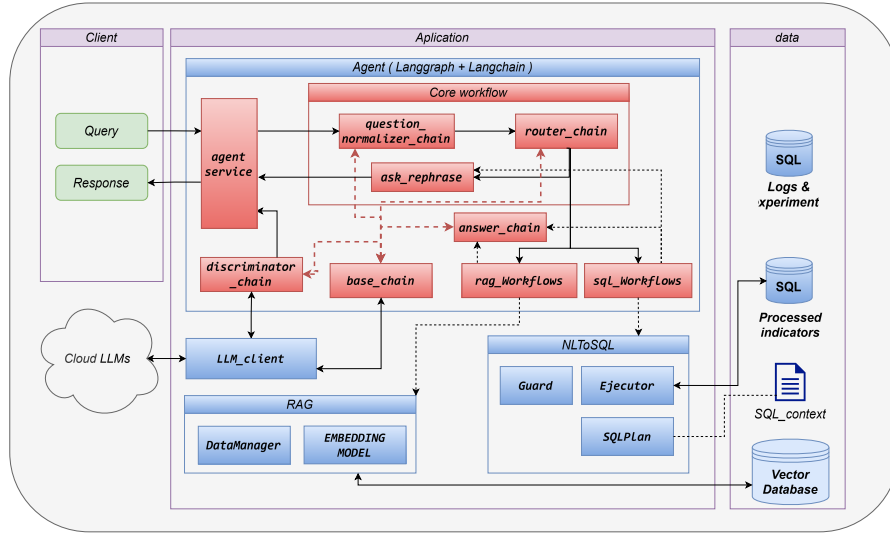


**Fig. 7** Arquitectura del agente. Contempla normalización, enrutamiento adaptativo, procesamiento híbrido SQL/RAG, discriminación y logs para obtener métricas de evaluación. Fuente: Elaboración propia.

*Ruta SQL*, orientada a cálculos cuantitativos, la *Ruta RAG*, enfocada en recuperación semántica, o bien derivarse hacia el *Módulo de Reformulación* en caso de que la inteligencia artificial detecte ambigüedad.

El enrutador aplica el principio de enrutamiento adaptativo (*Adaptive-RAG*) formulado por [21]. En lugar de forzar un único flujo de recuperación, el sistema evalúa la naturaleza de la consulta para dirigirla dinámicamente hacia:

- **Ruta SQL:** Para consultas cuantitativas y analíticas estructuradas sobre la base de datos de indicadores financieros.
- **Ruta RAG:** Para consultas cualitativas sobre regulaciones, notas aclaratorias o documentación no estructurada.
- **Módulo de Reformulación:** Cuando la consulta resulta ambigua o incompleta, solicitando al usuario una aclaración.

La respuesta final pasa por el *Discriminador de respuestas*, que evalúa si la salida generada responde a la pregunta inicial. En caso contrario, activa el flujo de nuevo hacia la normalización, pero con una sugerencia adicional.

### 3.4.2 Procesamiento Híbrido (SQL y RAG)

La separación explícita entre las ramas cuantitativa (*Ruta SQL*) y cualitativa (*Ruta RAG*) se fundamenta en el paradigma *Table-Augmented Generation* (TAG) de [22]. Este enfoque demuestra que los modelos RAG puros fallan en la agregación numérica precisa, mientras que los modelos *Text-to-SQL* carecen de contexto cualitativo; por tanto, la integración de ambas ramas garantiza precisión en el cálculo de ratios bancarios y contextualización en la generación del reporte.

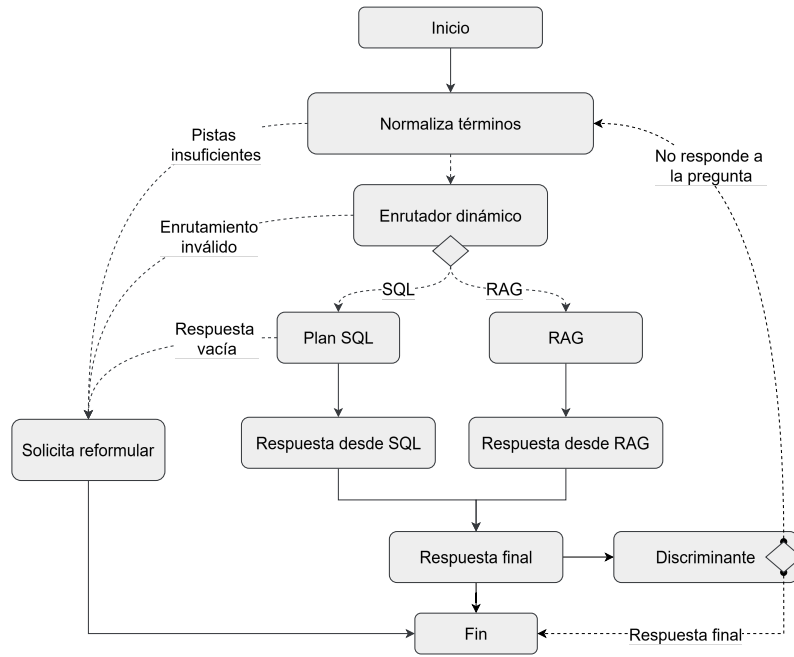


**Fig. 8** Flujo principal del agente. Desde la normalización de la consulta, el enrutamiento hacia la ruta más acertada, hasta retornar una respuesta razonable o solicitar reformulación. Fuente: Elaboración propia.

En el flujo RAG, la *Ruta RAG* activa la recuperación semántica y posterior generación de respuestas, mientras que la *Ruta SQL* ejecuta el camino estructurado hacia la salida cuantitativa. Ambas rutas convergen en el *Nodo de Respuesta final*, que representa la salida entregada al usuario.

### 3.4.3 Discriminación, Retroalimentación y Auto-Corrección

Inspirado en el marco de *Corrective RAG* (CRAG) [23], las respuestas generadas son evaluadas por el *Discriminador de respuestas* antes de su entrega final. Si la respuesta no satisface la intención de la pregunta o falla en el cumplimiento de los datos de negocio, el flujo retroalimenta al agente activando un reintento o solicitando una aclaración al usuario mediante el *Módulo de Reformulación*.

Este nodo opera como punto de control dentro del flujo, evaluando la salida de la Ruta SQL y la Ruta RAG antes de permitir la transición al nodo *Fin* o redirigir la consulta nuevamente hacia la reformulación.

### 3.4.4 Métricas de Evaluación

La evaluación del prototipo se fundamenta en métricas que permiten cuantificar la calidad de las respuestas, la fidelidad al contexto y la estabilidad operativa del agente. Estas dimensiones son importantes porque un agente conversacional no solo

debe generar respuestas correctas, sino también evitar alucinaciones, seleccionar adecuadamente las herramientas y completar las tareas de manera consistente.

Para medir la calidad de las respuestas *RAG*, *SQL* y el nivel de alucinaciones, se emplea una versión adaptada del marco *RAGAS* [24], utilizando las métricas de *Faithfulness* (grado en que la respuesta se mantiene fiel al contexto recuperado), y *Answer Relevance* (pertinencia semántica respecto a la consulta). Estas métricas son ampliamente reconocidas en la literatura para evaluar sistemas de generación aumentada por recuperación. Una limitación de este marco es que utiliza una IA para la evaluación y puede heredar un sesgo del juez [24], por lo que se opta por contar las consultas resueltas y evaluar la fidelidad de dichos resultados manualmente.

Adicionalmente, la precisión en la selección de caminos dentro del grafo de estados (*Tool/Route Selection Accuracy*) y la tasa de completitud de tareas (*Task Completion Rate*) se cuantifican siguiendo los estándares de *AgentBench* [25], un benchmark diseñado para evaluar la capacidad de los LLMs como agentes en entornos interactivos. Ampliando este marco, se incorporan métricas de latencia de inferencia y eficiencia en el uso de tokens basadas en los propuestos por [26]. Estas métricas permiten verificar que el agente no solo navega correctamente entre rutas SQL, RAG y reformulación manteniendo estabilidad en la ejecución de consultas, sino que lo hace optimizando el coste operativo y la velocidad de respuesta del sistema.

## 4 Resultados

Esta sección presenta los principales resultados obtenidos a partir del desarrollo y aplicación de la metodología WISE y de su implementación en la plataforma WISE Investors. En primer lugar, se presentan los resultados analíticos derivados de la metodología, incluyendo la identificación de observaciones atípicas, la construcción de una taxonomía de perfiles financieros y los rankings utilizados para comparar el comportamiento relativo de las entidades supervisadas. Posteriormente, se presentan los resultados asociados con la implementación de la plataforma, el agente conversacional basado en modelos de lenguaje de gran escala y su evaluación experimental.

### 4.1 Metodología WISE

El principal resultado analítico de esta investigación es la Metodología WISE, un marco diseñado para transformar información financiera pública en conocimiento útil para la evaluación comparativa del riesgo de crédito asumido por ahorrantes e inversionistas. La propuesta combina una representación vectorial del comportamiento financiero de las entidades supervisadas por la SUGEF con técnicas de aprendizaje automático y análisis multivariado, permitiendo caracterizar la posición relativa de cada institución dentro del sistema financiero supervisado mediante una formulación matemática unificada.

A diferencia de un enfoque basado en la interpretación aislada de indicadores financieros, la metodología representa cada entidad como un vector en un espacio financiero definido por cuatro indicadores estandarizados. Sobre esta representación se construyen medidas de similitud, se identifican patrones de comportamiento, se detectan observaciones atípicas y se genera un indicador sintético de alineamiento

respecto a un perfil financiero ideal. Este enfoque proporciona una interpretación geométrica del comportamiento financiero de las entidades y constituye el fundamento para la construcción de los rankings propuestos.

Bajo esta perspectiva, el riesgo de crédito se interpreta como una medida relativa del comportamiento financiero observado y no como una probabilidad de incumplimiento ni como una predicción de eventos futuros. Esta decisión metodológica responde tanto a la naturaleza del problema de investigación como a la información disponible. Mientras que los modelos supervisados requieren un número suficiente de eventos históricos de incumplimiento para estimar probabilidades de riesgo, en el caso de las entidades de intermediación financiera estos eventos son excepcionalmente poco frecuentes, lo que limita el desarrollo de modelos predictivos supervisados. En consecuencia, la metodología WISE adopta un enfoque no supervisado orientado a caracterizar y comparar el comportamiento financiero relativo de las entidades, en lugar de estimar probabilidades de insolvencia o quiebra.

Como resultado de este procedimiento se generan cuatro productos analíticos principales: la identificación de observaciones atípicas, una taxonomía de perfiles financieros, un “Ranking Mensual” construido a partir del *Alignment Score* y un *Risk-aware Ranking* que incorpora la evolución histórica reciente bajo un enfoque prudencial. En conjunto, estos productos transforman información financiera pública en instrumentos comparativos e interpretables para apoyar la toma de decisiones de ahorro e inversión.

La aplicación de *Isolation Forest* permite identificar observaciones cuyo comportamiento financiero resulta inusual respecto al patrón general observado en el sistema. La Fig. 9 presenta una representación tridimensional de esta clasificación, distinguiendo entre observaciones regulares e inusuales. La visualización utiliza tres de los cuatro indicadores que conforman el espacio financiero con el propósito de representar gráficamente la distribución de las observaciones; la clasificación mediante *Isolation Forest*, sin embargo, se realiza considerando conjuntamente los cuatro indicadores definidos por la metodología.

Este resultado complementa el análisis comparativo de las entidades al permitir identificar comportamientos financieros poco frecuentes que podrían quedar ocultos al examinar individualmente los indicadores. La condición de observación inusual no implica por sí misma un comportamiento financiero favorable o desfavorable, sino que señala una combinación de indicadores que se aparta del patrón predominante en el conjunto analizado.

Por otra parte, la aplicación del algoritmo *K-Means* permite identificar una taxonomía de seis perfiles financieros, presentada en la Tabla 3. Los conglomerados obtenidos representan patrones relativamente homogéneos de comportamiento dentro del sistema financiero supervisado y se interpretan desde una perspectiva económico-financiera a partir de las características predominantes de los cuatro indicadores que conforman el espacio financiero. Estos perfiles proporcionan un contexto adicional para interpretar el posicionamiento relativo de las entidades.

La Fig. 10 complementa la taxonomía anterior mediante una representación tridimensional de los conglomerados identificados. Los colores permiten distinguir los seis perfiles financieros y observar cómo las entidades con características similares tienden

Isolation Forest: CarteraAldía\_CT -- Mora90d\_CarteraDirecta -- UOB\_GastosAdmin

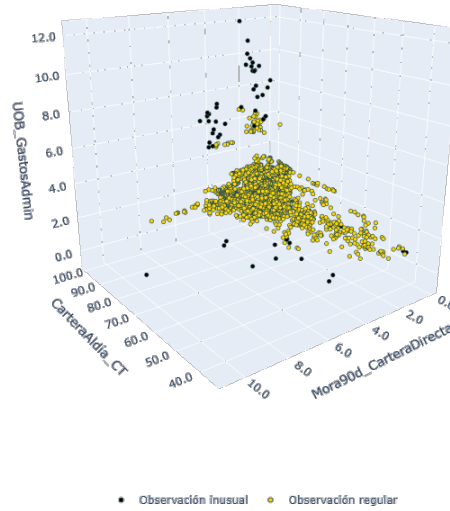


**Fig. 9** Representación tridimensional de las observaciones históricas clasificadas mediante *Isolation Forest*. La figura distingue entre observaciones regulares e inusuales identificadas utilizando un nivel de contaminación del 1% y representa tres de las cuatro dimensiones del espacio financiero. Fuente: Elaboración propia.

a concentrarse en determinadas regiones del espacio financiero. Al igual que en la visualización de *Isolation Forest*, se representan tres de los cuatro indicadores con fines gráficos, mientras que la segmentación mediante *K-Means* considera conjuntamente las cuatro dimensiones utilizadas por la metodología.

La taxonomía y la identificación de observaciones atípicas proporcionan dos perspectivas complementarias sobre el comportamiento financiero de las entidades. Mientras *Isolation Forest* permite detectar observaciones que se apartan del patrón general del sistema, *K-Means* identifica estructuras recurrentes y las organiza en perfiles financieros interpretables. Ambos resultados complementan el *Alignment Score* y los rankings derivados de este, al aportar información adicional sobre las características del posicionamiento relativo de cada entidad.

## 4.2 Risk-Aware Ranking

La Fig. 11 presenta el ordenamiento de los bancos con corte a abril de 2026 según el *Risk-Aware Ranking*. Para este período, BAC San José ocupa la primera posición, seguido por el Banco Nacional y Banco General. El ordenamiento se construye a partir de los cuatro indicadores ajustados bajo el enfoque *risk-aware*, los cuales incorporan

**Table 3** Taxonomía de perfiles financieros identificada mediante el algoritmo *K-Means*.

| Perfil financiero            | Caracterización                                                                                                            |
|------------------------------|----------------------------------------------------------------------------------------------------------------------------|
| Rentabilidad agresiva        | Rentabilidad y eficiencia operativa excepcionalmente altas, acompañadas de una exposición crediticia superior al promedio. |
| Alto desempeño integral      | Baja mora, elevada cartera al día, alta eficiencia operativa y rentabilidad superior al promedio.                          |
| Desempeño promedio saludable | Indicadores estables y saludables, con eficiencia y rentabilidad inferiores al perfil de alto desempeño integral.          |
| Presión crediticia           | Deterioro relativo de la calidad de cartera y niveles elevados de mora.                                                    |
| Baja cartera al día          | Menor proporción de cartera al día, frecuente en algunas financieras y cooperativas con modelos de negocio diferenciados.  |
| Rentabilidad crítica         | Rentabilidad persistentemente negativa y baja eficiencia operativa.                                                        |

el comportamiento de las entidades durante una ventana histórica de 36 meses y permiten evaluar su posicionamiento desde una perspectiva prudencial.

BAC San José presenta los resultados más favorables del grupo en las cuatro dimensiones consideradas: una Cartera al Día sobre Cartera Total ajustada de 92,51%, una Mora mayor a 90 días sobre Cartera Directa de 1,46%, una relación de Utilidad Operativa Bruta sobre Gastos Administrativos de 1,60 y una Rentabilidad sobre el Patrimonio de 8,44%. En comparación con los demás bancos, BAC San José combina la mayor proporción ajustada de cartera al día entre las primeras posiciones, el menor nivel de morosidad, la mayor eficiencia operativa y la mayor rentabilidad. Este comportamiento conjunto es consistente con su ubicación en la primera posición del *Risk-Aware Ranking*.

El Banco Nacional, ubicado en la segunda posición, presenta una Cartera al Día ajustada de 90,22%, una Mora mayor a 90 días de 2,79%, una relación de Utilidad Operativa Bruta sobre Gastos Administrativos de 1,45 y una Rentabilidad sobre el Patrimonio de 4,78%. Banco General ocupa la tercera posición con una Cartera al Día de 86,45%, una mora de 1,58%, una relación de eficiencia operativa de 1,25 y una rentabilidad de 2,56%. Estos resultados permiten observar el carácter multivariado del ordenamiento: mientras el Banco Nacional presenta una mayor cartera al día, eficiencia y rentabilidad, Banco General exhibe un nivel de morosidad considerablemente menor.

Las posiciones intermedias muestran con mayor claridad la utilidad de sintetizar conjuntamente las distintas dimensiones financieras. El BPDC, por ejemplo, presenta la mayor Cartera al Día ajustada entre los bancos analizados, con 94,02%, pero registra una mora de 2,60%, una relación de eficiencia operativa de 1,21 y una rentabilidad de 1,70%, ubicándose en la cuarta posición. Por su parte, Promerica ocupa la sexta posición a pesar de presentar una Cartera al Día de 90,56%, una mora relativamente baja de 1,83% y una relación de eficiencia operativa de 1,32, debido a que estos resultados coexisten con una Rentabilidad sobre el Patrimonio de  $-3,37\%$ . Estos casos muestran que una posición favorable en un indicador particular no determina por sí sola el ordenamiento de una entidad.

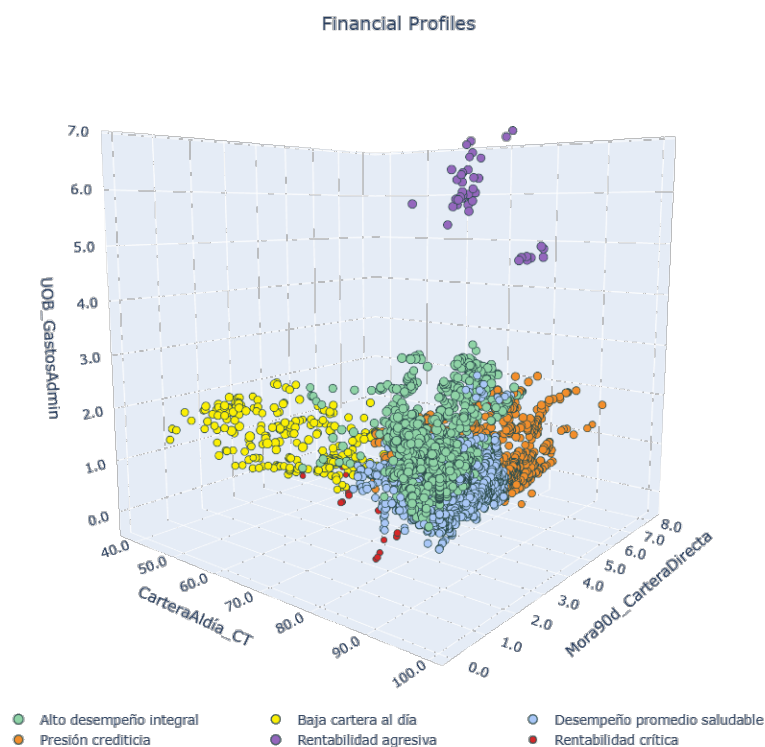


**Fig. 10** Representación tridimensional de los perfiles financieros identificados mediante *K-Means*. Los colores distinguen los seis perfiles resultantes de la segmentación y permiten visualizar su distribución en tres de las cuatro dimensiones del espacio financiero. Fuente: Elaboración propia.

En el extremo inferior, Improsa ocupa la décima posición y presenta una Cartera al Día ajustada de 73,13%, una Mora mayor a 90 días de 2,20%, una relación de Utilidad Operativa Bruta sobre Gastos Administrativos de 0,83 y una Rentabilidad sobre el Patrimonio de -2,91%. La combinación de una menor calidad de cartera, una eficiencia operativa inferior a la unidad y una rentabilidad negativa resulta consistente con su posición relativa en el extremo inferior del *ranking*. Cathay, ubicado en la novena posición, también presenta una proporción reducida de Cartera al Día, de 78,49%, aunque mantiene una rentabilidad positiva de 3,10%, lo que nuevamente evidencia que las posiciones resultan de la interacción entre las diferentes dimensiones y no de un único indicador.

En conjunto, los resultados muestran que el *Risk-Aware Ranking* permite sintetizar configuraciones financieras que resultan relativamente claras en los extremos y más heterogéneas en las posiciones intermedias. En entidades como BAC San José e Improsa, las fortalezas o debilidades relativas son observables simultáneamente en varias de las dimensiones consideradas, facilitando la interpretación de sus posiciones. En la zona intermedia, en cambio, las entidades presentan combinaciones de fortalezas

y debilidades que dificultan establecer un orden mediante la inspección individual de los indicadores. En estos casos, la metodología proporciona una regla consistente para integrar las cuatro dimensiones y establecer un posicionamiento relativo.

Asimismo, la comparación evidencia el aporte específico del componente *risk-aware*: el ordenamiento no depende exclusivamente de las condiciones financieras observadas en abril de 2026, sino de indicadores que incorporan el comportamiento de cada entidad durante una ventana histórica de 36 meses. De esta forma, la metodología introduce una perspectiva prudencial que penaliza comportamientos históricos menos favorables y evita que una observación mensual particularmente positiva determine por sí sola la posición de una entidad.

| #  | Entity     | Risk_CarteraAldia_CT | Risk_Mora90d_CarteraDirecta | Risk_UOB_GastosAdmin | Risk_Rentabilidad_Patrimonio |
|----|------------|----------------------|-----------------------------|----------------------|------------------------------|
| 1  | BAC_SJ     | 92.51                | 1.46                        | 1.6                  | 8.44                         |
| 2  | BN         | 90.22                | 2.79                        | 1.45                 | 4.78                         |
| 3  | General    | 86.45                | 1.58                        | 1.25                 | 2.56                         |
| 4  | BPDC       | 94.02                | 2.6                         | 1.21                 | 1.7                          |
| 5  | BCR        | 88.9                 | 2.97                        | 1.18                 | 4                            |
| 6  | Promerica  | 90.56                | 1.83                        | 1.32                 | -3.37                        |
| 7  | DAVIVIENDA | 89.44                | 2.8                         | 1.26                 | -0.09                        |
| 8  | BCT        | 91.11                | 2.39                        | 0.99                 | -2.56                        |
| 9  | CATHAY     | 78.49                | 2.75                        | 1.24                 | 3.1                          |
| 10 | Improsa    | 73.13                | 2.2                         | 0.83                 | -2.91                        |

**Fig. 11** *Risk-Aware Ranking* de los bancos con corte a abril de 2026. El ordenamiento compara el posicionamiento financiero relativo de las entidades considerando el comportamiento conjunto de los indicadores de calidad de cartera, morosidad, eficiencia operativa y rentabilidad durante una ventana histórica de 36 meses. Fuente: Elaboración propia.

### 4.3 Plataforma WISE Investors

Como resultado de la investigación se desarrolla el prototipo funcional de WISE Investors, una plataforma interactiva que materializa la metodología propuesta y permite explorar, analizar e interpretar la información financiera de las entidades supervisadas por la SUGEF. La plataforma integra los principales productos analíticos desarrollados en un entorno gráfico orientado a facilitar su consulta por parte de usuarios con distintos niveles de conocimiento financiero y técnico. La Fig. 12 muestra la interfaz principal de la aplicación.

WISE Investors centraliza en una misma herramienta la información correspondiente a bancos, cooperativas, empresas financieras, mutuales y otras entidades de intermediación financiera incluidas dentro del alcance de la investigación. A partir del último corte de información disponible, el usuario puede consultar el posicionamiento relativo de las instituciones, analizar la evolución histórica de sus indicadores y explorar los resultados generados por la metodología WISE.

Las principales funcionalidades implementadas en el prototipo son las siguientes:

- *Overview*: presenta una visión general del sistema financiero analizado y organiza las entidades según su tipo de intermediario, permitiendo al usuario acceder a la información disponible para cada grupo.
- *Rankings*: permite consultar y comparar el posicionamiento de las entidades a partir del *Alignment Score*, sintetizando el comportamiento conjunto de los cuatro indicadores financieros considerados por la metodología WISE.
- *Risk-aware*: presenta el *Risk-Aware Ranking*, que complementa el posicionamiento observado en el corte más reciente con el comportamiento histórico de las entidades durante una ventana móvil de 36 meses, incorporando una perspectiva prudencial al ordenamiento.
- *Time Series*: permite explorar la evolución histórica de los indicadores financieros y comparar su comportamiento entre entidades, proporcionando al usuario una perspectiva temporal que complementa los resultados sintetizados en los rankings.
- *WiChat*: incorpora una interfaz conversacional basada en inteligencia artificial generativa mediante la cual el usuario puede realizar consultas en lenguaje natural sobre las entidades financieras y la información disponible en WISE, reduciendo la necesidad de conocer previamente la estructura de los datos o los mecanismos de consulta de la plataforma.

De esta forma, el prototipo no se limita a presentar un ordenamiento estático de las entidades, sino que integra diferentes niveles de análisis dentro de una misma aplicación. El usuario puede partir de una visión general del sistema, consultar el posicionamiento relativo de una institución, examinar los indicadores que sustentan dicho resultado, analizar su comportamiento histórico y complementar la exploración mediante consultas en lenguaje natural. Esta integración constituye uno de los principales resultados tecnológicos de la investigación, al trasladar la metodología WISE desde su formulación analítica hacia una herramienta funcional orientada al pequeño ahorrrante e inversionista.

#### 4.4 Agente conversacional basado en LLM

El agente conversacional constituye la capa de interacción en lenguaje natural de WISE Investors. Su propósito es permitir que los usuarios consulten e interpreten la información producida por la metodología mediante preguntas formuladas en lenguaje cotidiano, reduciendo las barreras técnicas asociadas al análisis financiero. Como se muestra en la Fig. 7 y descrito en la sección anterior, el agente integra rutas SQL y RAG dentro de un grafo de estados que garantiza trazabilidad y control de calidad en cada respuesta.

A diferencia de un chatbot genérico, sus respuestas se fundamentan en información generada por el sistema y en herramientas especializadas para resolver consultas descriptivas, comparativas y metodológicas. La evaluación experimental con 50 preguntas de control evidenció que el agente alcanza 100% de fidelidad, con una similaridad semántica máxima de 28.1% en el modelo *OpenAI gpt-5.4-mini* y una cobertura de hasta 86% en Anthropic Claude Haiku. En términos de eficiencia, gpt-5.4-mini respondió en 7.3 segundos promedio con un consumo de 4955 tokens por consulta, lo que lo convierte en la configuración más recomendable para un

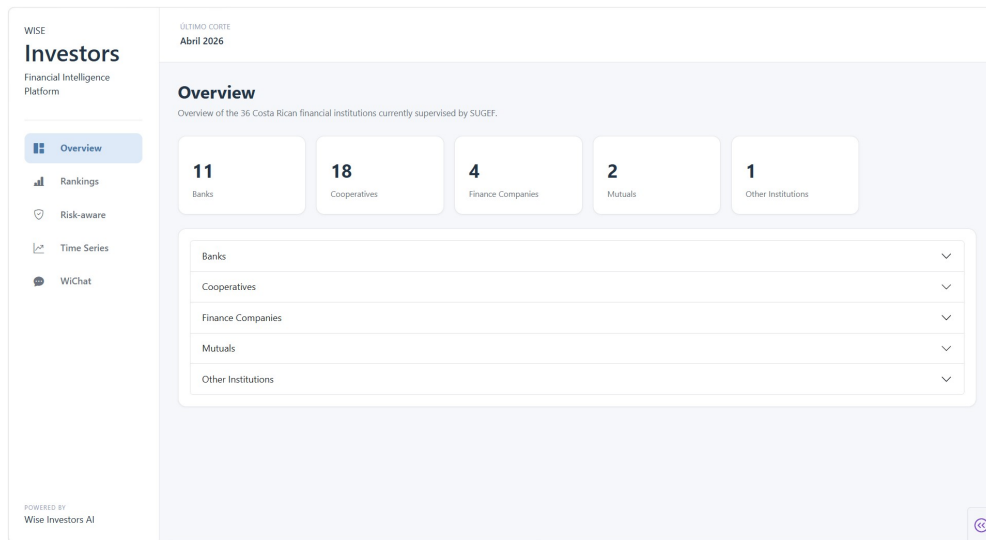


**Fig. 12** Interfaz gráfica de WISE Investors. La plataforma integra en una misma aplicación la consulta de las entidades financieras supervisadas por la SUGEF, los *rankings* desarrollados por la metodología WISE, el análisis de series de tiempo y el acceso al agente conversacional *WiChat*. Fuente: Elaboración propia.

despliegue práctico por su balance entre calidad, costo y escalabilidad. Estos resultados confirman que el agente conversacional es un complemento valioso para ampliar la accesibilidad de WISE, ofreciendo respuestas interpretables y consistentes que fortalecen el proceso de toma de decisiones financieras.

## 4.5 Evaluación experimental del agente conversacional

Para validar la robustez, precisión y viabilidad operativa del agente conversacional propuesto, se diseñó un entorno experimental comparativo que contrasta el comportamiento con dos proveedores de Modelos de Lenguaje Grande (LLMs), evaluando en cada uno de ellos una variante de alta capacidad (modelo “grande”) y una optimizada en escala (modelo “mini”). Específicamente, el banco de pruebas incluye a OpenAI (*gpt-5.4* y *gpt-5.4-mini*) y Anthropic (*Claude 4.5 Sonnet* y *Claude 4.5 Haiku*).

El pipeline de evaluación se ejecuta de forma iterativa sobre un conjunto de datos compuesto por 50 preguntas de control del dominio financiero, resultando en un total de 6 ejecuciones independientes de la matriz de pruebas. La evaluación se segmenta tridimensionalmente para cubrir la calidad semántica, la lógica de control del agente y el impacto en la infraestructura.

### 4.5.1 Calidad de Respuesta y Fidelidad

Este conjunto se evalúa porque la utilidad del agente conversacional depende directamente de la calidad de las respuestas que entrega al usuario, y no solo una salida cualquiera del sistema: es necesario verificar que dicha respuesta sea semánticamente

coherente con la pregunta y que los datos numéricos provengan estrictamente del contexto financiero recuperado. De esta forma, se asegura que el agente no introduzca alucinaciones ni errores que puedan afectar la interpretación de los indicadores. Combina técnicas automatizadas y auditoría manual para evaluar la calidad semántica y la fidelidad de las respuestas. La *Similaridad Semántica* se calcula, a partir de cada respuesta generada con una referencia base vectorizada, con valores entre 0 y 1, si una pregunta y respuesta son exactamente iguales, significas una similaridad del 100

La *Fidelidad* se obtiene mediante revisión manual, verificando si los datos numéricos provienen estrictamente del contexto recuperado; por cada pregunta, se asigna un valor de 1 si los datos son fieles a las fuentes y 0 en caso contrario, esta revisión fue manual para garantizar la métrica para las pruebas realizadas. Adicionalmente, se introduce la métrica de Completitud, que refleja el porcentaje de preguntas respondidas de manera válida (es decir, aquellas que no derivaron en salidas vacías ni en solicitudes de reformulación). Esta métrica permite cuantificar no solo la calidad de las respuestas entregadas, sino también la capacidad del agente para cubrir efectivamente el conjunto de consultas planteadas.

La Tabla 4 muestra los resultados de este eje de evaluación. Se observa que el modelo OpenAI *gpt-5.4-mini* alcanza la mayor similaridad de coseno (28.1%), superando incluso a su versión de mayor capacidad (24.9%). Todos los modelos presentan un 100% de fidelidad promedio, lo que implica que, independientemente de la variante, los datos numéricos generados provienen estrictamente de las fuentes recuperadas. En cuanto a la Completitud, los modelos de Anthropic destacan con valores superiores (86% en *Haiku* y 82% en *Sonnet*), mientras que los modelos de OpenAI se mantienen en 58%. Esto sugiere que, aunque OpenAI garantiza fidelidad perfecta, su cobertura efectiva de preguntas es menor, mientras que Anthropic logra responder un mayor número de consultas de forma válida.

**Table 4** Calidad de respuesta: Similaridad, Fidelidad y Completitud

| Modelo / Proveedor          | Similaridad Coseno (Prom.) | Fidelidad promedio | Completitud |
|-----------------------------|----------------------------|--------------------|-------------|
| OpenAI gpt-5.4-mini         | 28.1%                      | 100%               | 58%         |
| OpenAI gpt-5.4              | 24.9%                      | 100%               | 58%         |
| Anthropic claude-haiku-4-5  | 25.4%                      | 100%               | 86%         |
| Anthropic claude-sonnet-4-5 | 23.1%                      | 100%               | 82%         |

Fuente: Elaboración propia con datos del experimento.

<sup>1</sup>La tabla presenta los resultados del eje de calidad de respuesta, evaluando la Similaridad Semántica mediante distancia de coseno, la Fidelidad mediante auditoría manual y la Completitud como el porcentaje de preguntas respondidas correctamente sin derivar en reformulación.

#### 4.5.2 Orquestación, Trayectoria y Uso de Herramientas

Este eje se evalúa porque afecta directamente la rapidez y el costo computacional con que se generan las respuestas, por lo que este eje evalúa la capacidad del agente

para dirigir correctamente cada consulta dentro del grafo de estados y seleccionar la herramienta adecuada según el tipo de pregunta. Dado que el experimento consta de 50 consultas por modelo, se cuantifica el porcentaje de direccionamiento correcto promedio, así como la presencia de reintentos y fallos durante la ejecución. Esta métrica refleja la estabilidad del agente, su habilidad para evitar rutas incorrectas y su consistencia en la interacción con los nodos de SQL, RAG y reformulación.

La Tabla 5 muestra que los modelos de OpenAI alcanzan los mejores resultados en estabilidad, con trayectorias superiores al 90% y tasas de fallo reducidas (8–10%). Esto implica que, aunque requieren algunos reintentos, logran mantener un flujo consistente dentro del grafo. En contraste, los modelos de *Anthropic* presentan mayor variabilidad: *Claude Haiku* alcanza solo 66% de trayectorias correctas y un 34% de fallos, mientras que *Claude Sonnet* mejora a 74% pero aún con tasas de error superiores a las de *OpenAI*. Estos resultados sugieren que, a nivel de complejidad operativa, OpenAI ofrece mayor confiabilidad en la navegación del agente, mientras que Anthropic, aunque más completo en cobertura de preguntas, requiere mayor tolerancia a errores y reintentos.

**Table 5** Orquestación del agente: direccionamiento y estabilidad (50 preguntas por modelo)

| Modelo                      | Trayectoria (%) | Reintentos (Prom.) | Fallos (Prom.) |
|-----------------------------|-----------------|--------------------|----------------|
| gpt-5.4-mini                | 92%             | 16%                | 8%             |
| gpt-5.4                     | 90%             | 18%                | 10%            |
| Anthropic claude-haiku-4-5  | 66%             | 30%                | 34%            |
| Anthropic claude-sonnet-4-5 | 74%             | 18%                | 26%            |

Fuente: Elaboración propia con datos del experimento.

<sup>1</sup>Resume el desempeño del agente en términos de estabilidad operacional, mostrando el porcentaje de direccionamiento correcto dentro del grafo (RAG o SQL), junto con el promedio de reintentos y fallos registrados durante las 50 consultas evaluadas por modelo.

### 4.5.3 Eficiencia Operativa

La eficiencia operativa se evalúa mediante el tiempo promedio por consulta y el consumo promedio de tokens, indicadores que reflejan el costo computacional asociado al uso del agente.

La Tabla 6 evidencia que los modelos “mini” (*OpenAI gpt-5.4-mini* y *Anthropic Haiku*) son los más eficientes, con tiempos de respuesta más bajos (7.3–9.0 segundos) y menor consumo de tokens (alrededor de 4950–5030 por consulta). En cambio, las variantes “grandes” (*OpenAI gpt-5.4* y *Anthropic Sonnet*) muestran tiempos significativamente más altos (10.6–17.9 segundos) y mayor consumo de tokens (5200–5700). Esto implica que, aunque los modelos grandes pueden ofrecer mayor robustez, su uso conlleva un costo computacional notablemente más elevado. En términos prácticos, los modelos mini (y en este caso *gpt-5.4-mini*) resultan más adecuados para aplicaciones que priorizan eficiencia y escalabilidad, mientras que los modelos grandes se justifican en contextos donde la precisión y la completitud de las respuestas son críticas, aun a costa de mayor consumo de recursos.

**Table 6** Eficiencia operativa: tiempo y tokens

| Modelo                      | Tiempo prom. por Pregunta (ms) | Tokens Prom. por pregunta |
|-----------------------------|--------------------------------|---------------------------|
| gpt-5.4-mini                | 7385.36                        | 4954.72                   |
| gpt-5.4                     | 10628.50                       | 5207.08                   |
| Anthropic claude-haiku-4-5  | 9082,75                        | 5032.84                   |
| Anthropic claude-sonnet-4-5 | 17985,48                       | 5719.44                   |

Fuente: Elaboración propia con datos del experimento.

<sup>1</sup>Esta presenta los resultados del eje de eficiencia operativa, comparando el tiempo promedio requerido por cada modelo para responder una consulta y el consumo promedio de tokens promedio por cada pregunta.

## 5 Discusión

Los resultados muestran que WISE ofrece una vía para transformar información financiera pública en instrumentos objetivos de apoyo a la decisión. Su principal aporte no consiste en sustituir los análisis prudenciales tradicionales ni en predecir eventos de incumplimiento, sino en reducir la brecha existente entre la disponibilidad de información financiera y la capacidad de ahorrantes e inversionistas para integrarla e interpretarla de manera sistemática. En este sentido, la metodología permite responder, de forma objetiva y fundamentada, preguntas de interés práctico como: ¿qué entidad financiera presenta el mejor comportamiento financiero reciente dentro del sistema supervisado?

La combinación de representación vectorial, detección de anomalías, segmentación no supervisada y proyección sobre un perfil ideal permite sintetizar información multidimensional sin reducir el análisis a un único indicador aislado. La taxonomía de perfiles aporta contexto interpretativo, mientras que el *Alignment Score* proporciona una base común para comparar la posición relativa de las entidades.

El *Risk-aware Ranking* constituye un aporte especialmente relevante porque incorpora información histórica mediante un criterio de gestión de riesgo, no mediante un promedio. Esta regla reduce la posibilidad de que una mejora puntual produzca una valoración excesivamente favorable y, al mismo tiempo, evita suavizar deterioros recientes. En consecuencia, el *Ranking Mensual* y el *Risk-aware Ranking* responden a preguntas distintas y deben interpretarse como herramientas complementarias.

La capa conversacional amplía la accesibilidad de la propuesta, pero no sustituye el motor analítico. El agente debe actuar como interfaz para consultar y explicar resultados producidos por WISE, con trazabilidad hacia los datos y procedimientos que sustentan cada respuesta. Su valor depende de la exactitud, fundamentación, selección de herramientas, completitud de tareas, latencia y costo observados en la evaluación experimental. En este sentido, los resultados muestran que *OpenAI gpt-5.4-mini* ofrece el mejor balance: alcanza la mayor similaridad semántica (28.1%), fidelidad perfecta (100%) y tiempos de respuesta más bajos (7.3 segundos en promedio con 4955 tokens por consulta). Aunque su completitud es moderada (58%), resulta más eficiente y económico que las variantes grandes, que incrementan el costo computacional (hasta 17.9 segundos y 5719 tokens en *Claude Sonnet*) sin mejorar la

fidelidad. Por su parte, los modelos de Anthropic destacan en cobertura (86% en *Haiku* y 82% en *Sonnet*), pero con menor estabilidad en la orquestación y mayor latencia. En consecuencia, se recomienda el uso de *OpenAI gpt-5.4-mini* como modelo base para aplicaciones prácticas, reservando las variantes grandes únicamente para escenarios donde la completitud y robustez sean críticas aun a costa de mayor consumo de recursos.

Entre las fortalezas del trabajo destacan el uso de datos públicos, la interpretabilidad de los indicadores, la modularidad del sistema y la integración de componentes analíticos, visuales y conversacionales. La separación entre metodología y plataforma favorece la reutilización del motor WISE y permite que la interfaz evolucione sin modificar la lógica central del análisis.

Si bien los resultados obtenidos evidencian el potencial de la metodología propuesta como herramienta de apoyo a la decisión, existen algunas consideraciones que deben tenerse presentes al interpretar sus resultados y delimitar el alcance de su aplicación. En primer lugar, el estudio se circunscribe a las entidades de intermediación financiera supervisadas por la SUGEF: bancos, cooperativas de ahorro y crédito, mutuales y financieras. Esta delimitación responde tanto al alcance de la investigación, orientada a apoyar decisiones de ahorro e inversión en intermediarios financieros, como a la naturaleza de los indicadores utilizados por WISE, seleccionados específicamente para caracterizar el desempeño de este tipo de entidades. En consecuencia, la metodología no es directamente aplicable a otros segmentos del sistema financiero regulado, como aseguradoras, operadoras de pensiones, puestos de bolsa o fondos de inversión, cuya evaluación responde a dinámicas financieras distintas y requeriría el desarrollo de indicadores y criterios metodológicos específicos.

Una segunda consideración se relaciona con la posibilidad de establecer una medida objetiva de confianza o exactitud para los ordenamientos generados. El *Risk-Aware Ranking* incorpora un criterio prudencial basado en los percentiles de la distribución histórica de cada indicador durante una ventana móvil de 36 meses. Para los indicadores en los que un mayor valor representa un mejor desempeño se utiliza el percentil 10, mientras que para aquellos en los que un menor valor es preferible se emplea el percentil 90. Estos umbrales delimitan el 80% central de la distribución empírica y dejan un 10% de las observaciones históricas en cada extremo. De esta forma, la metodología toma como referencia valores ubicados en zonas relativamente desfavorables de la distribución, reduciendo la influencia de mejoras puntuales y favoreciendo una valoración conservadora de la trayectoria reciente de las entidades.

No obstante, esta cobertura del 80% no constituye un nivel de confianza estadístico ni una medida de precisión del ranking. La naturaleza no supervisada del problema implica que no existe una etiqueta o un ordenamiento verdadero de las entidades contra el cual contrastar los resultados y calcular una medida objetiva de exactitud. Por esta razón, la validación de WISE se sustenta principalmente en la coherencia económico-financiera de los resultados, su interpretabilidad y el comportamiento empírico observado. En este contexto, los percentiles 10 y 90 constituyen una decisión metodológica que determina el grado de conservadurismo del componente *risk-aware*, pero no permiten atribuir al ranking un determinado porcentaje de confianza estadística.

El alcance actual de WISE se concentra, además, en el comportamiento financiero individual de cada intermediario. Por esta razón, el análisis no incorpora el riesgo intragrupo derivado de las interacciones entre entidades pertenecientes a un mismo conglomerado financiero ni las posibles relaciones con otras actividades desarrolladas dentro de dichos conglomerados. La incorporación de estas interdependencias representa una posible extensión de la metodología que permitiría complementar el análisis individual con una perspectiva sobre las relaciones financieras existentes dentro de los grupos económicos.

Por otra parte, WISE se construye deliberadamente a partir de cuatro indicadores financieros cuantitativos, públicos e interpretables, lo que favorece la reproducibilidad y trazabilidad de los resultados. Como contrapartida, el análisis no incorpora dimensiones cualitativas como el gobierno corporativo, la calidad de la gestión o determinados componentes del riesgo operativo, ni información interna de las entidades que no se encuentra disponible públicamente. Asimismo, la oportunidad y precisión de los resultados están condicionadas por la calidad, periodicidad y nivel de desagregación de la información divulgada por el supervisor financiero. La incorporación futura de nuevas fuentes públicas de información podría ampliar las dimensiones consideradas sin perder el carácter transparente y reproducible de la propuesta.

Finalmente, tanto el *Alignment Score* como el *Risk-Aware Ranking* incorporan decisiones metodológicas, entre ellas la definición del vector ideal, el procedimiento de estandarización, los percentiles utilizados y la ventana histórica de 36 meses. Estas decisiones se sustentan en criterios financieros, análisis exploratorio y experimentación empírica, pero constituyen parámetros susceptibles de análisis de sensibilidad y eventual recalibración ante cambios en el universo de intermediarios analizados, la disponibilidad de información o las condiciones del sistema financiero. Esta flexibilidad permite que la metodología pueda evolucionar a medida que se disponga de nueva evidencia, manteniendo al mismo tiempo una formulación transparente e interpretable. En consecuencia, los rankings obtenidos deben entenderse como medidas comparativas del comportamiento financiero relativo de las entidades y no como estimaciones de probabilidades de quiebra, intervención o insolvencia futura.

## 6 Conclusiones

WISE Investors demuestra que es posible transformar información financiera pública, dispersa y técnicamente compleja en conocimiento comparativo accesible para ahorrantes e inversionistas. El objetivo de la investigación se cumple mediante el desarrollo de una metodología interpretable capaz de sintetizar información financiera multidimensional y su implementación en una plataforma interactiva que facilita el análisis comparativo del comportamiento de las entidades supervisadas.

La principal contribución metodológica del trabajo corresponde a la Metodología WISE, un marco analítico que representa el comportamiento financiero de las entidades como puntos en un espacio vectorial definido por cuatro indicadores, identifica observaciones atípicas y perfiles financieros mediante técnicas de aprendizaje automático, cuantifica el alineamiento respecto a un perfil financiero ideal y construye rankings complementarios desde una perspectiva puntual y prudencial. El valor de la

propuesta radica en la integración coherente de representación matemática, aprendizaje automático e inteligencia artificial dentro de un sistema de apoyo a la decisión orientado al usuario final.

Además de construir rankings comparativos, la metodología permitió identificar una taxonomía compuesta por seis perfiles financieros diferenciados: Rentabilidad Agresiva, Alto Desempeño Integral, Desempeño Promedio Saludable, Presión Crediticia, Baja Cartera al Día y Rentabilidad Crítica. Esta clasificación constituye un aporte adicional de la investigación, ya que proporciona un marco interpretable para comprender la heterogeneidad del sistema financiero supervisado y complementar el análisis comparativo entre entidades.

El uso combinado de rutas SQL y RAG, en el agente del chat, permitió cubrir tanto consultas cuantitativas como cualitativas, validando la pertinencia de un enfoque híbrido. Esta integración redujo el riesgo de alucinaciones y mejoró la interpretabilidad de las respuestas. La evaluación de calidad, basada en métricas automáticas de similitud coseno y en auditoría manual de fidelidad, evidenció que el agente mantiene alineación con el contexto financiero recuperado, ofreciendo respuestas consistentes y útiles para el usuario final. En particular, la Tabla 4 muestra que el modelo *OpenAI gpt-5.4-mini* alcanzó la mayor similitud semántica (28.1%) y fidelidad perfecta (100%), aunque con una completitud moderada (58%). Los modelos de “Anthropic” lograron mayor cobertura (86% en *Haiku* y 82% en *Sonnet*), pero con menor estabilidad en la orquestación y tiempos de respuesta más altos. En términos de eficiencia operativa (Tabla 6), *gpt-5.4-mini* fue el más rápido (7.3 segundos promedio) y con menor consumo de tokens (4955 por consulta), mientras que *Claude Sonnet* alcanzó hasta 17.9 segundos y 5719 tokens. Estos resultados respaldan la viabilidad técnica de incorporar el agente en una plataforma interactiva y sugieren que *OpenAI gpt-5.4-mini* es el modelo más recomendable para un despliegue práctico, dado su balance favorable entre calidad, eficiencia y costo computacional.

El agente conversacional constituye un complemento valioso para reducir la asimetría de información y facilitar la exploración comparativa entre entidades financieras supervisadas. Su integración con WISE permite responder consultas tanto cuantitativas como cualitativas, apoyándose en rutas SQL y RAG para mejorar la interpretabilidad y reducir el riesgo de alucinaciones. Los resultados experimentales muestran que el modelo *OpenAI gpt-5.4-mini* ofrece el mejor balance: alcanzó la mayor similitud semántica (28.1%), fidelidad perfecta (100%) y tiempos de respuesta más bajos (7.3 segundos en promedio con 4955 tokens por consulta). Estos indicadores respaldan su uso como modelo base en aplicaciones prácticas, dado que combina eficiencia, calidad y costo computacional favorable.

WISE Investors debe interpretarse como una herramienta comparativa que complementa el análisis financiero y el criterio del usuario dentro del proceso de toma de decisiones. Los *rankings* generados, como el *Alignment Score* y el *Risk-aware Ranking*, aportan medidas objetivas del comportamiento financiero relativo de las entidades, con fidelidad del 100% en todas las respuestas evaluadas y completitud superior en los modelos de Anthropic (86% en *Haiku* y 82% en *Sonnet*). Sin embargo, considerando la eficiencia operativa y la estabilidad en la orquestación, se recomienda priorizar el uso de *OpenAI gpt-5.4-mini*, que ofrece un desempeño más consistente y económico,

reservando las variantes de mayor escala para escenarios donde la cobertura y robustez sean críticas aun a costa de mayor consumo de recursos.

Como líneas futuras de investigación se propone incorporar variables macroeconómicas y cualitativas, evaluar configuraciones alternativas del vector ideal y de la metodología *Risk-aware*, explorar esquemas de ponderación diferenciada para los indicadores financieros y validar la metodología en otros sistemas financieros. Asimismo, la arquitectura modular desarrollada facilita la incorporación de nuevas capacidades analíticas y conversacionales, ampliando el potencial de WISE Investors como plataforma de apoyo para decisiones de ahorro e inversión mejor fundamentadas.

## Reconocimientos

En primera instancia a la Universidad Cenfotec por permitir un espacio adecuado para la elaboración de este proyecto de investigación. También se reconoce que el uso de la Inteligencia Artificial, como asistente, durante el desarrollo del proyecto, aunque el análisis, la arquitectura, diseño y revisión son reconiminiento total de los autores.

## Material complementario

El material complementario que acompaña este artículo consiste en 3 diferentes dashboards metodológicos interactivos, disponibles en zenodo. [27];, son los siguientes:

1. *Dashboard I. Exploratory Univariate Analysis*: Dashboard para documentar, de forma completa, el tratamiento de valores faltantes e incidencias en los datos muestrales, así como el análisis exploratorio y univariado de la información de entidades financieras.
2. *Dashboard II. Multivariate Analysis*: Incluye el análisis multivariado de indicadores financieros, proceso de segmentación y perfiles de *clustering* utilizados para la creación del ranking.
3. *Dashboard III. WISE Rankings*: Documentación del ranking mensual y el *Risk-aware Ranking*.

## References

- [1] Mishkin, F.S.: The Economics of Money, Banking, and Financial Markets, 12 edn. Pearson, Boston, MA (2018)
- [2] Entidades Financieras (SUGEF), S.G.: Sitio web oficial de SUGEF (2026). <https://www.sugef.fi.cr>
- [3] Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd edn. Springer, New York (2009). <https://doi.org/10.1007/978-0-387-84858-7>
- [4] Entidades Financieras (SUGEF), S.G.: Indicadores financieros (2025). [https://www.sugef.fi.cr/rsw\\_reportes/reporteIndicadoresFinancieros](https://www.sugef.fi.cr/rsw_reportes/reporteIndicadoresFinancieros)

- [5] Ortiz, D., Liberman, L.: El sistema de intermediación financiera costarricense: evolución, desafíos y perspectivas 2014-2024. Academia de Centroamérica, San José, Costa Rica. Estudio sectorial sobre el sistema financiero costarricense (2025). [https://www.academiaca.or.cr/wp-content/uploads/2025/12/scif\\_20142024\\_presentacion.pdf](https://www.academiaca.or.cr/wp-content/uploads/2025/12/scif_20142024_presentacion.pdf)
- [6] Capital, F.: Análisis & Estrategia: Perspectivas Económicas y Políticas de Costa Rica 2025-2026 (2026). <https://fcscapital.cr/publicaciones/>
- [7] Costa Rica, C.: Comunicados de prensa y estudios financieros. Sitio oficial. Publicaciones sobre estabilidad financiera y educación económica (2026). <https://www.camaradebancos.fi.cr>
- [8] Financiero, P.: Informes técnicos y plataformas de visualización del sistema financiero costarricense. Portal institucional. Informes sectoriales y análisis financiero especializado (2026). <https://www.perfilfinanciero.com>
- [9] Ratings, M.: Global Credit Ratings & Research Insights. Moody's Official Website. Evaluaciones independientes sobre riesgo crediticio y fortaleza financiera (2026). <https://ratings.moody.com>
- [10] Ratings, F.: Credit Ratings & Analysis for Financial Markets. Fitch Ratings Official Website. Informes de calificación crediticia y análisis sectorial (2026). <https://www.fitchratings.com>
- [11] Ratings, S.G.: Credit Ratings: Independent Opinions of Credit Risk. S&P Global Official Website. Calificaciones de emisores y emisiones financieras a nivel global (2026). <https://www.spglobal.com/ratings>
- [12] Jain, A.K.: Data clustering: 50 years beyond k-means. Pattern Recognition Letters **31**(8), 651–666 (2010) <https://doi.org/10.1016/j.patrec.2009.09.011>
- [13] Mercadier, M., Tarazi, A., Armand, P., Lardy, J.-P.: Monitoring bank risk around the world using unsupervised learning. European Journal of Operational Research **324**(2), 590–615 (2025) <https://doi.org/10.1016/j.ejor.2025.01.036>
- [14] Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffmann, A., Giovannini, E.: Handbook on constructing composite indicators: Methodology and user guide. Technical Report 2005/03, OECD Publishing (2005). <https://doi.org/10.1787/533411815016>
- [15] OECD, European Union, European Commission, Joint Research Centre: Handbook on Constructing Composite Indicators: Methodology and User Guide. OECD Publishing, Paris (2008). <https://doi.org/10.1787/9789264043466-en>
- [16] Jara, A., Lizano, J.: wise\_investors: Python tools for processing and analyzing SUGEF financial indicators. [Software]. GitHub (2026). <https://github.com/>

- [17] Deisenroth, M.P., Faisal, A.A., Ong, C.S.: Mathematics for Machine Learning. Cambridge University Press, Cambridge (2020). <https://doi.org/10.1017/9781108679930>
- [18] Axler, S.: Linear Algebra Done Right, 3rd edn. Springer, Cham (2015). <https://doi.org/10.1007/978-3-319-11080-6>
- [19] Liu, F.T., Ting, K.M., Zhou, Z.-H.: Isolation forest. In: Proceedings of the 2008 Eighth IEEE International Conference on Data Mining, pp. 413–422. IEEE, Piscataway, NJ (2008). <https://doi.org/10.1109/ICDM.2008.17>
- [20] MacQueen, J.: Some methods for classification and analysis of multivariate observations. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, pp. 281–297. University of California Press, Berkeley, CA (1967)
- [21] Jeong, S., Baek, J., Cho, S., Sung, M., Park, J.C.: Adaptive-rag: Learning to adapt retrieval-augmented large language models through question complexity. arXiv preprint arXiv:2403.14403 (2024)
- [22] Biswal, A., Patel, L., Jha, S., Kamsetty, A., Liu, S., Gonzalez, J.E., Guestrin, C., Zaharia, M.: Text2sql is not enough: Unifying ai and databases with tag. arXiv preprint arXiv:2408.14717 (2024)
- [23] Yan, S.-Q., Gu, J.-C., Zhu, Y., Ling, Z.-H.: Corrective retrieval augmented generation. arXiv preprint arXiv:2401.15884 (2024)
- [24] Es, S., James, J., Espinosa-Anke, L., Schockaert, S.: Ragas: Automated evaluation of retrieval augmented generation. arXiv preprint arXiv:2309.15217 (2023)
- [25] Liu, X., Yu, H., Zhang, H., Xu, Z., Lei, X., Lai, Y., Gu, Y., Ding, H., Men, K., Yang, K., *et al.*: Agentbench: Evaluating llms as agents. In: International Conference on Learning Representations (ICLR) (2024)
- [26] Bai, L., *et al.*: How do coding agents spend your money? analyzing and predicting token consumption patterns. In: The Fourteenth International Conference on Learning Representations (ICLR) (2025). <https://openreview.net/forum?id=1bUeVB3fov>
- [27] Cordero, A.J., Gallegos, J.A.L.: Material complementario - Prototipo Funcional Basado en Inteligencia Artificial para Evaluar el Riesgo Crediticio que asumen los Ahorrantes en Entidades Financieras Costarricenses Reguladas por la SUGEF. Zenodo (2026). <https://doi.org/10.5281/zenodo.21908137> . <https://doi.org/10.5281/zenodo.21908137>