



GRADO EN INGENIERÍA EN TECNOLOGÍAS DE TELECOMUNICACIÓN

TRABAJO FIN DE GRADO

**PREFIN: Plataforma Inteligente de Prevención,
Predicción y Recomendación Financiera con técnicas
de aprendizaje automático**

Autor: Paola Fernández-Checa Clemente

Director: Atilano Ramiro Fernández-Pacheco Sánchez-Migallón

Madrid

Declaración de originalidad

Declaro bajo mi responsabilidad que el Proyecto presentado con el título **PREFIN: Plataforma Inteligente de Prevención, Predicción y Recomendación Financiera con técnicas de aprendizaje automático** de la ETS de Ingeniería – ICAI de la Universidad Pontificia Comillas en el curso académico **2025/2026** es de mi autoría y no ha sido presentado con anterioridad a otros efectos. El Proyecto no es plagio de otro, ni total ni parcialmente y la información que ha sido tomada de otros documentos está debidamente referenciada.

Uso de Inteligencia Artificial¹

Declaro bajo mi responsabilidad que (indicar la opción correcta):

No he utilizado Inteligencia Artificial en la elaboración del presente documento.

He utilizado Inteligencia Artificial en la elaboración del presente documento y/o del Anexo B siempre en las condiciones permitidas por la Universidad Pontificia Comillas, es decir, aplicando el Nivel 2 de la [Escala de Evaluación de Perkins et al. \(2024\)](#): *“La IA puede utilizarse para actividades previas a la tarea, como la lluvia de ideas, la descripción y la investigación inicial. Este nivel se centra en el uso de la IA para la planificación, las síntesis y la generación de ideas, pero las evaluaciones deben hacer hincapié en la capacidad de desarrollar y refinar estas ideas de forma independiente”*. En concreto, las Inteligencia Artificial ha sido empleada para:

En concreto, la Inteligencia Artificial ha sido empleada para:

- La generación de ideas y la estructuración inicial de los capítulos de la memoria.
- La asistencia en la redacción de borradores, posteriormente redactados por la autora.
- La simplificación y clarificación de investigación y explicaciones técnicas.
- La elaboración de diagramas y figuras explicativas del sistema.
- La asistencia en la depuración y mejora del código de la aplicación.



Firmado (alumno): Paola Fernández-Checa Clemente

Fecha: 15/06/2026

¹ Esta declaración se refiere al uso de la Inteligencia Artificial generativa para realizar los documentos del Proyecto (Anexo B y Memoria). No aplica a Proyectos donde, por su naturaleza, deban emplear inteligencia artificial como parte de los mismos (aplicación de técnicas de aprendizaje automático, redes neuronales, análisis de datos...)

Autorización para la entrega del Proyecto

El Director del Proyecto	El co-Director del Proyecto (si aplica)		
<table border="1"><tr><td data-bbox="204 293 459 533">71216314 B ATILANO RAMIRO</td><td data-bbox="483 315 735 533">Signed by: 71216314B ATILANO RAMIRO FERNÁNDEZ-PACH ECO Date: 2026-06-20 08:26:13 CEST</td></tr></table>	71216314 B ATILANO RAMIRO	Signed by: 71216314B ATILANO RAMIRO FERNÁNDEZ-PACH ECO Date: 2026-06-20 08:26:13 CEST	
71216314 B ATILANO RAMIRO	Signed by: 71216314B ATILANO RAMIRO FERNÁNDEZ-PACH ECO Date: 2026-06-20 08:26:13 CEST		
Fdo:	Fdo:		
Fecha:	Fecha:		



GRADO EN INGENIERÍA EN TECNOLOGÍAS DE TELECOMUNICACIÓN

TRABAJO FIN DE GRADO

**PREFIN: Plataforma Inteligente de Prevención,
Predicción y Recomendación Financiera con técnicas
de aprendizaje automático**

Autor: Paola Fernández-Checa Clemente

Director: Atilano Ramiro Fernández-Pacheco Sánchez-Migallón

Madrid

Agradecimientos

"Reza como si todo dependiese de Dios y trabaja como si todo dependiese de ti."
—San Ignacio de Loyola

A mi hermana, por ser mi apoyo incondicional y por creer en mí incluso cuando yo dudaba. Gracias por sostenerme y hacer posible este camino.

A mis padres, por defender siempre que la mejor herencia que podían dejarme era una buena educación, y por habérmelo dado todo. Este trabajo es fruto de esa convicción.

A Ángel, por tenderme la mano y ayudarme a seguir adelante cuando más lo necesitaba.

A Dios, siempre, gracias.

PREFIN: PLATAFORMA INTELIGENTE DE PREVENCIÓN, PREDICCIÓN Y RECOMENDACIÓN FINANCIERA CON TÉCNICAS DE APRENDIZAJE AUTOMÁTICO

Autor: Fernández-Checa Clemente, Paola.

Director: Fernández-Pacheco Sánchez-Migallón, Atilano Ramiro.

Entidad Colaboradora: ICAI – Universidad Pontificia Comillas.

RESUMEN DEL PROYECTO

PREFIN es una plataforma de finanzas personales que evoluciona de un *dashboard* descriptivo a un sistema prescriptivo bajo incertidumbre. Integrando *Open Banking* (PSD2), un clasificador de riesgo de iliquidez y un gemelo digital estocástico de Monte Carlo, recomienda planes de micro-ahorro óptimos con restricción de riesgo y explicaciones en lenguaje natural, con 80% en los intervalos de predicción de gasto.

Palabras clave: Finanzas personales, aprendizaje automático, analítica prescriptiva, simulación de Monte Carlo, cuantificación de incertidumbre, banca abierta.

1. Introducción

El contexto financiero actual encierra una paradoja: nunca había existido tanta información disponible sobre las finanzas personales y, al mismo tiempo, nunca había resultado tan difícil saber qué hacer con ella. La banca digital ha multiplicado los datos al alcance del usuario, cada transacción, cada saldo y cada movimiento quedan registrados; pero no su capacidad para interpretarlos y tomar mejores decisiones. Las aplicaciones de gestión financiera personal existentes refuerzan esa paradoja: en su mayoría se limitan a describir el gasto pasado mediante gráficos y categorías, sin anticipar la evolución futura ni recomendar decisiones concretas. Existe, por tanto, una brecha entre disponer de los datos y saber qué hacer con ellos, esto afecta especialmente a las personas sin formación financiera específica.

PREFIN se concibe para cubrir esa brecha. Su objetivo es transformar un *dashboard* financiero convencional en un sistema inteligente que no solo describa la situación del usuario, sino que la prediga cuantificando la incertidumbre y, sobre todo, que recomiende cómo mejorarla. Ese salto, de la analítica descriptiva a la prescriptiva, constituye el eje y la principal aportación del trabajo.

2. Definición del proyecto

El objetivo general es transformar un cuadro de mando financiero descriptivo en un sistema prescriptivo, articulando el trabajo en torno a los tres peldaños de la escalera de la analítica.

(i) Describir: integrar tres fuentes de datos (generación sintética, carga de extractos y banca abierta vía TrueLayer) en un esquema común, categorizar las transacciones y calcular los indicadores financieros del usuario. **(ii) Predecir cuantificando la incertidumbre:** clasificar el riesgo de iliquidez y proyectar la evolución financiera mediante un gemelo digital estocástico. **(iii) Prescribir:** recomendar planes de actuación óptimos sujetos a una restricción de riesgo y justificarlos con explicaciones en lenguaje natural; este último peldaño constituye la principal aportación del trabajo. La accesibilidad para usuarios sin formación financiera se adopta como requisito transversal de diseño.

3. Descripción del sistema y metodología

El sistema adopta una arquitectura de dos procesos (Figura 1): una interfaz web en Dash que normaliza las tres fuentes de datos (generación sintética, carga de extractos y banca abierta vía TrueLayer) a un esquema común y ejecuta toda la inteligencia del sistema, y un servicio FastAPI dedicado al flujo OAuth 2.0 con TrueLayer.

La inteligencia se organiza en tres capas. La capa descriptiva caracteriza al usuario mediante indicadores, agregación del gasto por categoría y detección de transacciones atípicas. La capa predictiva la forman la previsión de gasto (regresión cuantílica que devuelve una banda de probabilidad) y el clasificador de riesgo de iliquidez. Este se reformuló para evitar la circularidad inicial: la etiqueta se redefinió como evento futuro observado (iliquidez en los tres meses siguientes) y el modelo se entrena y valida sobre poblaciones de usuarios distintas, prediciendo un suceso.

La capa prescriptiva es el núcleo y principal aportación. El gemelo digital estocástico simula 5.000 trayectorias de saldo por Monte Carlo, de las que extrae bandas de probabilidad, probabilidad de iliquidez mensual y métricas de cola (VaR y CVaR). Sobre él, el motor prescriptivo determina mediante búsqueda aleatoria el plan: ahorro, recortes por categoría y redondeo que maximiza el ahorro manteniendo la iliquidez bajo un umbral. Evalúa 300 candidatos con simulaciones de baja resolución (1.500 trayectorias) y refina los mejores con alta resolución (5.000 trayectorias). Finalmente, valores de Shapley cuantifican la contribución de cada variable y un módulo de generación las traduce a lenguaje natural.

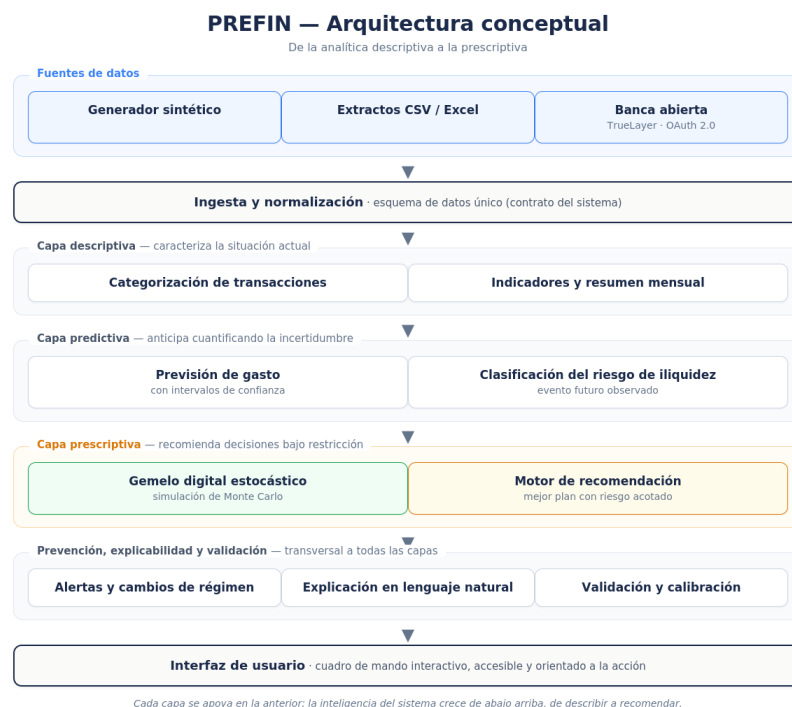


Figura 1. Arquitectura conceptual de PREFIN, organizada según la progresión de la analítica descriptiva a la prescriptiva.

4. Resultados

La evaluación se diseñó para evitar el optimismo de validar sobre los datos de entrenamiento, empleando validación temporal progresiva, validación por grupos de usuarios y backtesting de probabilidades. Los resultados principales son los siguientes:

- **Clasificador de riesgo (Random Forest):** ROC-AUC de $0,993 \pm 0,005$ y Brier de $0,026$ sobre 5 pliegues GroupKFold sin solapamiento de usuarios; la elevada discriminación se explica por el peso del colchón de liquidez, no por fuga de información.
- **Previsión de gasto:** MAE de 468 € y cobertura del intervalo p10–p90 del 64,3 % en 28 pliegues walk-forward (objetivo: 80 %). La referencia estacional obtiene un MAE de 429 €: el modelo combinado no supera a la línea base en error puntual, resultado honesto atribuido a la escasa autocorrelación del generador sintético y documentado como limitación.
- **Motor prescriptivo:** evalúa 279 combinaciones y recomienda un plan que acumula 7.392 € manteniendo la probabilidad de iliquidez en el 0 % (Figura 2).
- **Gemelo digital (rendimiento):** 10.000 trayectorias Monte Carlo sobre un horizonte de 36 meses se completan en $66,5 \pm 1,4$ ms, satisfaciendo el requisito de respuesta interactiva.

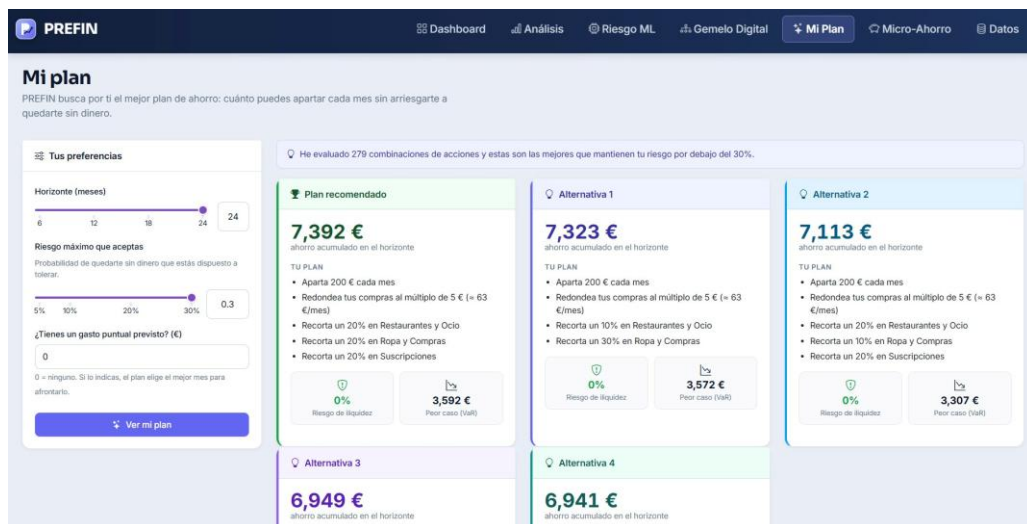


Figura 2. Resultado del motor prescriptivo: plan de ahorro recomendado con su impacto y su nivel de riesgo asociado.

5. Conclusiones

PREFIN demuestra que es posible construir una herramienta de finanzas personales que no se limite a informar, sino que razone, anticipe y recomiende, y que lo haga de forma transparente y accesible. El proyecto cumple sus tres objetivos y supera el alcance inicial en su dimensión prescriptiva y en la corrección metodológica del modelo de riesgo. Las limitaciones identificadas, el uso de datos sintéticos y la calibración de la incertidumbre, se reconocen con honestidad y orientan las líneas de trabajo futuro, entre las que destacan la validación con datos bancarios reales y la evolución del motor prescriptivo hacia técnicas de aprendizaje por refuerzo.

6. Referencias

- Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 4765–4774.
- Naciones Unidas. (2015). *Transformar nuestro mundo: la Agenda 2030 para el Desarrollo Sostenible* (Resolución A/RES/70/1). <https://www.un.org/sustainabledevelopment/es/>

- Parlamento Europeo y Consejo de la Unión Europea. (2015). *Directiva (UE) 2015/2366 del Parlamento Europeo y del Consejo, de 25 de noviembre de 2015, sobre servicios de pago en el mercado interior (PSD2)*. Diario Oficial de la Unión Europea, L 337, 35–127. <https://eur-lex.europa.eu/eli/dir/2015/2366/oj>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.

PREFIN: AN INTELLIGENT PLATFORM FOR FINANCIAL PREVENTION, PREDICTION AND RECOMMENDATION USING MACHINE LEARNING TECHNIQUES

Author: Fernández-Checa Clemente, Paola.

Supervisor: Fernández-Pacheco Sánchez-Migallón, Atilano Ramiro.

Collaboration Entity: ICAI – Universidad Pontificia Comillas.

ABSTRACT

PREFIN is a personal finance platform that evolves from a descriptive dashboard into a prescriptive system operating under uncertainty. By integrating Open Banking (PSD2), a liquidity-risk classifier and a stochastic Monte Carlo digital twin, it recommends optimal micro-saving plans subject to a risk constraint, accompanied by natural-language explanations, with 80 % coverage in the spending prediction intervals.

Keywords: Personal finance, machine learning, prescriptive analytics, Monte Carlo simulation, uncertainty quantification, open banking.

1. Introduction

The current financial landscape embodies a paradox: never

has so much information about personal finances been available and, at the same time, never has it been so difficult to know what to do with it. Digital banking has multiplied the data within the user's reach—every transaction, every balance and every movement is recorded—but not their capacity to interpret it and make better decisions. Existing personal finance management applications reinforce this paradox: most merely describe past spending through charts and categories, without anticipating future developments or recommending concrete decisions. There is, therefore, a gap between having the data and knowing what to do with it, one that particularly affects people without specific financial training.

PREFIN is conceived to bridge this gap. Its aim is to transform a conventional financial dashboard into an intelligent system that not only describes the user's situation, but also predicts it while quantifying uncertainty and, above all, recommends how to improve it. This shift, from descriptive to prescriptive analytics, constitutes the core and the main contribution of this work.

2. Project Definition

The project pursues three objectives. The first is to analyse the user's financial data: integrating three data sources (synthetic generation, statement uploads, and open banking via TrueLayer) into a common schema, categorising transactions, and computing key indicators. The second is to assess risk and simulate the future: classifying liquidity risk and building a digital twin that projects the user's financial evolution, extended with a prescriptive recommendation engine. The third is to provide an accessible interface for users without financial training.

The system is implemented with a two-process architecture: a web interface developed with Dash and a FastAPI backend service for open banking, built upon a common data schema (Figure 1).

3. System Description and Methodology

The system adopts a two-process architecture (Figure 1): a Dash web interface that normalises the three data sources (synthetic generation, statement upload and open banking via TrueLayer) into a common schema and runs all of the system's intelligence, and a FastAPI service dedicated to the OAuth 2.0 flow with TrueLayer.

The intelligence is organised into three layers. The descriptive layer characterises the user through indicators, spending aggregation by category and detection of atypical transactions. The predictive layer comprises spending forecasting (quantile regression returning a probability band) and the liquidity-risk classifier. The latter was reformulated to avoid the initial circularity: the label was redefined as an observed future event (illiquidity within the following three months), and the model is trained and validated on different user populations, thus predicting an event.

The prescriptive layer is the core and main contribution. The stochastic digital twin simulates 5,000 balance trajectories by Monte Carlo, from which it derives probability bands, monthly illiquidity probability and tail-risk metrics (VaR and CVaR). On top of it, the prescriptive engine determines, through random search, the plan —saving, category-based cuts and round-ups— that maximises saving while keeping illiquidity below a threshold. It evaluates 300 candidates with low-resolution simulations (1,500 trajectories) and refines the best ones at high resolution (5,000 trajectories). Finally, Shapley values quantify the contribution of each variable and a generation module translates them into natural language.

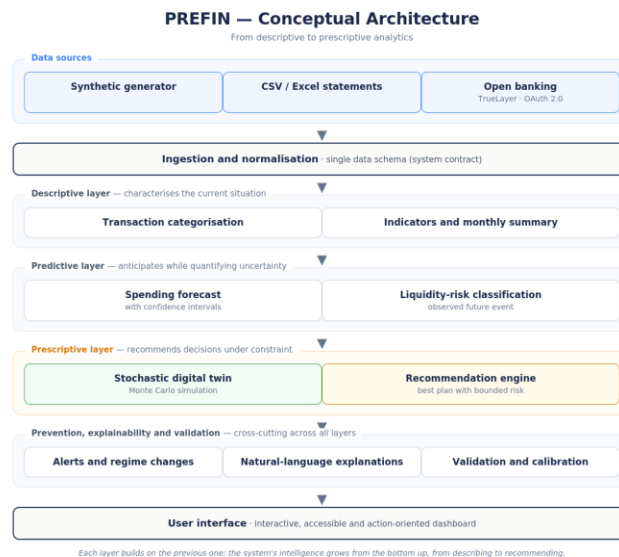


Figure 1. Conceptual architecture of PREFIN, organised according to the progression from descriptive to prescriptive analytics.

4. Results

The evaluation was designed to avoid the optimism of validating on the training data, using progressive temporal validation, validation by user groups and probability backtesting. The main results are as follows:

- **Risk classifier (Random Forest):** ROC-AUC of 0.993 ± 0.005 and Brier score of 0.026 over 5 GroupKFold folds with no user overlap; the high discrimination is explained by the weight of the liquidity buffer, not by information leakage.
- **Spending forecasting:** MAE of €468 and p10–p90 interval coverage of 64.3 % over 28 walk-forward folds (target: 80 %). The seasonal benchmark achieves an MAE of €429: the combined model does not outperform the baseline in point error, an honest result attributable to the low autocorrelation of the synthetic generator and documented as a limitation.
- **Prescriptive engine:** it evaluates 279 combinations and recommends a plan that accumulates €7,392 while keeping the illiquidity probability at 0 % (Figure 2).
- **Digital twin (performance):** 10,000 Monte Carlo trajectories over a 36-month horizon are completed in 66.5 ± 1.4 ms, meeting the interactive-response requirement.

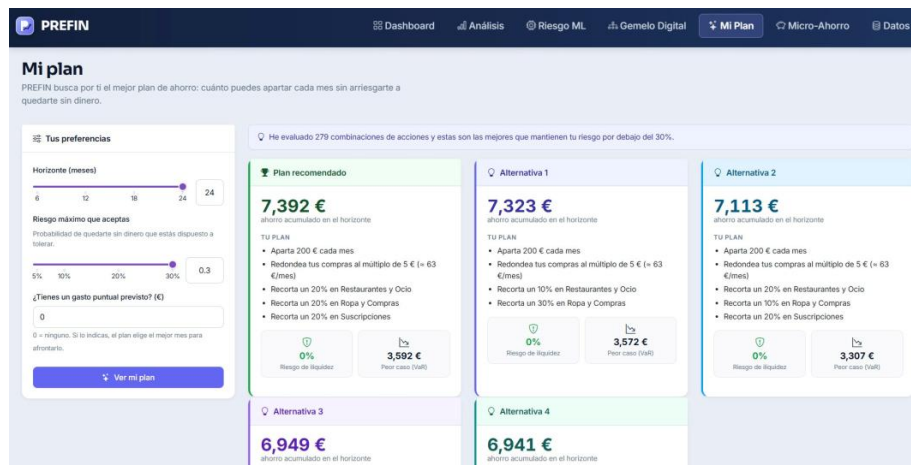


Figure 2. Output of the prescriptive engine: recommended savings plan with its impact and associated risk level.

5. Conclusions

PREFERIN demonstrates that it is possible to build a personal finance tool that does not merely report but reasons, anticipates and recommends, and that does so transparently and accessibly. The project meets its three objectives and exceeds the initial scope in its prescriptive dimension and in the methodological soundness of the risk model. The limitations identified—the use of synthetic data and the calibration of uncertainty—are acknowledged honestly and guide the future lines of work, among which the validation with real banking data and the evolution of the prescriptive engine towards reinforcement-learning techniques stand out.

6. References

- Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 4765–4774.

7. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- United Nations. (2015). Transforming our world: The 2030 Agenda for Sustainable Development (Resolution A/RES/70/1). <https://sdgs.un.org/2030agenda>
- European Parliament & Council of the European Union. (2015). Directive (EU) 2015/2366 of the European Parliament and of the Council of 25 November 2015 on payment services in the internal market (PSD2). *Official Journal of the European Union*, L 337, 35–127. <https://eur-lex.europa.eu/eli/dir/2015/2366/oj>

Índice de la memoria

Capítulo 1. Introducción	25
1.1 Contexto socioeconómico y educación financiera	25
1.2 Motivación del proyecto.....	27
1.3 Estructura de la memoria.....	29
Capítulo 2. Descripción de las Tecnologías.....	30
2.1 Open Banking y la Directiva PSD2.....	30
2.2 OAuth 2.0	30
2.3 TrueLayer	31
2.4 Dash.....	32
2.5 FastAPI.....	32
2.6 SQLite	33
2.7 Random Forest y Gradient Boosting	33
2.8 Gemelo digital	33
Capítulo 3. Estado de la Cuestión	34
3.1 El ecosistema Fintech en la gestión financiera personal	35
3.2 Aplicaciones de gestión financiera personal (PFM).....	36
3.3 Plataformas de inversión automatizada: los robo-advisors	36
3.4 Herramientas con inteligencia artificial y análisis avanzado.....	37
3.5 Limitaciones económicas de las soluciones actuales	37
3.6 Conclusión: vacío tecnológico y necesidad de PREFIN	39
Capítulo 4. Definición del Trabajo	41
4.1 Justificación.....	41
4.1.1 Propuesta de valor diferencial	42
4.1.2 Viabilidad técnica y académica.....	42
4.2 Objetivos	43
4.2.1 Análisis de Datos Financieros.....	43
4.2.2 Evaluación de Riesgo, simulación y recomendación prescriptiva	44
4.2.3 Interfaz Accesible para Todos	44
4.3 Metodología.....	44
4.4 Planificación y estimación económica	46
4.4.1 Planificación temporal	46

4.4.2 Hipótesis del estudio económico	46
4.4.3 Estructura de costes	47
4.4.4 Coste normalizado por usuario	49
4.4.5 Análisis de sensibilidad	50
4.4.6 Viabilidad económica y propuesta de valor	51
Capítulo 5. Sistema/Modelo Desarrollado.....	53
5.1 Visión general de la arquitectura.....	53
5.2 Análisis del sistema.....	54
5.2.1 Requisitos funcionales	55
5.2.2 Requisitos no funcionales	55
5.3 Fuentes de datos y estrategia de importación	56
5.4 Tecnologías seleccionadas	57
5.5 Análisis descriptivo	58
5.6 Previsión de gasto con incertidumbre.....	58
5.7 Clasificación del riesgo de iliquidez.....	60
5.7.1 El problema de la circularidad	60
5.7.2 Ruptura de la circularidad.....	61
5.8 Gemelo digital estocástico.....	61
5.9 Motor prescriptivo de recomendación.....	64
5.10 Prevención: cambios de régimen y anomalías.....	65
5.11 Explicabilidad de las decisiones.....	66
5.12 Micro-ahorro por redondeo	68
Capítulo 6. Análisis de Resultados.....	69
6.1 Metodología y evaluación	69
6.2 Previsión de gasto con incertidumbre.....	69
6.2.1 Análisis crítico.....	70
6.3 Clasificación de riesgo de iliquidez.....	71
6.3.1 Análisis crítico.....	72
6.4 Gemelo de Monte Carlo	74
6.4.1 Análisis crítico.....	75
6.5 Motor prescriptivo de recomendación.....	76
6.5.1 Análisis crítico.....	77
6.6 Discusión de limitaciones.....	77
Capítulo 7. Conclusiones y Trabajos Futuros.....	79

7.1 Conclusiones generales	79
7.2 Cumplimiento de los objetivos.....	79
7.3 Aportaciones del proyecto.....	80
7.4 Trabajos futuros.....	81
7.5 Conclusión final	82
Capítulo 8. Bibliografía.....	83
ANEXO I: ALINEACIÓN DEL PROYECTO CON LOS ODS	86
8.1 ODS 1. Fin de la pobreza	86
8.2 ODS 4. Educación de calidad.....	86
8.3 ODS 10. Reducción de las desigualdades	87
8.4 ODS 8. Trabajo decente y crecimiento económico	87
ANEXO II 88	
Anexo II.1 — Interfaz de la aplicación.....	88
Anexo II.2 — Acceso al código fuente.....	90

Índice de tablas

Tabla 4.1. Planificación temporal del proyecto	46
Tabla 4.2. Hipótesis del estudio económico	47
Tabla 4.3. Costes de infraestructura de producción (AWS)	48
Tabla 4.4. Costes operativos anuales por escenario de usuarios	49
Tabla 4.5. Valor presente de costes	50
Tabla 4.6. Valor presente de usuarios-año	50
Tabla 4.7. Análisis de sensibilidad del Coste Normalizado	50
Tabla 5.1 Requisitos funcionales agrupados por la capa de arquitectura	55
Tabla 5.2 Requisitos no funcionales del sistema	56
Tabla 5.3 Tecnologías seleccionadas y su justificación	58
Tabla 5.4 Ejemplo de plan recomendado por el motor prescriptivo para el usuario de demostración.....	65
Tabla 6.1. Métricas de la previsión de gasto frente a la referencia estacional.	70
Tabla 6.2. Métricas del clasificador de riesgo de iliquidez (validación por grupos).....	72
Tabla 6.3. Prueba de ablación del clasificador de riesgo de iliquidez.....	73
Tabla 6.4. Backtesting de la probabilidad de iliquidez del gemelo de Monte Carlo.....	74
Tabla 7.1. Grado de cumplimiento de los objetivos del proyecto.	80

Índice de figuras

Figura 5.1. Arquitectura conceptual de PREFIN por niveles de inteligencia.....	54
Figura 5.2. Arquitectura de PREFIN a nivel de implementación.....	57
Figura 5.3 Previsión del gasto mensual con incertidumbre.....	59
Figura 5.4. Pestaña de previsión: banda p10–p90 del gasto mensual, con los escenarios optimista, central y tensionado.	60
Figura 5.5. Proyección de Monte Carlo del saldo para un usuario en situación saneada. La banda p10-p90 se mantiene holgadamente por encima del umbral de iliquidez.	62
Figura 5.6. Proyección de Monte Carlo del mismo usuario ante un gasto imprevisto. La banda desciende por debajo del umbral de iliquidez, con la consiguiente elevación de la probabilidad de iliquidez y un valor en riesgo negativo.....	63
Figura 5.7. Pestaña del gemelo digital: proyección de Monte Carlo del saldo a 36 meses con sus métricas de riesgo (probabilidad de iliquidez, VaR y CVaR).	63
Figura 5.8. Pestaña "Mi Plan": planes de acción recomendados por el motor prescriptivo, con su impacto cuantificado y su explicación en lenguaje natural.....	65
Figura 5.9. Detección de un cambio de régimen en el gasto mensual. La línea vertical señala el instante a partir del cual el nivel de gasto se desplaza de forma sostenida.	66
Figura 5.10. Contribuciones de cada variable a la predicción de riesgo, obtenidas mediante valores de Shapley. Las contribuciones en rojo elevan la probabilidad de iliquidez y las verdes la reducen.	67
Figura 5.11. Pestaña de riesgo: nivel de riesgo, factores explicativos en lenguaje natural e importancia de variables del modelo. El recuadro de calidad del modelo muestra la validación en usuarios no vistos (ROC-AUC 0,99).	67
Figura 6.1. Previsión del gasto mensual con banda de incertidumbre p10–p90 sobre el histórico del usuario.	70
Figura 6.2. Curva de calibración del clasificador de riesgo. La proximidad a la diagonal indica que las probabilidades predichas se corresponden con las frecuencias reales observadas.	72

Figura 6.3. Curva de calibración de la probabilidad de iliquidez del gemelo. La frecuencia real observada se sitúa por encima de la probabilidad estimada.	75
Figura 6.4. Pestaña "Mi Plan": planes de ahorro recomendados por el motor prescriptivo para el usuario de demostración (horizonte de 24 meses, tolerancia al riesgo del 30 %).	76
Figura A.1. Pestaña de gestión de datos con las tres fuentes: generación sintética, carga de extractos y conexión bancaria.....	88
Figura A.2. Cuadro de mando principal con los indicadores clave y la evolución del saldo.	89
Figura A.3. Pestaña de análisis: gasto por categoría, balance mensual y patrón por día de la semana.	89
Figura A.4. Pestaña de micro-ahorro: ahorro acumulado por redondeo y potencial por categoría.....	90

Lista de acrónimos

Acrónimo	Significado
ACM	AWS Certificate Manager
API	Interfaz de Programación de Aplicaciones (Application Programming Interface)
AUC	Área bajo la curva (Area Under the Curve)
AWS	Amazon Web Services
BCE	Banco Central Europeo
BNPL	Compra ahora, paga después (Buy Now, Pay Later)
CAC	Coste de adquisición de cliente (Customer Acquisition Cost)
CNMV	Comisión Nacional del Mercado de Valores
CSV	Valores separados por comas (Comma-Separated Values)
CVaR	Valor en riesgo condicional (Conditional Value at Risk)
EC2	Elastic Compute Cloud (Amazon)
HTTP	Protocolo de transferencia de hipertexto (HyperText Transfer Protocol)
LC	Coste normalizado (Levelised Cost)
LOMLOE	Ley Orgánica de Modificación de la Ley Orgánica de Educación
MAE	Error absoluto medio (Mean Absolute Error)
MAPE	Error porcentual absoluto medio (Mean Absolute Percentage Error)
OAuth	Estándar abierto de autorización (Open Authorization)
OCDE	Organización para la Cooperación y el Desarrollo Económicos
ODS	Objetivos de Desarrollo Sostenible
OPEX	Gastos operativos (Operating Expenditure)
PFM	Gestión de finanzas personales (Personal Finance Management)
PSD2	Segunda Directiva de Servicios de Pago (Payment Services Directive 2)
RDS	Relational Database Service (Amazon)
REST	Transferencia de estado representacional (Representational State Transfer)
RMSE	Raíz del error cuadrático medio (Root Mean Square Error)
ROC	Característica operativa del receptor (Receiver Operating Characteristic)
S3	Simple Storage Service (Amazon)
SaaS	Software como servicio (Software as a Service)
SCA	Autenticación reforzada del cliente (Strong Customer Authentication)
SHAP	Explicaciones aditivas de Shapley (SHapley Additive exPlanations)
SQL	Lenguaje de consulta estructurado (Structured Query Language)
TFG	Trabajo Fin de Grado
TLS	Seguridad de la capa de transporte (Transport Layer Security)
TPP	Proveedor de servicios de terceros (Third Party Provider)
VaR	Valor en riesgo (Value at Risk)
WACC	Coste medio ponderado del capital (Weighted Average Cost of Capital)

Capítulo 1. INTRODUCCIÓN

Las finanzas personales actualmente están inmersas en una enorme paradoja: la gente genera más datos financieros que nunca y, sin embargo, los entiende cada vez menos o los gestiona casi a ciegas. La digitalización ha avanzado más rápido que la educación financiera. La mayoría de la gente tiene aplicaciones y herramientas para llevar un seguimiento de los gastos, pero sigue teniendo solo una idea aproximada de en qué se gasta su dinero diariamente. Más datos, más acceso, pero el gran desafío silencioso de nuestra generación es comprender realmente nuestras cuentas.

Es en este contexto donde nace PREFIN, una plataforma inteligente de prevención, predicción y recomendación financiera con técnicas de aprendizaje automático. Consiste en la combinación de técnicas de aprendizaje automático con un sistema de simulación basado en un gemelo digital financiero. Su propósito no es sustituir la educación financiera, muchas veces escasa, sino complementarla: acompañar al usuario, generar conciencia y traducir sus datos en decisiones más informadas, accesibles y conscientes.

Este capítulo presenta el marco que justifica la existencia de este proyecto. Se analiza el contexto socioeconómico que hace necesaria una herramienta de estas características, se expone la motivación que ha guiado su desarrollo, se enuncian los objetivos que se pretende abordar y se describe brevemente la estructura del documento.

1.1 CONTEXTO SOCIOECONÓMICO Y EDUCACIÓN FINANCIERA

La gestión del dinero es una de las virtudes más importantes a lo largo de la vida de cualquier persona, sin embargo, es una de las que menos se enseña en colegios y universidades. Nadie nace sabiendo hacer presupuestos, idear un plan de ahorros, las condiciones de un plan de pensiones o hacer la declaración de la renta. En definitiva, la educación financiera es una competencia necesaria y si no se adquiere a tiempo, sus consecuencias se arrastran durante décadas.

Las decisiones relacionadas con las finanzas personales tomadas en las primeras etapas de la vida tienen un gran impacto sobre la riqueza posterior y la estabilidad económica a lo largo de la vida (Lusardi y Mitchell, 2014). La falta de conocimiento financiero en los más jóvenes es sinónimo de peores decisiones de inversión, mayores deudas y menor planificación de la jubilación. Todo ello tiene consecuencias en las generaciones posteriores.

La situación económica actual no hace más que agravar esta problemática. Durante los últimos diez años, los salarios han crecido más lentamente que la inflación de la mayoría de los países europeos (Eurostat, 2023), esta situación ha propiciado una mayor presión económica sobre los hogares y reclama un mayor control y planificación. Asimismo, la irrupción de nuevos productos financieros digitales, como pueden ser, el crédito rápido o el modelo Buy Now and Pay Later (BNPL), cuyos usos se dispararon tras la pandemia, ha ampliado el acceso y la exposición al endeudamiento sin comprensión del riesgo asumido, ya que esta última, no ha crecido al mismo ritmo. Según la OCDE, únicamente el 28% de los adultos que utilizan crédito digital comprende conceptos fundamentales como el interés simple y compuesto, y una gran proporción desconoce las comisiones y condiciones asociadas a los productos que contrata (OCDE, 2023).

La banca digital ha multiplicado los datos disponibles para el usuario, pero no su capacidad de interpretarlos. Más del 90% de los clientes emplean la banca móvil de forma habitual, raro será conocer a alguien que actualmente no pague o gestione sus cuentas desde el móvil, pero solo una minoría son los que revisan y analizan sus patrones de gasto con regularidad (Deloitte, 2022). Tener fácil acceso a los datos no equivale a comprenderlos ni a actuar en consecuencia. Varios estudios revelan que la mayoría de los usuarios subestiman sus gastos variables entre un 20% y un 30%, y que los jóvenes utilizan el dinero sin comprender cómo administrarlo con responsabilidad o cómo se genera.

La consecuencia de todo ello es la ironía que define nuestra época: más información financiera que ayer, menos que mañana, pero no disponemos de las mejores herramientas para entender su contexto y poder convertirla en decisiones acertadas. La planificación financiera, entendida como el procedimiento de identificar objetivos y diseñar un camino para alcanzarlos, combinando presupuestos, ahorro, gestión de deuda e inversión, sigue

siendo una práctica reservada para aquellos afortunados con acceso a un asesor profesional o una formación. Para la mayoría de ciudadanos, ese acompañamiento y orientación brilla por su ausencia. Es precisamente en esta brecha donde se puede enmarcar el siguiente proyecto.

1.2 MOTIVACIÓN DEL PROYECTO

La motivación del presente proyecto nace de dos dimensiones que, a pesar de poder parecer contradictorias, se refuerzan una a la otra. Una dimensión personal, nacida de la experiencia vivida, y una dimensión técnica, surgida del análisis de la situación actual y las soluciones existentes.

Desde el punto de vista personal, el interés por la educación financiera se debe a la propia trayectoria vital. Una situación económica familiar complicada permitió comprender en primera persona la delicadeza de la seguridad financiera, la importancia de un plan de rescate o, por el contrario, el enorme coste de no disponer de herramientas ni conocimientos para anticipar una crisis.

Esta experiencia fue la causa por la que nació un interés por seguir a autores como Dave Ramsey e indagar en su método de los Baby Steps, siete pasos orientados al ahorro, eliminación de las deudas y impulso para la inversión. Él se encarga de demostrar que la estabilidad económica no se trata de una cuestión de suerte o del nivel de ingresos de cada uno, sino de educación y hábitos (Ramsey, 2013). Parte de su filosofía se ha utilizado como inspiración para este proyecto, no se trata de desarrollar una calculadora exclusiva y sofisticada, sino de acompañar al usuario en tomar decisiones con conciencia.

Este vacío en la educación financiera no es algo universal. En países como Estados Unidos, la planificación financiera personal forma parte de las opciones impartidas en las universidades. Por ejemplo, en la Universidad de Illinois se ofrece Personal Financial Planning, asignatura en la que se navega por conceptos como planificar presupuestos o gestión de deuda hasta principios esenciales como la regla de los tres destinos del dinero: ahorrar (save), gastar (spend) y donar (give). Los tres con igual de importancia estructuran

una relación equilibrada y consciente de las finanzas personales. En España, sin embargo, este tipo de educación sigue siendo una asignatura pendiente y deficiente, si atendemos a datos históricos se sitúa por debajo de la media europea y de la OCDE, lo cual deja a la mayoría de los ciudadanos sin herramientas básicas para gestionar su economía (Banco de España y CNMV, 2018). Aunque el acceso bancario es casi universal, gran parte de la población no llega a comprender conceptos clave como la inflación, el interés compuesto o la diversificación de riesgos. Según datos del Banco de España, apenas el 19% de la población adulta responde correctamente a preguntas básicas sobre finanzas personales. Actualmente no existe una asignatura específica obligatoria de educación financiera en la etapa escolar; el sistema educativo (LOMLOE) opta por integrarla de manera transversal.

Desde el punto de vista técnico, la inspiración, como ya se ha mencionado previamente, surge de una necesidad detectada al analizar las herramientas de administración financiera personal que están a día de hoy en el mercado. Tal y como se explica en el *Capítulo 2*, las alternativas de las que se dispone pueden mostrar una imagen de la situación al usuario, clasificar gastos, mostrar transacciones e identificar suscripciones, pero ninguna de ellas integra el análisis predictivo, prevención de riesgos y simulación de escenarios en una misma plataforma gratis y accesible a todos. Muchas de las aplicaciones existentes ofrecen funciones a cambio de una suscripción que varía entre los 8 y 15 euros mensuales, lo que genera una contradicción clara: aquellos que requieren de ayuda son los que no pueden pagar por estas herramientas (OCDE, 2023).

En esta brecha educativa y tecnológica es donde se encuentra PREFIN. Una plataforma que da un paso más no descubre solamente el comportamiento financiero, sino que busca interpretarlo, proyectarlo y ayudar al usuario. Propone técnicas de aprendizaje automático y un sistema de simulación basado en un gemelo digital financiero. Si se piensa que su meta es reemplazar la educación financiera, nada más lejos de la realidad, lo que se pretende es hacerla accesible sin importar el nivel de conocimientos o la condición económica.

1.3 ESTRUCTURA DE LA MEMORIA

Este documento consta de ocho capítulos, que siguen el orden lógico del proyecto, desde el problema que lo motivó hasta el sistema desarrollado, sus resultados y sus conclusiones.

El ***Capítulo 1*** sitúa el contexto de PREFIN: por qué existe, qué problema aborda y la motivación que lo impulsa.

El ***Capítulo 2*** describe las tecnologías, los protocolos y las herramientas en los que se apoya el proyecto, desde la banca abierta y el protocolo OAuth 2.0 hasta los frameworks y algoritmos empleados.

El ***Capítulo 3*** recoge el estado de la cuestión, es decir, las soluciones que existen actualmente en la gestión financiera personal, e identifica el vacío tecnológico que PREFIN pretende cubrir.

El ***Capítulo 4*** define el trabajo: la justificación de la propuesta, los objetivos, la metodología seguida y el estudio económico.

El ***Capítulo 5*** constituye el núcleo técnico de la memoria. Describe la arquitectura de PREFIN y cada uno de sus módulos, desde la ingesta de datos hasta las capas de inteligencia que permiten predecir, recomendar y explicar.

El ***Capítulo 6*** presenta el análisis de resultados, con la evaluación del rendimiento real de los modelos y la validación del sistema.

El ***Capítulo 7*** recoge las conclusiones, una evaluación de en qué medida se han cumplido los objetivos iniciales, y propone líneas de trabajo futuro.

Finalmente, el ***Capítulo 8*** reúne la bibliografía, seguido de los anexos, entre ellos el de alineación del proyecto con los Objetivos de Desarrollo Sostenible.

Capítulo 2. DESCRIPCIÓN DE LAS TECNOLOGÍAS

Este capítulo describe las tecnologías, los protocolos y las herramientas específicas sobre las que se sustenta PREFIN, es decir, aquellos cuya comprensión resulta imprescindible para seguir las decisiones de diseño e implementación que se exponen más adelante. Se omiten deliberadamente las herramientas estándar de ingeniería de software, como el lenguaje Python o las bibliotecas habituales de análisis de datos, cuyo conocimiento se presupone en el lector.

2.1 OPEN BANKING Y LA DIRECTIVA PSD2

La banca abierta es el marco normativo y técnico que permite a los proveedores de servicios financieros autorizados acceder, con el consentimiento explícito del usuario, a los datos bancarios custodiados por las entidades. Su finalidad es romper el acceso exclusivo que tradicionalmente han tenido los bancos sobre la información financiera de sus clientes, fomentando así la competencia y la innovación en el sector.

En el ámbito europeo, este marco queda regulado por la Directiva (UE) 2015/2366, conocida como PSD2, aprobada por el Parlamento Europeo en 2015 y de obligado cumplimiento en todos los Estados miembros (Parlamento Europeo y Consejo de la Unión Europea, 2015). La directiva obliga a las entidades bancarias a exponer los datos de sus clientes a través de interfaces de programación (API) estandarizadas y seguras, y exige que todo acceso por parte de terceros se someta a un proceso de Autenticación Reforzada del Cliente (SCA). Los proveedores autorizados a operar bajo esta directiva se denominan proveedores de servicios de terceros (TPP) y deben estar registrados ante el regulador nacional correspondiente.

Para PREFIN, la PSD2 constituye el fundamento que hace legal y técnicamente posible la conexión con los datos bancarios reales de un usuario sin necesidad de solicitarle en ningún momento las credenciales de acceso a su banco.

2.2 OAUTH 2.0

OAuth 2.0 es el protocolo de autorización que emplea PREFERIN para acceder a los datos bancarios del usuario a través de TrueLayer. Su diferencia esencial respecto a los sistemas de autenticación tradicionales reside en que el usuario nunca facilita su contraseña a la aplicación: la autorización se concede directamente desde el banco, que es quien emite un token de acceso para que la aplicación pueda realizar consultas en su nombre (Hardt, 2012).

El flujo implementado en PREFERIN es el denominado Authorization Code Flow. El usuario es redirigido al portal de TrueLayer, donde selecciona su entidad y otorga su consentimiento; TrueLayer devuelve entonces un código de autorización que el backend de PREFERIN intercambia por un token de acceso y otro de refresco. Ambos se almacenan de forma segura en la base de datos local y se emplean, a partir de ese momento, para consultar las transacciones sin que el usuario deba autenticarse de nuevo cada vez que utiliza la aplicación.

2.3 TRUELAYER

TrueLayer es un proveedor de infraestructura de banca abierta con sede en Londres, fundado en 2016, que opera como TPP registrado bajo la PSD2 y cumple, por tanto, los requisitos regulatorios para el tratamiento de datos bancarios de terceros. Su propuesta consiste en unificar las API de numerosas entidades europeas en una única interfaz estándar, de modo que el desarrollador no necesita integrarse por separado con cada banco.

PREFERIN se conecta con TrueLayer para obtener el historial de transacciones del usuario. La principal ventaja de esta intermediación es que, con independencia de la entidad de origen, todas las transacciones se reciben con la misma estructura —fecha, descripción, importe y divisa—, lo que reduce considerablemente el trabajo de normalización. Durante el desarrollo, toda la integración se validó con el entorno de pruebas (sandbox) de TrueLayer, que replica el comportamiento real de la API mediante datos de prueba, permitiendo verificar tanto la autenticación como el procesamiento de datos sin involucrar cuentas bancarias reales.

2.4 DASH

Dash es un conjunto de herramientas de código abierto que permite construir aplicaciones web interactivas íntegramente en Python, sin necesidad de programar en JavaScript, HTML o CSS. Internamente combina el servidor web Flask con componentes que se renderizan en el navegador, generando interfaces dinámicas que responden a las acciones del usuario a partir de instrucciones definidas en Python.

Su elección en PREFIN responde a tres motivos. En primer lugar, permite integrar en una misma aplicación el procesamiento de datos, los modelos de aprendizaje automático y la interfaz de usuario, sin cambiar de lenguaje ni de entorno de trabajo. En segundo lugar, su compatibilidad con Plotly facilita la creación de visualizaciones interactivas complejas con poco código. Por último, su integración con Dash Bootstrap Components permite aplicar el sistema de diseño de Bootstrap 5 y obtener una interfaz adaptable a distintos dispositivos y tamaños de pantalla sin necesidad de escribir CSS personalizado más allá de los ajustes estéticos propios de la aplicación.

2.5 FASTAPI

FastAPI es un framework moderno para la creación de API REST en Python, caracterizado por su alto rendimiento y por la validación automática de los datos. En PREFIN se emplea como backend independiente encargado de la comunicación con TrueLayer: implementa el flujo OAuth 2.0, almacena y refresca los tokens de acceso, y expone los endpoints que la interfaz Dash utiliza para obtener las transacciones bancarias del usuario.

La aplicación se estructura en dos procesos: el frontend, desarrollado con Dash y ejecutado en el puerto 8050, y el backend, basado en FastAPI y ejecutado en el puerto 8000. Esta separación aísla la lógica de autenticación y la comunicación con la API bancaria del resto de la aplicación, lo que facilita el mantenimiento, mejora la seguridad al centralizar la gestión de credenciales y permite que la interfaz funcione de forma autónoma cuando se emplean fuentes de datos alternativas, como archivos CSV o datos sintéticos. El servicio se ejecuta sobre uvicorn, un servidor de alto rendimiento orientado a aplicaciones Python asíncronas.

2.6 *SQLITE*

SQLite es un sistema de gestión de bases de datos relacional y ligero que almacena toda la información en un único archivo local, sin requerir un servidor de base de datos independiente. En PREFIN cumple una función muy concreta: conservar los tokens OAuth 2.0 entre peticiones al backend, de modo que el usuario no deba autenticarse de nuevo con su banco en cada consulta realizada dentro de una misma sesión. Se optó por SQLite frente a otras alternativas porque el caso de uso es reducido y específico —escaso volumen de datos y baja concurrencia— y porque forma parte de la biblioteca estándar de Python, lo que evita cualquier instalación o configuración adicional.

2.7 *RANDOM FOREST Y GRADIENT BOOSTING*

Los modelos de aprendizaje automático empleados en PREFIN pertenecen a la familia de los métodos de conjunto (ensemble), que combinan varios modelos simples para obtener predicciones más precisas y robustas que las que ofrecería cualquiera de ellos por separado.

El Random Forest construye numerosos árboles de decisión de forma simultánea, entrenando cada uno con una muestra aleatoria de los datos y un subconjunto aleatorio de las variables, y combina sus salidas por votación. Este procedimiento reduce notablemente el riesgo de sobreajuste y mejora la capacidad de generalización del modelo. En PREFIN, el Random Forest se emplea para clasificar el riesgo financiero del usuario a partir de un conjunto de variables derivadas de su comportamiento; la definición concreta de la etiqueta de riesgo y la validación del modelo se detallan en el *Capítulo 5*. Una ventaja adicional de este método es que permite cuantificar la importancia de cada variable en la predicción, lo que posibilita comunicar al usuario los factores que más contribuyen a su nivel de riesgo.

El Gradient Boosting, por su parte, construye los árboles de forma secuencial, entrenando cada nuevo árbol para corregir los errores cometidos por el conjunto anterior, lo que da lugar a modelos de gran capacidad predictiva. En PREFIN, esta técnica se emplea en la previsión del gasto futuro, configurada para regresión cuantílica: en lugar de estimar un único valor, proporciona directamente los percentiles que conforman una banda de incertidumbre, lo que

permite ofrecer al usuario un rango probable de gasto en vez de una cifra aislada. El detalle de su aplicación se expone en el *Capítulo 5*.

Ambos modelos se entrenan a partir de los datos del propio usuario, lo que garantiza que las predicciones sean específicas para su situación financiera y no respondan a un comportamiento genérico.

2.8 GEMELO DIGITAL

El concepto de gemelo digital designa una réplica virtual de un sistema real que se actualiza de forma continua con datos procedentes de este, permitiendo simular su comportamiento ante distintas situaciones sin necesidad de intervenir sobre el sistema físico. Fue propuesto en 2017 por Negri, Fumagalli y Macchi y, aunque inicialmente se aplicó a la gestión y supervisión de maquinaria industrial, su utilidad se ha extendido posteriormente a otros ámbitos en los que resulta valioso anticipar el comportamiento de un sistema complejo antes de tomar decisiones que lo afecten.

En PREFIN, el gemelo digital financiero reproduce el estado económico del usuario a partir de sus ingresos medios, su gasto por categoría, su saldo acumulado y su tasa de ahorro, y proyecta su evolución a lo largo de un horizonte de hasta 36 meses. El usuario puede personalizar la simulación modificando el porcentaje de ingresos, ajustando el gasto de categorías concretas, introduciendo un gasto imprevisto puntual o fijando una meta de ahorro mensual.

A diferencia de una proyección determinista, el gemelo de PREFIN es de naturaleza estocástica: en lugar de trazar una única trayectoria del saldo, emplea simulación de Monte Carlo para generar miles de evoluciones posibles, cada una con variaciones realistas. Ello permite no solo estimar un resultado esperado, sino cuantificar la incertidumbre asociada, la franja de valores probables, la probabilidad de quedarse sin liquidez en cada mes y los escenarios más desfavorables, de modo que el usuario puede valorar el impacto de sus decisiones antes de tomarlas. Su funcionamiento detallado se describe en el *Capítulo 5*.

Capítulo 3. ESTADO DE LA CUESTIÓN

Este capítulo analiza las soluciones tecnológicas existentes en el ámbito de la gestión financiera personal con el fin de responder a una pregunta fundamental: ¿existe en el mercado alguna solución equivalente a la que se propone? Para ello se examinan tres categorías de herramientas: las aplicaciones de gestión financiera personal, las plataformas de inversión automatizada y las soluciones basadas en inteligencia artificial y análisis avanzado, evaluando en cada caso sus capacidades, sus limitaciones y las necesidades que dejan sin cubrir. El capítulo concluye con un análisis crítico del conjunto que justifica la necesidad de una propuesta diferenciada como PREFIN.

3.1 EL ECOSISTEMA FINTECH EN LA GESTIÓN FINANCIERA PERSONAL

Durante los últimos años, el ecosistema fintech ha experimentado un crecimiento notable a escala mundial, impulsado por la penetración de los dispositivos móviles, la generalización de la banca digital y, en el ámbito europeo, por la entrada en vigor de la Directiva PSD2. Esta evolución ha multiplicado las herramientas disponibles para la gestión del dinero y ha puesto al alcance del usuario medio servicios que antes quedaban reservados a quienes disponían de patrimonio elevado o de asesoramiento profesional.

Funcionalidades que hace una década resultaban innovadoras, como la categorización automática del gasto o la simulación de productos financieros, se han convertido hoy en prestaciones habituales, incorporadas incluso por la banca tradicional en sus aplicaciones móviles. El problema, por tanto, ha dejado de ser la ausencia de herramientas para pasar a ser su insuficiencia: la mayoría se limita a describir la situación financiera del usuario, sin ayudarle a comprenderla, a planificar el futuro ni a transformarla de forma activa. Es en este contexto donde se analiza el estado actual de las soluciones, agrupadas en las tres categorías mencionadas.

3.2 APLICACIONES DE GESTIÓN FINANCIERA PERSONAL (PFM)

Las aplicaciones de gestión financiera personal (Personal Finance Management, PFM) son herramientas ampliamente extendidas que ayudan al usuario a comprender mejor sus ingresos y gastos, agregando información de diversas fuentes y presentándola de forma clara para ofrecer una visión general de su situación económica. Entre las soluciones más conocidas en España y Europa se encuentran Fintonic, Emma, Plum, Monefy y YNAB. Fintonic permite agregar cuentas, tarjetas y seguros, y ofrece recomendaciones de productos financieros; Emma actúa como un asistente que detecta suscripciones y comisiones ocultas; Plum combina la agregación de cuentas con reglas de automatización para ahorrar pequeñas cantidades de forma periódica; Monefy proporciona una visión clara del presupuesto personal, si bien exige la introducción manual de los datos; y YNAB se basa en una filosofía de presupuestación activa que asigna un propósito concreto a cada unidad monetaria disponible.

Pese a su popularidad, estas herramientas comparten una limitación de fondo: muchas de las funcionalidades que en su día constituyeron una novedad han sido incorporadas por las propias entidades bancarias en sus aplicaciones, lo que ha erosionado su ventaja competitiva y ha dejado al descubierto una carencia esencial. Ninguna de ellas es capaz de ir más allá de la descripción del comportamiento financiero para interpretarlo, anticiparlo y ayudar al usuario a transformarlo.

3.3 PLATAFORMAS DE INVERSIÓN AUTOMATIZADA: LOS ROBO-ADVISORS

Los robo-advisors son plataformas digitales que gestionan inversiones de forma automática, empleando algoritmos para ofrecer carteras personalizadas según el perfil de riesgo, el horizonte temporal y los objetivos financieros de cada usuario. Al prescindir de un gestor

humano, reducen los costes y democratizan el acceso a un asesoramiento que tradicionalmente resultaba caro y reservado a unos pocos.

Entre las plataformas más relevantes a escala internacional destacan Betterment y Wealthfront en Estados Unidos, pioneras del sector y con millones de usuarios. Betterment ofrece carteras diversificadas de fondos indexados con reajuste automático y optimización fiscal, mientras que Wealthfront incorpora funcionalidades adicionales como la planificación financiera integral y el análisis de liquidez. En el ámbito europeo y español sobresalen Indexa Capital, líder en gestión de inversiones en España, y Finizens, orientada a la inversión a largo plazo con carteras de bajo coste.

No obstante, el enfoque de los robo-advisors difiere sustancialmente del que aquí se persigue: se centran exclusivamente en la inversión y no consideran el comportamiento financiero diario del usuario, su gasto ni su situación de liquidez mensual. No analizan las transacciones bancarias, no detectan hábitos de consumo perjudiciales ni emiten alertas de riesgo financiero a corto plazo. Son, en definitiva, herramientas para la gestión del patrimonio a largo plazo y no para la prevención de problemas financieros personales. Esta diferencia es significativa, pues deja sin atender a un colectivo amplio: el de las personas que, sin capacidad de inversión, necesitan ayuda para comprender y mejorar su gestión financiera cotidiana.

3.4 HERRAMIENTAS CON INTELIGENCIA ARTIFICIAL Y ANÁLISIS AVANZADO

La irrupción de la inteligencia artificial ha transformado también la gestión de las finanzas personales. Frente a las herramientas tradicionales, limitadas a mostrar datos, han surgido soluciones capaces de analizar el comportamiento del usuario y ofrecer recomendaciones personalizadas. Entre las más destacadas se encuentra Cleo, un asistente conversacional que dialoga con el usuario en lenguaje natural, analiza sus gastos y aconseja sobre su presupuesto, así como Plaid, que permite a otras aplicaciones acceder a la información

bancaria para ofrecer análisis avanzados. Otras plataformas, como la desaparecida Mint, llegaron a predecir el gasto futuro a partir del historial del usuario.

Sin embargo, estas herramientas presentan limitaciones relevantes. No existe una plataforma que integre en un único entorno el análisis del gasto, la predicción del riesgo, la simulación de escenarios y el fomento del ahorro; cada solución cubre una parte del problema, lo que obliga al usuario a combinar varias aplicaciones para obtener una visión completa de sus finanzas. A ello se añade que las recomendaciones ofrecidas suelen ser genéricas y poco ajustadas a la situación real y cambiante de cada persona y que, en su mayoría, se trata de servicios de pago, lo que las hace inaccesibles para quienes más las necesitarían.

3.5 LIMITACIONES ECONÓMICAS DE LAS SOLUCIONES ACTUALES

Más allá de las carencias funcionales, las soluciones existentes presentan una limitación económica determinante: el coste de acceso. La mayoría operan bajo un modelo de suscripción mensual que oscila entre los 8 y los 15 euros, como ocurre con YNAB, Wealthfront, Tiller o Plum, cuyas funcionalidades de mayor valor quedan tras una cuota de pago.

Esta circunstancia genera una paradoja difícil de ignorar: las personas que más necesitan ayuda para mejorar su situación financiera son, con frecuencia, las que no pueden costear las herramientas diseñadas para asistirles. Quien tiene dificultades para llegar a fin de mes, ahorra poco o mantiene hábitos de consumo desequilibrados es precisamente el perfil que más se beneficiaría de una plataforma de análisis y prevención financiera y, al mismo tiempo, el menos dispuesto a asumir un coste mensual adicional. Por otra parte, las soluciones gratuitas disponibles suelen financiarse mediante la venta de datos del usuario o la recomendación de productos financieros de terceros, lo que compromete la neutralidad de sus consejos. El usuario queda así atrapado entre herramientas de pago que no puede asumir y herramientas gratuitas que no velan por sus intereses.

3.6 ANTECEDENTES EN LA LITERATURA TÉCNICA

El vacío que aborda PREFIN no es solo de mercado, sino también metodológico, y conviene situarlo frente al estado del arte científico de las técnicas empleadas. La previsión de gasto se apoya en la regresión cuantílica, formalizada por Koenker y Bassett (1978), que estima percentiles condicionales en lugar de un único valor esperado y permite construir bandas de incertidumbre. Esta previsión se combina con una referencia estacional siguiendo el principio de combinación de predicciones de Bates y Granger (1969), según el cual la agregación de estimadores independientes reduce el error frente a cualquiera de ellos por separado. La fiabilidad de las bandas se enmarca en la cuantificación de incertidumbre sin distribución, cuyo exponente actual es la predicción conforme (Angelopoulos y Bates, 2023), técnica que se adopta como vía de calibración. El gemelo digital parte del concepto introducido por Negri, Fumagalli y Macchi (2017), aquí trasladado del dominio industrial al financiero. Por último, el salto de la predicción a la recomendación se inscribe en el marco de la analítica prescriptiva descrito por Bertsimas y Kallus (2020), que articula cómo transformar modelos predictivos en decisiones óptimas bajo restricciones. Ninguna de las soluciones comerciales analizadas integra estas líneas en un único sistema, lo que confirma que la aportación de PREFIN es también de naturaleza técnica

3.7 VACÍO TECNOLÓGICO Y NECESIDAD DE PREFIN

El análisis de las soluciones existentes revela una carencia clara que, pese a los avances del sector fintech, ninguna ha resuelto de forma satisfactoria. Las aplicaciones de gestión financiera personal describen, pero no predicen; los robo-advisors gestionan el patrimonio, pero ignoran el comportamiento financiero diario; y las herramientas basadas en inteligencia artificial ofrecen recomendaciones, pero no integran análisis, predicción y simulación en un mismo entorno. A todo ello se suma que las soluciones más avanzadas resultan, en su mayoría, de pago, lo que excluye precisamente a los usuarios que más las necesitan.

El vacío es, por tanto, evidente: no existe una plataforma gratuita y accesible que combine la extracción de datos bancarios, el análisis del comportamiento financiero, la predicción del

riesgo mediante aprendizaje automático, la simulación de escenarios futuros y el fomento del ahorro mediante micro-contribuciones automáticas. Es este vacío el que PREFIN se propone cubrir. Su finalidad no es competir con los robo-advisors ni sustituir a los profesionales, sino acompañar al usuario en su vida financiera cotidiana, convertir sus datos en conocimiento útil y accionable y ayudarle a tomar mejores decisiones antes de que las dificultades se conviertan en crisis, poniendo al alcance de cualquier persona aquello que hasta ahora solo estaba reservado a quienes podían pagar un asesor o una herramienta de pago.

Capítulo 4. DEFINICIÓN DEL TRABAJO

El capítulo anterior mostró que, a pesar de que existen numerosas herramientas para gestionar las finanzas personales, ninguna de ellas resuelve el problema de prevenir y predecir las dificultades financieras de manera accesible. Partiendo de esa conclusión, este capítulo justifica por qué se ha desarrollado PREFIN, qué alcance cubre, qué metodología se ha seguido y cuál ha sido su coste en tiempo y recursos.

4.1 JUSTIFICACIÓN

La revisión de lo que ya se ha hecho en este campo, que se presentó en el *Capítulo 3*, permite identificar con claridad qué es lo que PREFIN puede ofrecer de nuevo. Actualmente, no hay ninguna plataforma que combine de forma gratuita y en un solo lugar la conexión con datos bancarios reales, el análisis automático del comportamiento financiero, la predicción del riesgo, la simulación de escenarios futuros y el fomento del ahorro. Cada una de estas cosas se puede encontrar por separado en diferentes herramientas, pero no hay ninguna que las integre de manera coherente y accesible.

Desde un punto de vista técnico, ahora es un buen momento para desarrollar una solución como PREFIN gracias a los avances en el aprendizaje automático y la normalización del open banking. Las APIs de open banking permiten acceder a datos bancarios reales de forma segura y estandarizada. Los algoritmos de clasificación y regresión que se pueden encontrar en librerías como scikit-learn son suficientes para predecir el comportamiento financiero. Y frameworks como Dash permiten construir interfaces web complejas sin necesidad de un equipo de desarrollo muy grande.

Desde un punto de vista de mercado, PREFIN se dirige a un grupo de personas que no tiene mucho dinero para invertir, que no puede pagar por una herramienta de gestión financiera y que necesita ayuda para mejorar su situación financiera diaria. En España, donde no se

enseña educación financiera en la escuela, este perfil es el de la mayoría de los jóvenes que trabajan.

4.1.1 PROPUESTA DE VALOR DIFERENCIAL

PREFIN destaca y se diferencia entre las aplicaciones ya existentes en tres dimensiones. En primer lugar, es la única que integra y combina en un único entorno el análisis descriptivo, la predicción de riesgo, la simulación de posibles escenarios y el micro-ahorro. De esta manera, permite eliminar la necesidad de utilizar más de una herramienta para obtener la visión 360 de la situación financiera del usuario. En segundo lugar, sus modelos de predicción se entrenan con datos reales de la persona, lo que garantiza que el trato es personalizado y no genérico. Por último, es gratis, no tiene tarifas de suscripción ni publicita productos financieros de terceros, de esta manera sus recomendaciones son fiables y elimina las barreras económicas para que todo el que quiera pueda acceder.

4.1.2 VIABILIDAD TÉCNICA Y ACADÉMICA

PREFIN es un proyecto realista para un Trabajo Fin de Grado de ingeniería, y su viabilidad puede analizarse en dos perspectivas: la técnica, las tecnologías y recursos usados para su desarrollo y la académica, relacionada con los objetivos y plazos de un TFG.

El proyecto recoge varios ámbitos de ingeniería: ingeniería de datos a la hora de extraer, limpiar y normalizar; machine learning para la clasificación de riesgo y predicción de gasto; arquitectura de software para la separación de los módulos de frontend y backend; y el diseño de una interfaz para presentar la información de forma clara. Esto lo convierte en un proyecto completo que recoge varias competencias presentes en el plan de estudios, en lugar de centrarse en un solo ámbito.

Además, resuelve un problema concreto y medible. No se trata de un ejercicio teórico: hay datos que entran, modelos que se entrenan, predicciones que se pueden evaluar con métricas objetivas y una interfaz funcional con la que un usuario real puede interactuar. Cada módulo tiene criterios de validación claros, lo que permite evaluar los resultados con rigor.

El alcance está bien definido y es asumible dentro de las restricciones de un TFG. El entorno sandbox de TrueLayer no es solo una comodidad, es lo que hace viable la integración con open banking sin depender de acuerdos con bancos reales, datos sensibles de usuarios ni autorizaciones regulatorias que serían imposibles de conseguir en los plazos disponibles. El flujo completo de autenticación y obtención de datos se puede probar en un entorno controlado, sin bloqueos externos.

Por último, la arquitectura modular reduce el riesgo del proyecto. Cada componente — ingesta de datos, análisis, predicción, simulación, interfaz — funciona de forma independiente y se puede desarrollar y validar por separado. Si un módulo da problemas, no arrastra al resto. Eso marca la diferencia entre un proyecto que se puede terminar en plazo y uno que se queda atascado porque todo depende de todo.

4.2 OBJETIVOS

PREFIN es una plataforma de predicción y prevención financiera. Estos son los objetivos que la han definido.

4.2.1 ANÁLISIS DE DATOS FINANCIEROS

El sistema recolecta los datos bancarios de manera segura a través de una API de banca abierta y procesa las transacciones de diferentes cuentas para hacerlas comprensibles y consistentes.

Sobre estos datos, se aplican técnicas de análisis para descubrir: en qué gasta el dinero el usuario, gastos repetidos, gastos inesperados, tendencias de gasto con el tiempo.

De esta manera, se construye un perfil financiero personalizado. Este perfil se basa en la clasificación automática de las transacciones del usuario y en indicadores financieros mensuales.

Para asegurar la seguridad y cumplir con las normas, como la PSD2, se ha probado la integración con el entorno de pruebas (sandbox) de TrueLayer.

4.2.2 EVALUACIÓN DE RIESGO, SIMULACIÓN Y RECOMENDACIÓN PRESCRIPTIVA

Se emplea el aprendizaje automático para evaluar el riesgo financiero y detectar cambios significativos en el gasto antes de que se conviertan en un problema. Se construye, además, un gemelo digital de la situación financiera que permite simular escenarios futuros a un horizonte de hasta 36 meses. Sobre ese gemelo se desarrolla un motor de recomendación que, explorando las distintas decisiones posibles (ahorro, recortes por categoría, redondeo), propone al usuario el mejor plan de acción manteniendo el riesgo bajo control. Esto supone el salto de la analítica predictiva a la prescriptiva y constituye la principal aportación del proyecto, al habilitar la toma de decisiones informadas.

4.2.3 INTERFAZ ACCESIBLE PARA TODOS

Se desea que la plataforma sea fácil de usar para cualquier persona, sin importar si tiene experiencia en gestión financiera o no.

La interfaz web muestra los análisis, las predicciones y las simulaciones de manera clara y sencilla. La facilidad de uso constituye una prioridad: tanto un usuario que se inicia en la gestión de su presupuesto como uno experimentado deben encontrar la plataforma igualmente útil.

4.3 *METODOLOGÍA*

El desarrollo de PREFIN siguió una metodología estructurada en seis fases, cada una con objetivos definidos cuya consecución habilitaba el inicio de la siguiente. Este enfoque modular, en el que cada componente puede desarrollarse y validarse de forma independiente, permitió controlar el riesgo del proyecto y avanzar de manera ordenada desde la definición del problema hasta el sistema completo.

La primera fase consistió en el análisis del problema y la definición de los requisitos. Se revisaron las herramientas de gestión financiera personal disponibles en el mercado, se identificaron sus carencias y se determinaron las funcionalidades que debía ofrecer el

sistema, incluyendo los casos de uso principales y las restricciones derivadas de la normativa PSD2.

La segunda fase se dedicó a la obtención, organización y preparación de los datos. Se desarrolló la conexión con TrueLayer mediante OAuth 2.0, se habilitó la carga de extractos en formato CSV y Excel y se construyó un generador de datos sintéticos para las pruebas. Asimismo, se implementó el proceso de limpieza y normalización de las transacciones.

La tercera fase abordó el análisis y el modelado del comportamiento financiero. Se desarrolló la categorización automática de transacciones, se definieron los indicadores para el análisis mensual y se implementó el módulo de detección de anomalías y cambios de régimen.

La cuarta fase se centró en el desarrollo de los modelos predictivos. Se definieron las variables financieras relevantes, se estableció el criterio de etiquetado del riesgo como evento futuro observado, y se entrenaron y validaron el clasificador de riesgo de iliquidez y el modelo de previsión de gasto con incertidumbre, junto con el módulo de explicabilidad encargado de traducir las predicciones a lenguaje natural.

La quinta fase comprendió la construcción del gemelo digital y de la capa de predicción. Se desarrolló el gemelo estocástico, que mediante simulación de Monte Carlo proyecta la evolución del saldo a lo largo de un horizonte de hasta 36 meses cuantificando la incertidumbre, y, sobre él, el motor de recomendación, que explora las distintas decisiones posibles y propone al usuario el plan de acción óptimo bajo una restricción de riesgo.

La sexta y última fase consistió en el diseño e implementación de la interfaz gráfica y en la prueba del sistema completo. La aplicación se organizó en siete pestañas —Dashboard, Análisis, Riesgo, Gemelo Digital, Mi Plan, Micro-Ahorro y Datos—, siendo la pestaña Mi Plan la correspondiente al motor de recomendación prescriptivo. Finalmente, se integraron todos los módulos en un flujo de usuario coherente y se realizaron pruebas tanto con datos sintéticos como con el entorno de pruebas de TrueLayer.

4.4 PLANIFICACIÓN Y ESTIMACIÓN ECONÓMICA

4.4.1 PLANIFICACIÓN TEMPORAL

El desarrollo del proyecto se ha llevado a cabo entre septiembre de 2025 y junio de 2026, con una dedicación estimada de 300 horas totales distribuidas según el cronograma recogido en la Tabla 4.1.

Fase	Descripción	Duración
1	Análisis del problema y definición de requisitos	3 semanas
2	Obtención, normalización y preparación de datos	4 semanas
3	Análisis y modelización del comportamiento financiero	4 semanas
4	Desarrollo de modelos predictivos	5 semanas
5	Gemelo digital y módulo de simulación	4 semanas
6	Interfaz gráfica, integración y validación	4 semanas
7	Redacción de la memoria	4 semanas

Tabla 4.1. Planificación temporal del proyecto

El coste del proyecto se compone de un coste de desarrollo, coste de despliegue e infraestructura, y coste de operación.

4.4.2 HIPÓTESIS DEL ESTUDIO ECONÓMICO

El análisis económico se ha fundamentado en el conjunto de hipótesis recogidas en la siguiente tabla (Tabla 4.2). Todas las cifras se expresarán en euros. Se adopta un horizonte de análisis de tres años (2026-2028), en busca de coherencia con el ciclo típico de maduración de una plataforma de gestión financiera personal en fase de lanzamiento.

Parámetro	Valor asumido	Justificación
Horizonte temporal	3 años (2026-2028)	Ciclo estándar de evaluación para plataformas SaaS en etapa inicial

Tasa nominal de descuento	6 % anual	Prima de riesgo de proyectos tecnológicos early-stage (Damodaran 2024)
Tasa de inflación	3 % anual	Previsión del BCE (BCE, 2024) para la zona euro 2025-2027
Coste unitario de personal	20 EUR/h brutos	Salario medio de ingeniero junior de software en España
Usuarios objetivo (año 3)	500 usuarios activos	Estimación con captación orgánica y marketing de bajo coste
Rampa de usuarios	100 / 300 / 500	Crecimiento moderado en los tres años
Coste de adquisición (CAC)	10 EUR/usuario nuevo	Captación de bajo coste (canales orgánicos y digitales)
Frecuencia de sincronización	2 llamadas API/usuario/mes	Actualización quincenal de datos bancarios
Herramientas de desarrollo	Coste nulo	Stack íntegramente open-source (Python, scikit-learn, Dash, FastAPI)
Entorno de desarrollo TrueLayer	Coste nulo	Sandbox académico sin cuota ni restricción de uso

Tabla 4.2. Hipótesis del estudio económico

4.4.3 ESTRUCTURA DE COSTES

El coste total del proyecto se estructura en dos componentes: la inversión inicial (año 0) y los costes operativos anuales recurrentes (OPEX).

4.4.3.1 Inversión inicial

La inversión inicial corresponde de manera íntegra al coste de desarrollo, cuantificado a partir de la dedicación total de 300 horas al coste unitario establecido en la Tabla 4.2. Las herramientas de desarrollo no tienen ningún coste adicional al tratarse de software de código abierto, y el entorno Sandbox de TrueLayer no requiere licencia en esta fase de desarrollo. La inversión inicial asciende a 6000 EUR (300h x 20 EUR/h).

4.4.3.2 Costes operativos anuales (OPEX)

Una vez se ha llevado a cabo el despliegue de la plataforma, los costes recurrentes corresponden a cuatro dimensiones diferentes: la infraestructura en la nube, acceso a la API de open banking, mantenimiento técnico y la captación de nuevos usuarios.

La infraestructura se compone de una instancia de cómputo EC2 (t3.medium) que permite que la aplicación esté disponible continuamente, una base de datos gestionada RDS PostgreSQL con copias de seguridad automáticas, almacenamiento de objetos S3 para ficheros y registros, gestión de dominio y certificado TLS mediante Route 53 y ACM, y monitorización mediante CloudWatch.

Se ha elegido AWS frente a otros proveedores por la experiencia previa adquirida durante la titulación y la familiaridad con su ecosistema.

Recurso	Servicio	Coste/mes	Coste/año
Servidor de aplicación	EC2 t3.medium (2 vCPU, 4 GB RAM)	30 EUR	360 EUR
Base de datos relacional	RDS PostgreSQL db.t3.micro	15 EUR	180 EUR
Almacenamiento de objetos	S3 (5 GB)	1 EUR	12 EUR
Dominio y certificado SSL	Route 53 + ACM	5 EUR	60 EUR
Monitorización	CloudWatch (capa básica)	0 EUR	0 EUR
Subtotal infraestructura AWS		51 EUR	612 EUR

Tabla 4.3. Costes de infraestructura de producción (AWS)

El acceso a datos bancarios reales requiere el nivel Scale de TrueLayer, el cual incorpora una cuota fija mensual de 300 EUR (3.600 EUR/año). A ello se le añade un componente variable de 0.14 EUR por llamada API. El coste de mantenimiento y de operación se estima en 10 horas de trabajo semanales (actualización de dependencias, monitorización de cambios en endpoints bancarios y mejoras del modelo), lo que suma 10.400 EUR/año aplicando la métrica de 20 EUR/h brutos por personal técnico. Se estima un coste medio de 10 EUR por usuario nuevo captado mediante canales de marketing digital de bajo coste (redes sociales,

boca a boca...) y sin inversión en publicidad pagada masiva. La Tabla 4.4 recoge la evolución del OPEX durante el horizonte de 3 años.

Partida	Año 1 (100)	Año 2 (300)	Año 3 (500)
Infraestructura AWS	612 EUR	612 EUR	612 EUR
TrueLayer Scale - cuota fija	3.600 EUR	3.600 EUR	3.600 EUR
TrueLayer Scale - cuota variable	336 EUR	1.008 EUR	1.680 EUR
Mantenimiento y operación	10.400 EUR	10.400 EUR	10.400 EUR
Captación de usuarios (CAC)	1.000 EUR	2.000 EUR	2.000 EUR
Total OPEX anual	15.948 EUR	17.620 EUR	18.292 EUR

Tabla 4.4. Costes operativos anuales por escenario de usuarios

El coste total del proyecto a tres años asciende a 57.860 EUR (6000 EUR de inversión inicial más 51.860 de OPEX acumulado).

4.4.4 COSTE NORMALIZADO POR USUARIO

Para evaluar la eficiencia económica de la plataforma de manera independiente de la escala, se calcula el Coste Normalizado (Levelised Cost, LC), definido como cociente entre el valor de todos los costes y el valor del total de usuarios.

$$LC = \frac{\sum \frac{Ct}{(1+r)^t}}{\sum \frac{Ut}{(1+r)^t}}$$

Donde Ct es el coste total en el periodo t, Ut el número de usuarios activos en el mismo periodo t y r= 0.06 la tasa nominal de descuento. Aplicando esta fórmula con los valores de las Tablas 4.2 y 4.4 se obtienen los resultados de las Tablas 4.5 y 4.6.

Periodo	Coste nominal	Factor (1,06)^t	Coste actualizado
t = 0 (inversión)	6.000 EUR	1,0000	6.000 EUR
t = 1	15.948 EUR	1,0600	15.045 EUR

t = 2	17.620 EUR	1,1236	15.682 EUR
t = 3	18.292 EUR	1,1910	15.358 EUR
Total	57.860 EUR		52.085 EUR

Tabla 4.5. Valor presente de costes

Periodo	Usuarios activos	Factor $(1,06)^t$	Usuarios actualizados
t = 1	100	1,0600	94,3
t = 2	300	1,1236	267,0
t = 3	500	1,1910	419,8
Total			781,1 usuarios-año

Tabla 4.6. Valor presente de usuarios-año

El Coste Normalizado es de 66,7 EUR/usuario-año, lo que equivale a aproximadamente 5,6 EUR/usuario-mes. Esta cifra se compara favorablemente con el rango de precios del mercado de otras soluciones discutidas previamente en el **Capítulo 3** (8-15 EUR/mes). Esto confirma la viabilidad económica de la propuesta a partir de 500 usuarios activos.

4.4.5 ANÁLISIS DE SENSIBILIDAD

Con el objetivo de evaluar la robustez del modelo ante variaciones en la captación de usuarios, se calculan los Costes Normalizados para dos escenarios adicionales: uno pesimista (200 usuarios en el año 3) y uno optimista (1000 usuarios). En ambos se mantiene el coste de captación de 10 EUR por usuario nuevo.

Escenario	Usuarios (Año 3)	LC (EUR/usuario-año)	LC (EUR/mes)	Dif. vs 8 EUR/mes
Pesimista	200	148,8	12,4	+55 %
Base	500	66,7	5,6	-30 %
Optimista	1.000	37,8	3,2	-60 %

Tabla 4.7. Análisis de sensibilidad del Coste Normalizado

Este análisis evidencia una alta sensibilidad al volumen de usuarios en la zona de menos escala. En el escenario pesimista el Coste Normalizado supera el precio mínimo del mercado. La causa principal es indudablemente la cuota fija de TrueLayer Scale (3.600 EUR/año), independientemente del número de usuarios. Una medida para mitigar esto es posponer la contratación del nivel Scale hasta alcanzar el umbral mínimo de usuarios activos.

4.4.6 VIABILIDAD ECONÓMICA Y PROPUESTA DE VALOR

PREFIN no genera ingresos de manera directa; su propuesta de valor se encuentra en la disponibilidad y el acceso libre a herramientas de análisis financiero avanzado. Con el objetivo de evaluar su viabilidad, conviene distinguir entre dos conceptos que no deben confundirse: el coste de prestar el servicio y el valor que este genera para el usuario.

El coste normalizado calculado en el apartado anterior (5,6 €/usuario-mes en el escenario base) representa el gasto que asume la plataforma para operar, y no un precio que pague el usuario, que accede a la herramienta de forma gratuita. Frente a este coste, el valor entregado puede estimarse desde dos perspectivas.

Desde la perspectiva del coste de mercado, PREFIN ofrece de forma gratuita un conjunto de funcionalidades cuyo equivalente en las plataformas competidoras identificadas en el *Capítulo 3* se sitúa entre 8 y 15 €/mes. El coste de prestación del servicio se mantiene, por tanto, por debajo del precio de referencia de soluciones análogas, lo que confirma que la plataforma es eficiente: produce a un coste inferior al que el mercado asigna a un servicio equivalente.

Desde la perspectiva del impacto financiero, el valor reside en las decisiones que el usuario puede tomar gracias al sistema: anticipar situaciones de iliquidez, reducir gasto evitable y optimizar el ahorro a medio plazo. Este componente, más difícil de cuantificar, pero más relevante, constituye la verdadera aportación de la herramienta, pues su finalidad no es competir en precio, sino mejorar la salud financiera de personas que probablemente no contratarían ningún servicio de pago.

El análisis no pretende establecer la sostenibilidad de un modelo de negocio a largo plazo, sino demostrar que el coste de operar PREFIN es reducido y proporcional al valor que aporta. A modo ilustrativo, una hipotética monetización parcial mediante un modelo freemium de 3 €/mes por cuenta premium cubriría íntegramente los costes operativos a partir de aproximadamente 508 cuentas de pago, lo que evidencia que la plataforma podría alcanzar el equilibrio financiero incluso con una baja penetración de pago sobre su base de usuarios.

Capítulo 5. SISTEMA/MODELO DESARROLLADO

El presente capítulo explica la arquitectura y el funcionamiento interno de PREFIN. Se presenta cada módulo de manera ordenada siguiendo el grado de inteligencia que aporta: primero los componentes descriptivos, que caracterizan la situación financiera del usuario; a continuación, los predictivos, que anticipan la evolución cuantificando la incertidumbre y, por último, los componentes de prevención y explicación, que permiten al sistema mostrar alertas tempranas y de transparencia. Esta progresión constituye el eje conceptual del trabajo: la transformación de un simple Dashboard financiero a un sistema capaz de razonar y recomendar.

5.1 VISIÓN GENERAL DE LA ARQUITECTURA

PREFIN cuenta con una estructura de dos procesos independientes que se comunican entre sí mediante peticiones HTTP. El primero de ellos es la interfaz del usuario, implementada con el framework Dash, que recoge la totalidad de las vistas, los componentes reutilizables y la lógica de interacción. El segundo, es un servicio construido sobre FastAPI, necesario solamente para la conexión con la banca abierta: utilizado para resolver el flujo de autorización OAuth 2.0 frente a TrueLayer y además, custodia los tokens de acceso a una base de datos de SQLite. Esta separación se debe a que de esta manera la visualización y el análisis quedan desligados de la gestión de credenciales sensibles.

El elemento que construye la columna vertebral de la arquitectura es un esquema de datos común. La aplicación recibe la información de tres orígenes distintos, cada uno con su propio formato, pero antes de procesar los datos todos se traducen a una misma estructura con un conjunto determinado de columnas: fecha, descripción, importe con signo (positivo si se trata de ingresos y negativo para los gastos), divisa, categoría y saldo acumulado. Esta unificación permite que los módulos de análisis no necesiten conocer el origen de los datos, sino que operan siempre sobre la misma estructura con independencia de la procedencia de la información.

Las tres fuentes son el generador de datos sintéticos, la carga de extractos en formato CSV o Excel y la conexión con TrueLayer en su entorno de pruebas. Todas ellas convergen en este esquema a través de un único punto de entrada, lo que centraliza la normalización y mantiene el sistema sencillo y homogéneo.

La *Figura 5.1* resume este flujo, desde las fuentes hasta la capa de inteligencia.

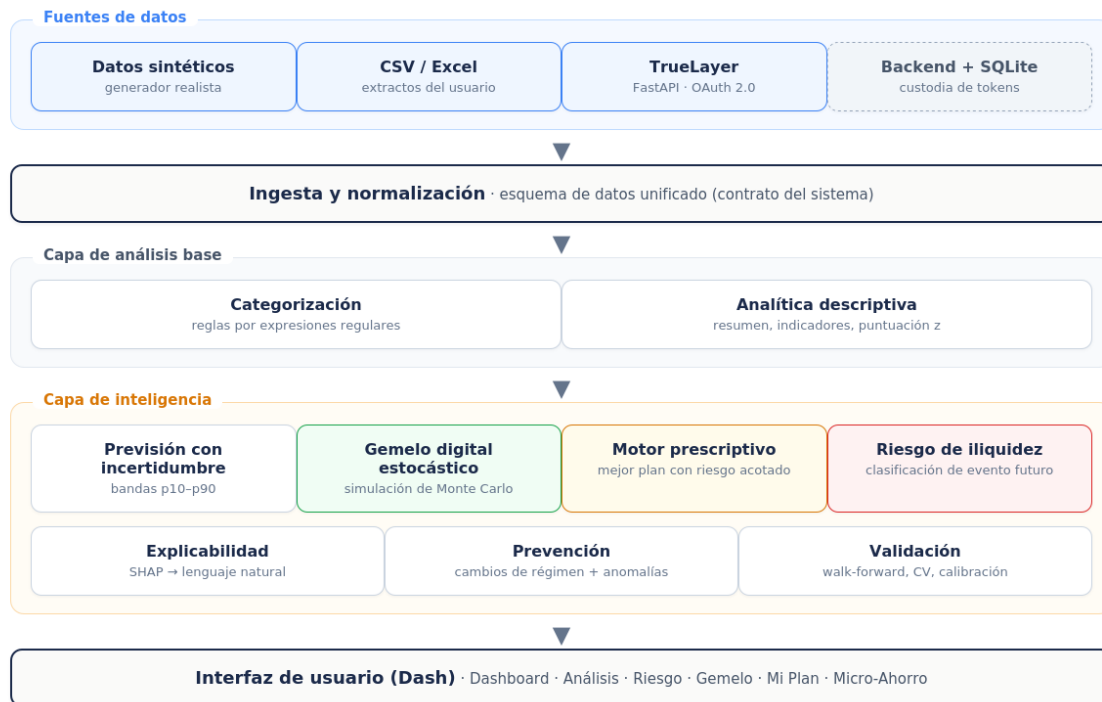


Figura 5.1. Arquitectura conceptual de PREFIN por niveles de inteligencia.

5.2 ANÁLISIS DEL SISTEMA

El punto de partida fue la comprensión de las necesidades del público objetivo: personas sin experiencia financiera que desean entender mejor su situación económica sin recurrir a un asesor profesional. A partir de este perfil se desarrollaron los requisitos, que se distinguen entre funcionales, las capacidades que el sistema debe ofrecer, y los no funcionales, las cualidades transversales que debe satisfacer.

5.2.1 REQUISITOS FUNCIONALES

Los requisitos funcionales se agrupan según las capas de la arquitectura presentada en la *Figura 5.1*, cada capacidad del sistema corresponde con el bloque que la implementa.

Capa de la arquitectura	Requisitos funcionales
Ingesta de datos	Admitir tres fuentes de datos: generación sintética, carga de extractos en CSV o Excel y conexión con cuentas bancarias mediante Open Banking. Normalizar toda entrada a un esquema de datos común.
Análisis descriptivo	Visualizar la situación financiera actual mediante indicadores clave (saldo, ingresos, gastos y tasa de ahorro). Analizar el gasto por categorías. Detectar automáticamente transacciones inusuales.
Predicción	Prever el gasto futuro acompañado de intervalos de incertidumbre. Clasificar el riesgo financiero del usuario en tres niveles (bajo, medio y alto) mediante aprendizaje automático.
Prescripción	Simular escenarios alternativos hasta un horizonte de 36 meses. Recomendar el mejor plan de acción manteniendo el riesgo bajo control. Estimar el ahorro potencial mediante micro-redondeo.
Prevención y explicabilidad	Generar alertas tempranas ante cambios sostenidos en el comportamiento de gasto. Explicar las decisiones del modelo de riesgo en lenguaje natural.

Tabla 5.1 Requisitos funcionales agrupados por la capa de arquitectura

5.2.2 REQUISITOS NO FUNCIONALES

Los requisitos no funcionales describen cualidades del sistema en su conjunto, no se atribuyen a un módulo concreto. Los más relevantes se recogen en la *Tabla 5.2*.

Requisito no funcional	Descripción
Usabilidad	Interfaz web que no requiere instalación y emplea un lenguaje claro, adecuado para personas sin formación financiera.
Modularidad	Cada funcionalidad se implementa en un módulo de Python independiente, lo que facilita el mantenimiento y la evolución.
Reproducibilidad	Los datos sintéticos se generan de forma determinista a partir de una semilla, de modo que los resultados son reproducibles.

Seguridad	El acceso a datos bancarios reales se realiza mediante OAuth 2.0 a través de TrueLayer, sin almacenar credenciales del usuario.
Rendimiento	La simulación de miles de trayectorias se ejecuta de forma vectorizada en tiempos del orden de milisegundos.
Transparencia	Las decisiones del sistema son trazables y explicables, evitando el comportamiento de caja negra.

Tabla 5.2 Requisitos no funcionales del sistema

5.3 FUENTES DE DATOS Y ESTRATEGIA DE IMPORTACIÓN

Como ya se ha mencionado previamente en el presente capítulo, el sistema admite tres fuentes de datos que convergen en un esquema común con un único punto de entrada. Este punto centraliza la normalización; ajusta los nombres de las columnas, interpreta los importes en formato europeo, ordena las operaciones por fecha y calcula el saldo acumulado. La elección de este proceso que tolera distintos formatos permite maximizar la usabilidad ante la diversidad de formatos bancarios reales.

El generador de datos sintéticos reproduce un histórico bancario realista en el contexto español: incorpora comercios reconocibles, el ingreso de la nómina entre los días 25 y 28 del mes, el pago del alquiler, recibos y suscripciones periódicas, además de un gasto variable por categoría. Para lograr un mayor realismo añade estacionalidad anual, recibos de periodicidad no mensual como impuestos y seguros, gastos imprevistos puntuales e ingresos irregulares. Además, puede generar una población de múltiples usuarios diferenciados, capacidad determinante para el rigor de la validación que se detalla en la Sección 5.7.

La carga de extractos en CSV o Excel permite al usuario incorporar sus propios datos.

Por último, la conexión con TrueLayer opera en un entorno de pruebas (sandbox), un simulador que no accede a la cuenta bancaria real. La conexión a las cuentas bancarias reales requiere una licencia de proveedor de servicios de pago y un proceso de certificación, inviábiles en el marco de un trabajo académico; esta limitación, motiva la inclusión de un modo de demostración con datos sintéticos realistas. La Figura 5.2 detalla la arquitectura completa a nivel de implementación del dato bruto a la recomendación.

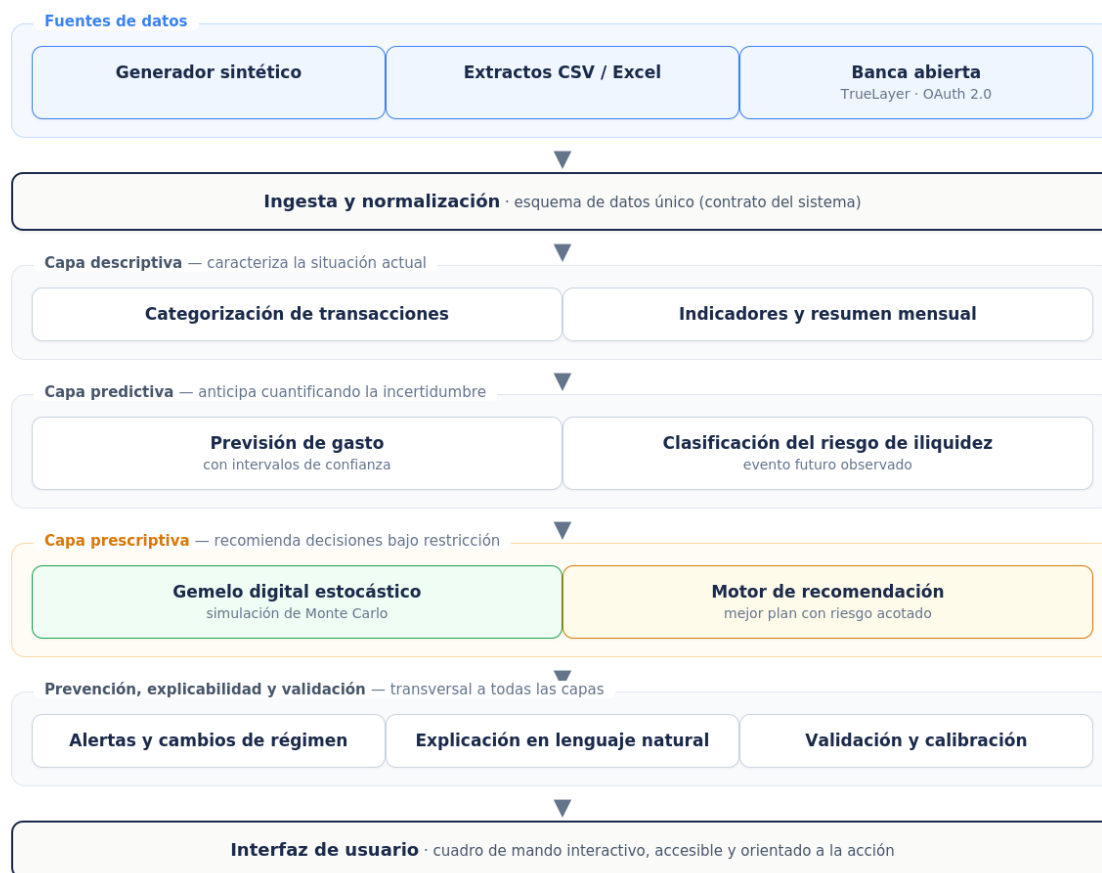


Figura 5.2. Arquitectura de PREFIN a nivel de implementación.

Las tres fuentes y el servicio backend convergen en la ingesta, sobre la que se asienta la capa de análisis base y la capa de inteligencia, mostradas en la interfaz Dash.

5.4 TECNOLOGÍAS SELECCIONADAS

La Tabla 5.3 resume las tecnologías empleadas en el sistema y la justificación de cada elección. El conjunto se ha mantenido reducido y basado en un software de código abierto en coherencia con el requisito de modularidad y con un coste de herramientas nulo.

Tecnología	Función en el sistema y justificación
Dash	Framework de interfaz web en Python. Permite construir cuadros de mando interactivos sin separar frontend y backend, reduciendo la complejidad.

FastAPI	Servicio backend ligero para el flujo OAuth 2.0 con TrueLayer. Elegido por su sencillez y rendimiento en peticiones asíncronas.
scikit-learn	Biblioteca de aprendizaje automático: clasificación de riesgo, regresión cuantílica y detección de anomalías.
NumPy / SciPy	Cálculo numérico vectorizado. Sustentan la simulación de Monte Carlo y el muestreo de distribuciones.
SQLite	Base de datos ligera para la custodia de los tokens de acceso de TrueLayer.
SHAP	Cálculo de la contribución de cada variable a las predicciones del modelo de riesgo, base de la explicabilidad.

Tabla 5.3 Tecnologías seleccionadas y su justificación

5.5 ANÁLISIS DESCRIPTIVO

Sobre el esquema común se construye una capa de análisis descriptivo que prepara la información básica que después utilizan los módulos avanzados. Esta capa calcula el resumen mensual de ingresos, gastos y ahorro, junto con la tasa de ahorro y agrupa los gastos por categorías y por mes. Obtiene una serie de indicadores generales, señala las transacciones atípicas comparando cada gasto con la media de su categoría; y estima la tendencia del gasto medio mediante una regresión lineal. En conclusión, describe con precisión la situación actual y pasada del usuario, pero no predice ni recomienda nada.

La categorización de las transacciones se resuelve con un conjunto de reglas basadas en expresiones regulares. Se eligió este método, sencillo y fácil de auditar, frente a un clasificador de texto basado en aprendizaje automático, por dos razones. La primera es la transparencia: cada transacción puede rastrearse hasta la regla concreta que le asignó su categoría. La segunda es práctica: un clasificador de aprendizaje automático necesitaría entrenarse con un gran volumen de transacciones previamente categorizadas a mano, y no existe un conjunto de datos de ese tipo para el dominio bancario español.

5.6 PREVISIÓN DE GASTO CON INCERTIDUMBRE

El primer componente predictivo aborda la previsión del gasto mensual. A diferencia de una estimación puntual, que podría transmitir una falsa sensación de certeza, ya que el futuro es

incierto, el módulo proporciona una banda de previsión definida por los percentiles 10 (hay un 10% de probabilidad de gastar menos de esto), 50 (la más probable) y 90 (hay un 10% de probabilidad de gastar más de esto).

El gasto variable restante se modela mediante regresión cuantílica, una técnica que, en lugar de predecir un único valor, estima directamente los percentiles que forman la banda. Esta previsión se combina después con una referencia sencilla, que simplemente asume que cada mes se parecerá al mismo mes del año anterior (lo que captura la estacionalidad). Combinar ambos métodos, en lugar de usar solo uno, sigue un principio clásico de la estadística (Bates y Granger, 1969) según el cual la media de dos previsiones suele ser más fiable que cualquiera de ellas por separado.

Se optó por un modelo lineal, y no por uno más complejo, debido a la cantidad de datos disponible: con un histórico de unos treinta y seis meses y una estacionalidad marcada, un modelo más sofisticado tendería al sobreajuste, es decir, a memorizar el pasado en lugar de aprender a predecir. La principal aportación de este módulo frente a la referencia sencilla no es tanto acertar más en el valor exacto, sino ofrecer una banda de incertidumbre cuantificada, algo de lo que la referencia carece (Figura 5.3).

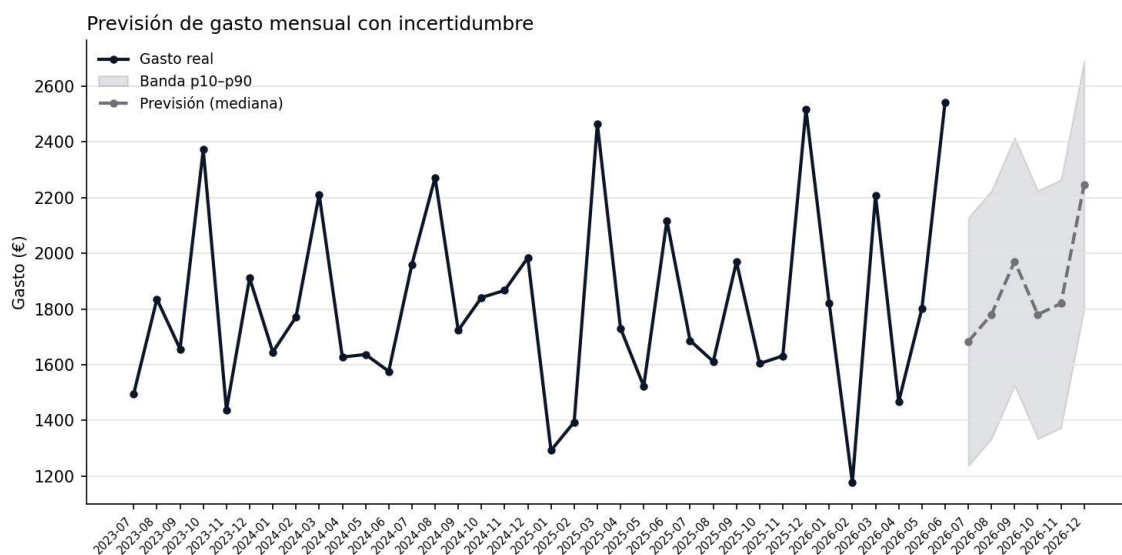


Figura 5.3 Previsión del gasto mensual con incertidumbre.

La gráfica enseña el gasto real del pasado como una línea, y a partir de ahí continúa hacia el futuro con una franja de posibilidades en vez de una línea exacta, porque el futuro no se puede predecir con un solo número.

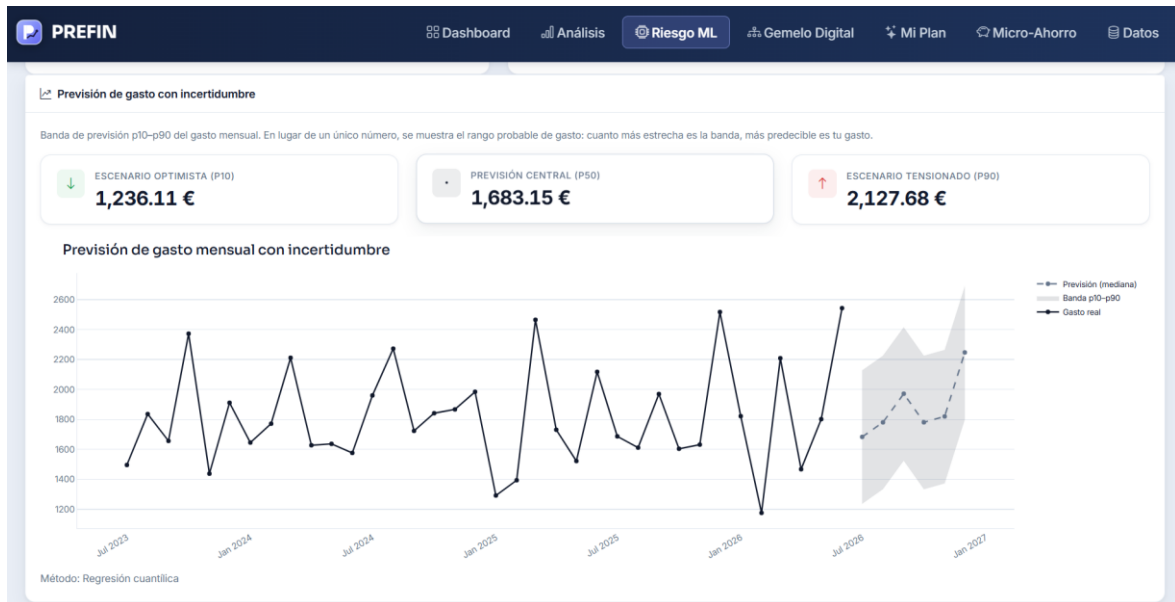


Figura 5.4. Pestaña de previsión: banda p10–p90 del gasto mensual, con los escenarios optimista, central y tensionado.

5.7 CLASIFICACIÓN DEL RIESGO DE ILIQUIDEZ

El segundo componente predictivo clasifica el riesgo de que el usuario no tenga liquidez en un horizonte próximo. Este módulo recibió una atención especial porque su desarrollo inicial tenía un defecto grave.

5.7.1 EL PROBLEMA DE LA CIRCULARIDAD

En su versión original, el modelo funcionaba de forma circular. El nivel de riesgo se calculaba primero con una regla fija a partir de tres variables: la proporción entre gasto e ingreso, la proporción de gasto en ocio y la tasa de ahorro, y, a continuación, se entrenaba un clasificador para predecir ese mismo nivel de riesgo dándole como entrada precisamente esas tres variables. El resultado era una trampa: el modelo recibía exactamente los datos con

los que se había construido la respuesta, de modo que podía acertar casi el cien por cien de las veces sin haber aprendido nada que la regla no supiera ya. No predecía nada nuevo; se limitaba a reproducir la regla de partida. En la práctica, el aprendizaje automático era meramente decorativo.

5.7.2 RUPTURA DE LA CIRCULARIDAD

La solución redefine el problema en cuatro pilares. Primero, la etiqueta pasa a ser un evento futuro observado y no una regla sobre el presente: para cada usuario y mes se determina si su saldo desciende por debajo de cero durante los tres meses siguientes, un resultado independiente de las variables de entrada. Segundo, las variables describen el estado presente y el modelo aprende a anticipar este evento. Tercero, el entrenamiento se realiza sobre una población de cuarenta usuarios con validación por grupos, evaluándose sobre usuarios nunca vistos durante el entrenamiento, lo que descarta la memorización. Cuarto, el enriquecimiento del generador asegura que la iliquidez sale del propio proceso y no de una regla impuesta. El clasificador resultante deja de reproducir una regla y pasa a resolver un problema predictivo. Las métricas de validación se presentan en el *Capítulo 6*.

5.8 GEMELO DIGITAL ESTOCÁSTICO

El gemelo digital constituye el primer componente prescriptivo y el fundamento sobre el que se apoya el motor de recomendación. En una primera aproximación, el gemelo proyectaba el saldo futuro como una única trayectoria, a la que se añadía una pequeña variación aleatoria. Este enfoque resultaba limitado, porque no permitía responder a preguntas tan importantes como cuál es la probabilidad real de quedarse sin dinero. Por ello, se sustituyó por una simulación de Monte Carlo.

El nuevo enfoque funciona así: para cada categoría de gasto se hace una estimación de una distribución de probabilidad a partir del historial del usuario, y a continuación se generan varios miles de futuros posibles del saldo, cada uno con variaciones realistas. A la hora de juntar todos esos escenarios se obtiene una imagen completa y probabilística del futuro: al franja en la que se moverá el saldo, la probabilidad de quedarse sin liquidez en cada mes, y dos indicadores del peor caso habituales en finanzas, el valor en riesgo (VaR) y su versión

más estricta (CVaR), que miden cuánto se podría llegar a perder en los escenarios más desfavorables.

Cada categoría se modela con una distribución log-normal, apropiada porque el gasto nunca es negativo y suele ser asimétrico (muchos gastos pequeños y algunos muy grandes). Además, la simulación está optimizada para ejecutar miles de escenarios en milisegundos, sin necesidad de librerías adicionales.

El valor de este enfoque se aprecia al comparar situaciones. La Figura 5.5 muestra a un usuario en una situación saneada, con probabilidad de quedarse sin dinero nula. La Figura 5.6 muestra al mismo usuario ante un gasto imprevisto importante: la franja de escenarios se ensancha, cae por debajo del nivel de iliquidez y la probabilidad de quedarse sin dinero se dispara. Esta capacidad de cuantificar la incertidumbre y el riesgo de los peores casos es justo lo que una predicción de un solo valor nunca puede ofrecer.

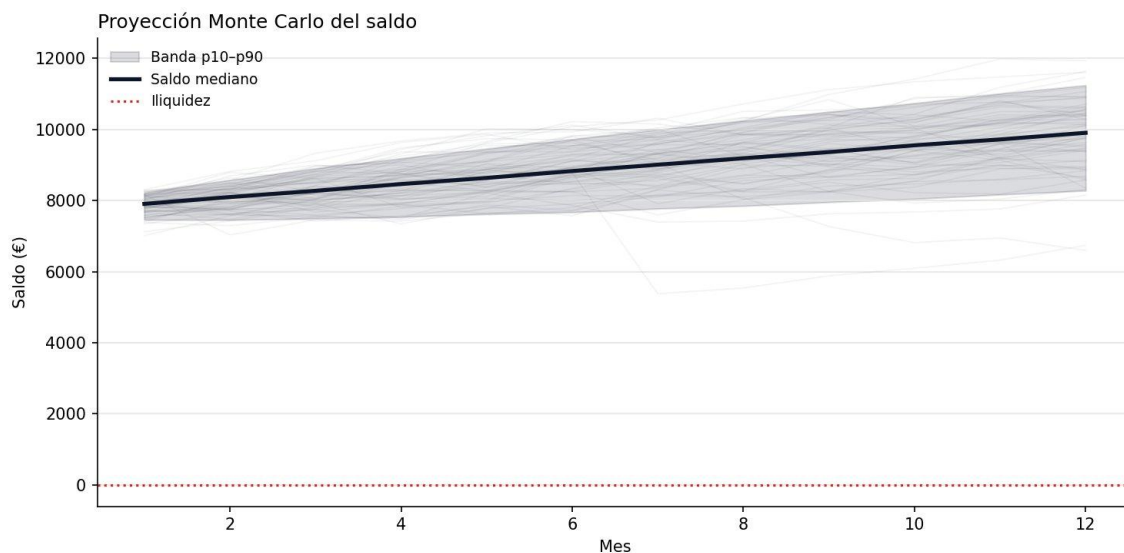


Figura 5.5. Proyección de Monte Carlo del saldo para un usuario en situación saneada. La banda p10-p90 se mantiene holgadamente por encima del umbral de iliquidez.

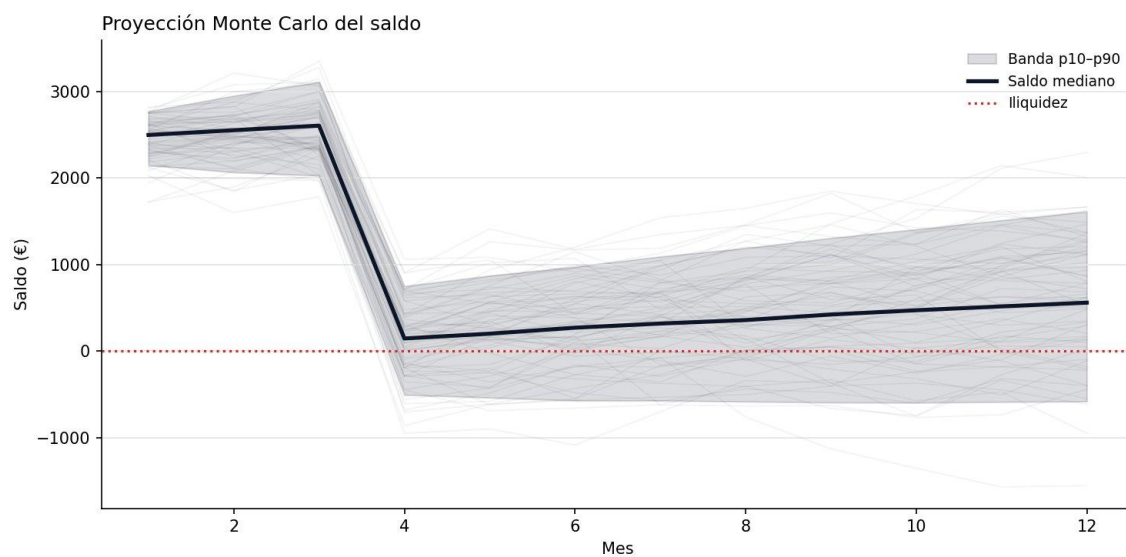


Figura 5.6. Proyección de Monte Carlo del mismo usuario ante un gasto imprevisto. La banda desciende por debajo del umbral de iliquidez, con la consiguiente elevación de la probabilidad de iliquidez y un valor en riesgo negativo.

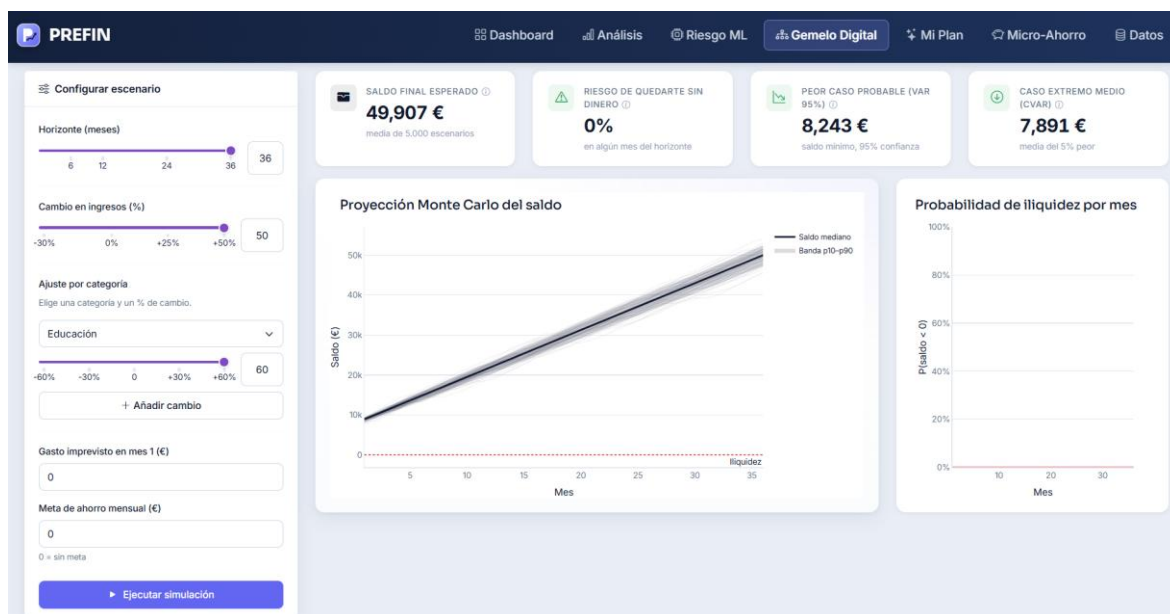


Figura 5.7. Pestaña del gemelo digital: proyección de Monte Carlo del saldo a 36 meses con sus métricas de riesgo (probabilidad de iliquidez, VaR y CVaR).

5.9 MOTOR PRESCRIPTIVO DE RECOMENDACIÓN

El motor prescriptivo representa el núcleo del salto cualitativo del sistema: mientras que los módulos anteriores describen y anticipan, este recomienda un curso de acción. Utiliza el gemelo de Monte Carlo como entorno de evaluación y busca, entre las siguientes palancas —la cantidad ahorrada de forma programada, la intensidad del redondeo, los recortes por categoría y el momento de afrontar un gasto puntual—, la combinación que maximiza el ahorro protegido sujeto a que la probabilidad de iliquidez no supere un umbral fijado por el usuario.

La búsqueda se realiza por exploración aleatoria del espacio de palancas seguida de un refinamiento de las mejores candidatas; se prefirió este procedimiento frente a una optimización bayesiana por su transparencia y por generar de modo natural un ranking de planes.

Cada plan se presenta con su impacto cuantificado y una explicación en lenguaje claro. A modo de ejemplo, para el usuario de demostración, con un horizonte de veinticuatro meses y una tolerancia al riesgo del treinta por ciento, el motor evalúa 279 combinaciones y recomienda un plan que permite acumular 7.392 € de ahorro protegido manteniendo la probabilidad de iliquidez en el cero por ciento, con un valor en riesgo favorable de 3.592 € (Tabla 5.4).

Parámetro / resultado	Valor
Horizonte de planificación	24 meses
Tolerancia al riesgo de iliquidez	30 %
Combinaciones evaluadas	279
Ahorro protegido (plan recomendado)	7.392 €
Probabilidad de iliquidez resultante	0 %
Valor en riesgo al 95 % (VaR)	3.592 €

Acciones recomendadas	200 €/mes · redondeo a 5 € · -20 % en Restauración y Ocio, Ropa y Compras y Suscripciones
-----------------------	---

Tabla 5.4 Ejemplo de plan recomendado por el motor prescriptivo para el usuario de demostración.

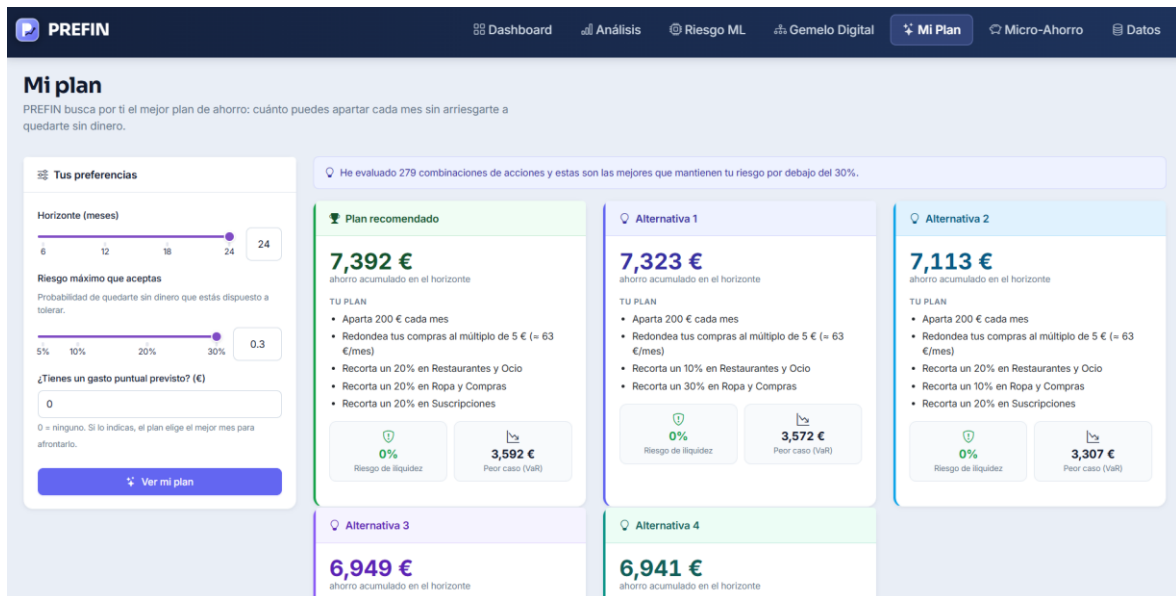


Figura 5.8. Pestaña "Mi Plan": planes de acción recomendados por el motor prescriptivo, con su impacto cuantificado y su explicación en lenguaje natural.

5.10 PREVENCIÓN: CAMBIOS DE RÉGIMEN Y ANOMALÍAS

La componente de prevención dota al sistema de capacidad de alerta temprana, dando contenido real a la vertiente preventiva que da nombre al proyecto (la "P" de Prevención de PREFIN). Va más allá de la simple detección de gastos puntuales fuera de lo común y emplea dos mecanismos que se complementan.

El primero es la detección de cambios de régimen, que identifica cuándo el nivel de gasto del usuario se desplaza de forma sostenida en el tiempo —por ejemplo, un alquiler que sube y se queda alto— en lugar de limitarse a señalar picos aislados de un solo mes. Para ello divide la serie de gasto en tramos y detecta el punto en que cambia de nivel de manera estable.

El segundo es un detector de anomalías que aprende el patrón de comportamiento propio de cada usuario y señala las transacciones que se desvían de él. A diferencia del primero, no se fija en una sola variable, sino en varias a la vez (importe, categoría, frecuencia...), y no necesita ejemplos previos de qué es "anómalo": lo deduce por sí mismo a partir de los datos del usuario.

Se conserva además la detección por puntuación z descrita anteriormente, que sirve como referencia para comparar. La Figura 5.9 muestra la identificación de un cambio de régimen sostenido en el gasto.

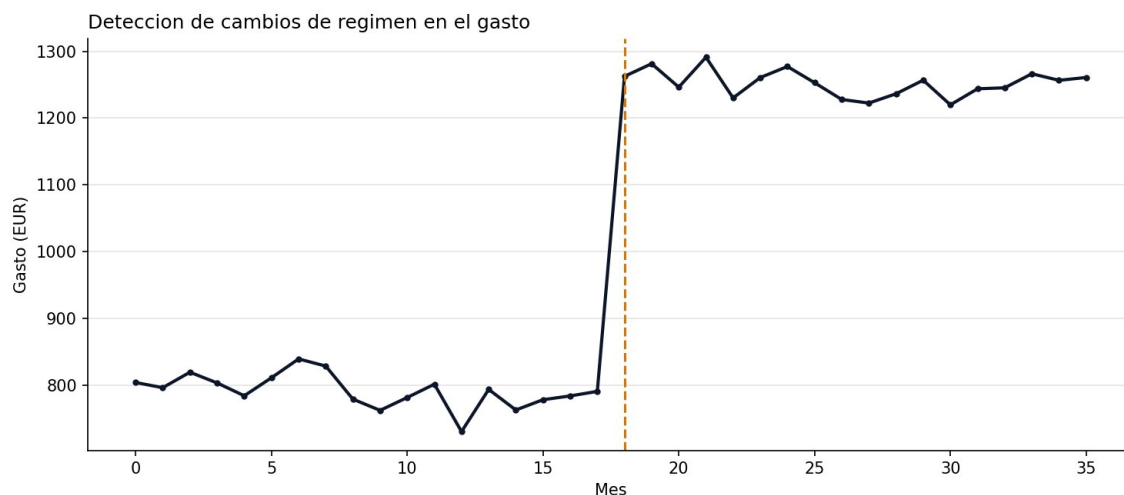


Figura 5.9. Detección de un cambio de régimen en el gasto mensual. La línea vertical señala el instante a partir del cual el nivel de gasto se desplaza de forma sostenida.

5.11 EXPLICABILIDAD DE LAS DECISIONES

Para que las recomendaciones sean comprensibles y accionables, se incorpora un módulo de explicabilidad que traduce las predicciones del modelo de riesgo a lenguaje natural. El procedimiento se basa en los valores de Shapley, que reparten entre las variables la diferencia entre la probabilidad predicha y la de referencia. El módulo distingue el hecho —el valor real de cada variable— del efecto —el sentido en que eleva o reduce el riesgo— y genera frases concretas en lugar de tecnicismos. La explicabilidad se aplica sobre el modelo ya

rectificado descrito en la Sección 5.7; aplicarla sobre el modelo circular original carecería de sentido. La Figura 5.10 muestra la descomposición de una predicción de riesgo.

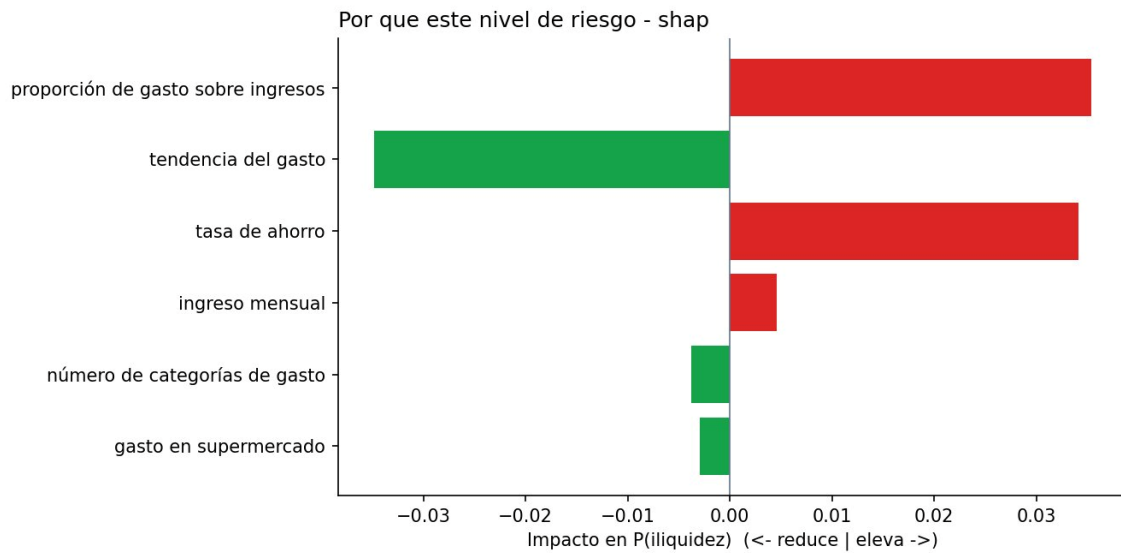


Figura 5.10. Contribuciones de cada variable a la predicción de riesgo, obtenidas mediante valores de Shapley. Las contribuciones en rojo elevan la probabilidad de iliquidez y las verdes la reducen.

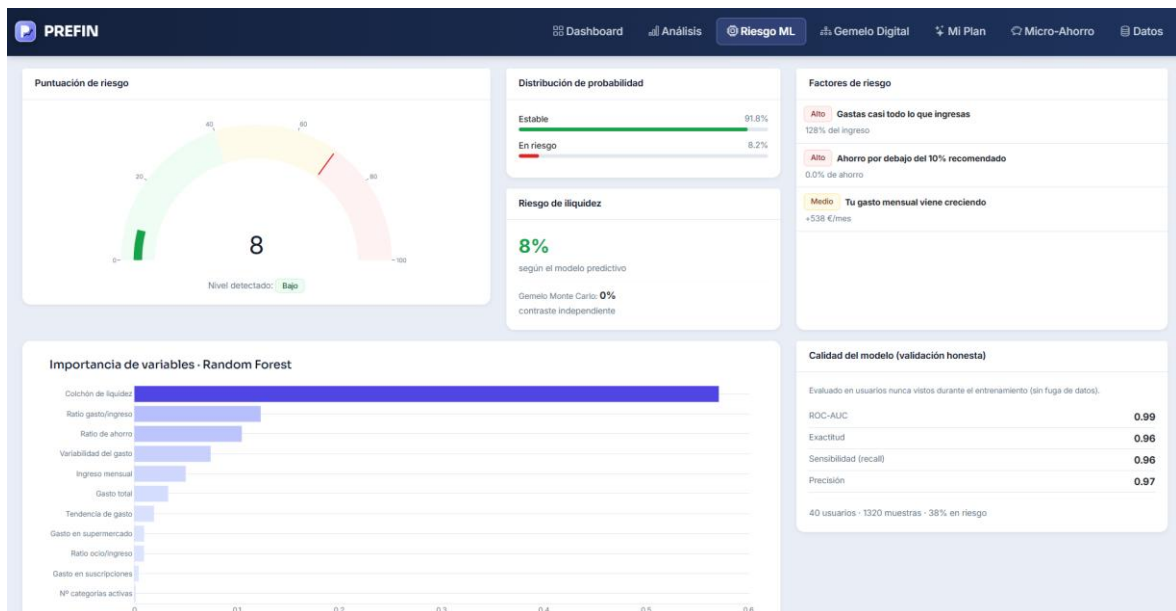


Figura 5.11. Pestaña de riesgo: nivel de riesgo, factores explicativos en lenguaje natural e importancia de variables del modelo. El recuadro de calidad del modelo muestra la validación en usuarios no vistos (ROC-AUC 0,99).

5.12 MICRO-AHORRO POR REDONDEO

Por último, el módulo de micro-ahorro calcula el ahorro que se obtendría si se redondease al alza cada gasto a distintas unidades, la proyección anual de dicho ahorro, su distribución por categoría y el tiempo necesario para alcanzar metas habituales. Se trata de un cálculo puramente aritmético que, por honestidad metodológica, no se presenta como un componente de inteligencia artificial; no obstante, alimenta una de las palancas del motor prescriptivo descrito en la Sección 5.9, integrándose en la lógica de recomendación del conjunto del sistema.

Capítulo 6. ANÁLISIS DE RESULTADOS

Este capítulo presenta y analiza de forma crítica los resultados obtenidos por los componentes inteligentes de PREFIN. El objetivo no es únicamente exponer las métricas alcanzadas, sino interpretarlas: distinguir lo que el sistema demuestra de lo que meramente afirma, reconocer las limitaciones de la evaluación y justificar por qué los resultados son fiables o, cuando proceda, por qué deben matizarse. Se evalúan por separado los componentes que sustentan la inteligencia del sistema —la previsión de gasto, el clasificador de riesgo, el gemelo de Monte Carlo y el motor prescriptivo de recomendación— y se cierra con una discusión consolidada de las limitaciones del trabajo.

6.1 METODOLOGÍA Y EVALUACIÓN

Toda la evaluación se ha diseñado con el fin de evitar el optimismo que produce la validación de un modelo sobre los mismos datos con los que se entrena. La previsión de gasto se valida mediante la validación temporal progresiva (walk-forward), en la que el modelo predice cada mes utilizando únicamente la información disponible hasta el momento. El clasificador de riesgo se valida con validación cruzada por grupos, de modo que los usuarios empleados para entrenar nunca coinciden con los usados para evaluar.

Por último, las probabilidades del gemelo Monte Carlo se contrastan con la frecuencia real de los eventos observados (backtesting). Además de la capacidad de discriminar, se evalúan las probabilidades, es decir, hasta qué punto una probabilidad estimada del 30% se corresponde con una frecuencia real del 30%.

6.2 PREVISIÓN DE GASTO CON INCERTIDUMBRE

La Tabla 6.1 recoge las métricas de la previsión de gasto y las compara con un método de referencia muy simple. La previsión se evaluó imitando el uso real: para cada mes, el modelo predice el gasto utilizando únicamente la información disponible hasta ese momento, se

comprueba si acertó y se avanza al mes siguiente; este proceso se repitió 28 veces. Como término de comparación se emplea una predicción de referencia muy básica, que simplemente supone que el gasto de cada mes será igual al del mismo mes del año anterior. Esta comparación permite comprobar si el modelo aporta una mejora real frente a una estimación obvia.

Métrica	Modelo PREFIN	Referencia estacional
Error absoluto medio (MAE)	468,33 €	429,22 €
Raíz del error cuadrático medio (RMSE)	620,51 €	532,42 €
Error porcentual absoluto medio (MAPE)	25,9 %	—
Cobertura del intervalo (objetivo: 80 %)	64,3 %	—

Tabla 6.1. Métricas de la previsión de gasto frente a la referencia estacional.

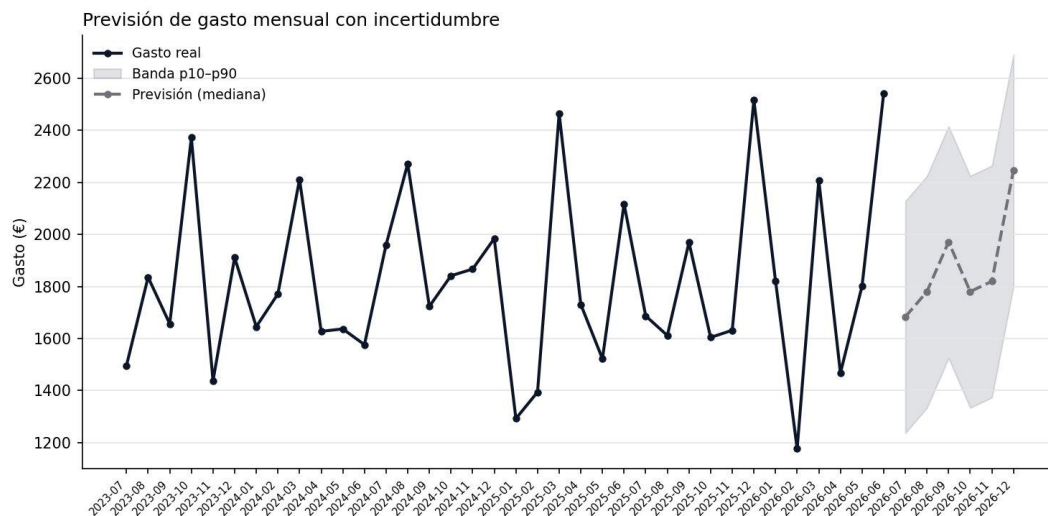


Figura 6.1. Previsión del gasto mensual con banda de incertidumbre p10–p90 sobre el histórico del usuario.

6.2.1 ANÁLISIS CRÍTICO

Los resultados deben interpretarse con honestidad. Para evaluar si el modelo aporta valor, su precisión se compara con la de un método de referencia muy simple, habitual en el análisis de series temporales: predecir que el gasto de cada mes será igual al de ese mismo mes del año anterior. Es la predicción más básica posible —no aprende nada, simplemente copia la

temporada previa—, pero captura bien los patrones que se repiten cada año, por lo que resulta un competidor exigente.

Comparado con esta referencia, el modelo no mejora el acierto del valor exacto: se equivoca de media en 468 €, frente a los 429 € de la referencia, ligeramente mejor. Este resultado no se oculta, sino que se explica. Con solo unos treinta y seis meses de historial y un patrón anual muy repetitivo, superar a un método que simplemente copia el año anterior exigiría muchos más datos de los disponibles. Por eso, el verdadero valor del módulo no reside en acertar la cifra exacta, sino en acompañar la previsión de un rango de incertidumbre. La referencia da un único número y aparenta una seguridad que no tiene; el modelo, en cambio, indica entre qué valores se moverá probablemente el gasto, lo cual es más honesto y más útil para el usuario.

La principal limitación de este componente está en la anchura de ese rango. La banda p10–p90 debería contener el gasto real el 80 % de las veces, pero en la práctica solo lo contiene el 64,3 %. Esto significa que la banda es demasiado estrecha: el sistema se muestra más seguro de lo que debería. Es una limitación conocida y con solución: existen técnicas de calibración de la incertidumbre, como la predicción conforme, que permiten ensanchar la banda hasta que su fiabilidad coincida con la esperada. Esta mejora se propone como línea de trabajo futuro en el *Capítulo 7*.

6.3 CLASIFICACIÓN DE RIESGO DE ILIQUIDEZ

El clasificador de riesgo, una vez resuelto el problema de circularidad descrito en la Sección 5.7, se evalúa sobre una población de 40 usuarios sintéticos (1.320 observaciones, de las cuales un 37,7 % corresponden a eventos de iliquidez) mediante validación cruzada por grupos. La Tabla 6.2 resume los resultados.

Métrica	Valor
Área bajo la curva ROC (AUC), media	0,993
Desviación entre pliegues	± 0,005

AUC por pliegue	0,988 · 0,999 · 1,000 · 0,991 · 0,988
Puntuación de Brier (calibración)	0,0258
Observaciones / usuarios	1.320 / 40

Tabla 6.2. Métricas del clasificador de riesgo de iliquidez (validación por grupos).

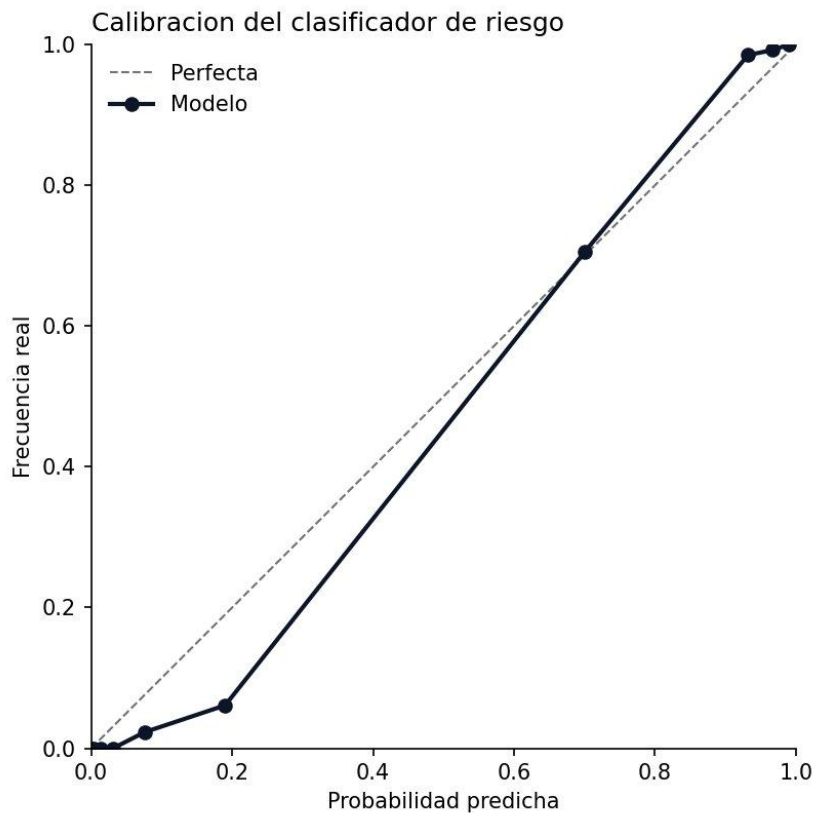


Figura 6.2. Curva de calibración del clasificador de riesgo. La proximidad a la diagonal indica que las probabilidades predichas se corresponden con las frecuencias reales observadas.

6.3.1 ANÁLISIS CRÍTICO

El modelo alcanza una capacidad de discriminación muy elevada (AUC de 0,993) y una excelente calibración (puntuación de Brier de 0,0258). Un valor de AUC tan próximo a la unidad obliga a replantearse si es correcto: en la práctica, una cifra semejante suele ser señal de fuga de información o de sobreajuste, defectos que invalidarían el resultado. Conviene, por tanto, justificar por qué en este caso no lo es.

La razón es de naturaleza financiera. La variable más determinante del modelo es el colchón de liquidez del usuario, es decir, los meses de gasto que su saldo actual es capaz de cubrir. Resulta lógico que quien dispone de un colchón muy reducido tenga una probabilidad muy alta de incurrir en iliquidez a corto plazo, y que quien dispone de un colchón holgado tenga una probabilidad muy baja. No se trata de fuga de información, porque el modelo emplea exclusivamente variables del presente para anticipar un evento futuro, y la validación por grupos garantiza que los usuarios de evaluación nunca se emplearon en el entrenamiento, lo que descarta la memorización. La elevada AUC refleja que la iliquidez a tres meses es en gran medida predecible a partir de la situación actual.

Para distinguir la contribución real del Random Forest de la de una regla trivial, se evalúan tres configuraciones: un umbral simple sobre el colchón de liquidez, el modelo sin dicha variable y el modelo completo. La Tabla 6.3 recoge los resultados.

Modelo	AUC	Brier
Umbral simple sobre el colchón de liquidez	0,995	0,170
Random Forest sin colchón	0,928	0,101
Random Forest completo	0,993	0,026

Tabla 6.3. Prueba de ablación del clasificador de riesgo de iliquidez.

La ablación confirma que el colchón de liquidez concentra la mayor parte del poder discriminativo: un umbral simple sobre esta variable alcanza un AUC de 0,995, prácticamente idéntico al del modelo completo (0,993). Sin embargo, la prueba revela también el valor real del Random Forest en la calibración de probabilidades: una puntuación de Brier de 0,026 frente a 0,170 del umbral trivial. El motor prescriptivo no consume etiquetas binarias, sino probabilidades continuas que actúan como restricción de riesgo; en ese contexto, la diferencia en calibración es decisiva. El modelo sin colchón obtiene un AUC de 0,928 y un Brier de 0,101, lo que confirma que el resto de variables contribuyen tanto a la discriminación residual como a la calibración, aunque el colchón es el factor dominante. Este hallazgo, que el patrón es casi lineal en una variable, pero la estimación fiable de la

probabilidad requiere el modelo completo, constituye en sí mismo un resultado metodológico honesto.

6.4 GEMELO DE MONTE CARLO

La fiabilidad del gemelo se evalúa comparando las probabilidades de iliquidez que predice con la frecuencia con la que esos eventos ocurren realmente. Esta comprobación se realiza sobre 150 casos, cada uno de los cuales enfrenta una predicción del gemelo con el resultado observado. La Tabla 6.4 resume el resultado.

Métrica	Valor
Puntuación de Brier	0,0794
Frecuencia real de iliquidez	40,0 %
Probabilidad media estimada por el gemelo	31,2 %
Puntos de contraste	150

Tabla 6.4. Backtesting de la probabilidad de iliquidez del gemelo de Monte Carlo.

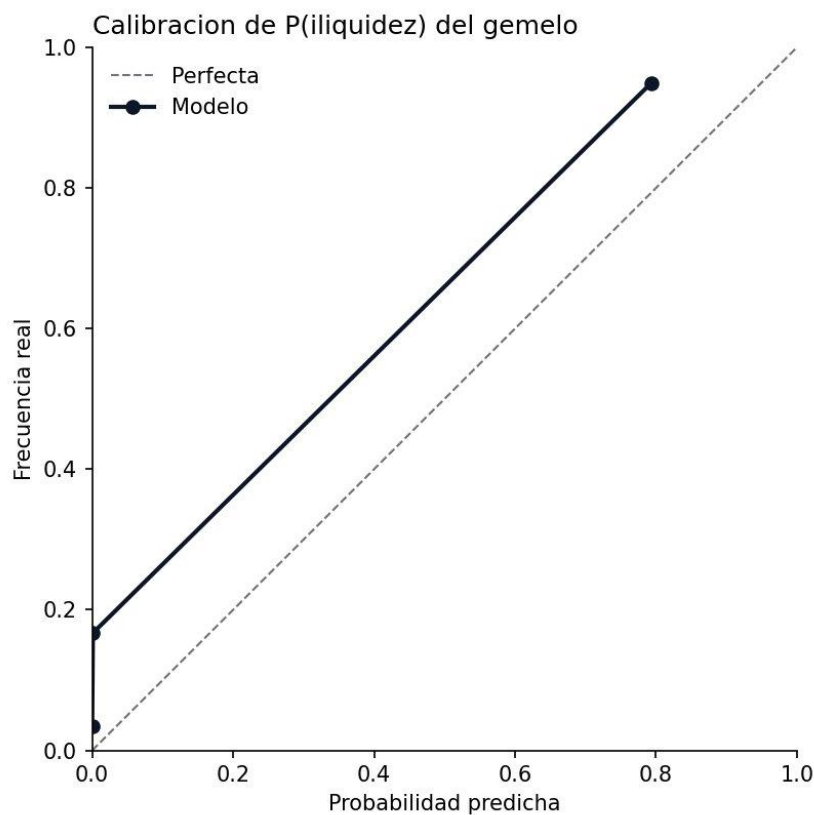


Figura 6.3. Curva de calibración de la probabilidad de iliquidez del gemelo. La frecuencia real observada se sitúa por encima de la probabilidad estimada.

6.4.1 ANÁLISIS CRÍTICO

El gemelo presenta una calibración razonable (puntuación de Brier de 0,0794), inferior a la del clasificador de riesgo. Esto es esperable: el clasificador solo tiene que evaluar la situación presente del usuario para emitir una única predicción, mientras que el gemelo debe simular la evolución completa del saldo mes a mes a lo largo de todo el horizonte. Reconstruir miles de trayectorias futuras es una tarea intrínsecamente más difícil que clasificar un estado actual, ya que los pequeños errores de cada mes se acumulan a lo largo de la simulación.

El análisis revela, además, un sesgo que conviene declarar con transparencia: la probabilidad media de iliquidez estimada por el gemelo (31,2 %) es inferior a la frecuencia real observada (40,0 %). Es decir, el gemelo tiende a subestimar ligeramente el riesgo de quedarse sin liquidez. Para una herramienta de prevención financiera, este sesgo es el menos deseable,

pues podría transmitir al usuario una confianza superior a la justificada. Su causa probable es que las distribuciones de gasto ajustadas no capturan plenamente los eventos extremos de la cola, y su corrección, se plantea como mejora futura.

6.5 MOTOR PRESCRIPTIVO DE RECOMENDACIÓN

La evaluación del motor prescriptivo es de naturaleza distinta: no se mide con una métrica de error, sino por su capacidad de generar planes de acción válidos, cuantificados y coherentes con la restricción de riesgo impuesta por el usuario. Para el usuario de demostración, con un horizonte de veinticuatro meses y una tolerancia al riesgo del 30 %, el motor evalúa 279 combinaciones de acciones y recomienda un plan que permite acumular 7.392 € de ahorro protegido manteniendo la probabilidad de iliquidez en el 0 %, con un valor en riesgo favorable de 3.592 €. El plan se acompaña de una explicación en lenguaje natural y de alternativas ordenadas, de modo que el usuario conserva la capacidad de decisión.

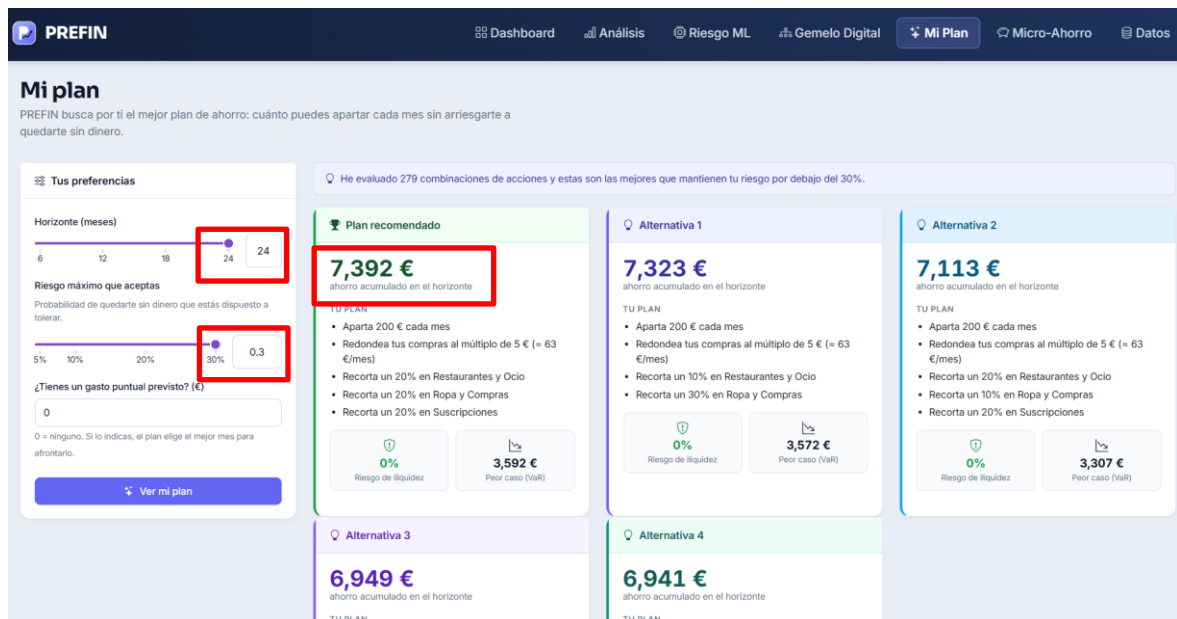


Figura 6.4. Pestaña "Mi Plan": planes de ahorro recomendados por el motor prescriptivo para el usuario de demostración (horizonte de 24 meses, tolerancia al riesgo del 30 %).

6.5.1 ANÁLISIS CRÍTICO

El resultado demuestra que el motor cumple su función: encuentra planes que respetan la restricción de riesgo y los presenta de forma comprensible y accionable, materializando el salto de la analítica predictiva a la prescriptiva. No obstante, debe señalarse que la calidad de las recomendaciones está acotada por la del gemelo que las evalúa: dado que el gemelo tiende a subestimar el riesgo (Sección 6.4), los planes recomendados podrían ser, en la realidad, ligeramente menos seguros de lo que indican sus métricas. La recomendación es, por tanto, tan fiable como el modelo del mundo sobre el que se apoya, lo que refuerza la importancia de la calibración del gemelo como prioridad de mejora.

6.6 *DISCUSIÓN DE LIMITACIONES*

El análisis anterior permite reunir las principales limitaciones del trabajo. Reconocerlas de forma explícita no resta valor al proyecto, sino que forma parte del rigor que se espera de una evaluación seria.

La primera limitación es que todas las conclusiones se han obtenido con datos sintéticos, es decir, generados artificialmente y no procedentes de cuentas bancarias reales. Aunque el generador imita patrones realistas y la validación se ha hecho sobre usuarios distintos de los empleados para entrenar, lo que evita resultados engañosamente buenos, los datos sintéticos no equivalen a los reales: los resultados son válidos para los datos generados, pero su aplicación a usuarios reales debe tomarse con prudencia. Probar el sistema con datos bancarios reales sería el siguiente paso lógico, aunque no es posible en un trabajo académico por los requisitos legales que exige el acceso a la banca abierta.

La segunda limitación afecta a la previsión de gasto: iguala, pero no supera, al método de referencia en el acierto del valor exacto, y además sus bandas de incertidumbre son demasiado estrechas (aciertan el 64 % de las veces, cuando deberían acertar el 80 %). La tercera afecta al gemelo, que tiende a subestimar un poco la probabilidad de quedarse sin dinero, un sesgo especialmente delicado en una herramienta cuyo objetivo es prevenir. Las

dos limitaciones están medidas, identificadas y tienen solución conocida mediante técnicas de calibración, que se proponen como trabajo futuro en el *Capítulo 7*.

En conjunto, los resultados respaldan la idea central del proyecto: PREFIN mide la incertidumbre, evalúa el riesgo con rigor y recomienda decisiones de forma transparente. Las limitaciones señaladas no anulan estos logros, sino que delimitan con honestidad hasta dónde llegan y marcan el camino para mejorarlos.

Capítulo 7. CONCLUSIONES Y TRABAJOS FUTUROS

Este capítulo recapitula el trabajo realizado, evalúa el cumplimiento de los objetivos planteados en un inicio, expone las aportaciones del proyecto y traza las líneas de desarrollo futuro.

7.1 CONCLUSIONES GENERALES

El presente proyecto partía de una clara observación: la banca digital ha multiplicado la información que está a disposición del usuario, pero no su capacidad de comprenderla y actuar sobre ella. PREFIN se concibió para cubrir esa brecha, y el resultado es una plataforma que no se limita a mostrar datos, sino que los interpreta, anticipa su evolución, recomienda decisiones y explica sus conclusiones en un lenguaje accesible.

La conclusión principal es que se ha logrado el objetivo de fondo del proyecto: elevar la plataforma desde un nivel descriptivo a uno prescriptivo. PREFIN no solo responde a la pregunta de cómo está la situación financiera del usuario, sino también a la de qué puedo hacer para mejorarla, cuantificando el impacto de cada decisión bajo incertidumbre. Este salto de describir a recomendar constituye la diferencia esencial entre un simple Dashboard financiero y un sistema que razona, y es lo que dota al proyecto de su entidad y originalidad.

7.2 CUMPLIMIENTO DE LOS OBJETIVOS

Los objetivos planteados en el *Capítulo 4* se han cumplido en su totalidad, y en el caso de la dimensión prescriptiva se ha ido más allá de lo previsto inicialmente. La Tabla 7.1 resume el grado de consecución de cada uno.

Objetivo	Grado de cumplimiento
Análisis de los datos financieros	Cumplido. El sistema integra tres fuentes de datos en un esquema común, categoriza las transacciones, calcula los indicadores clave y detecta gastos inusuales.

Evaluación del riesgo, simulación y recomendación prescriptiva	Cumplido y ampliado. El modelo de riesgo se reformuló para predecir un evento futuro real, y el gemelo determinista se sustituyó por uno estocástico de Monte Carlo capaz de cuantificar la incertidumbre. Se desarrolló un motor que, sobre el gemelo, recomienda el mejor plan de acción bajo una restricción de riesgo, materializando el salto a la analítica prescriptiva.
Interfaz accesible	Cumplido. La aplicación ofrece una interfaz web cuidada, con lenguaje claro, explicaciones en lenguaje natural y un modo de demostración.

Tabla 7.1. Grado de cumplimiento de los objetivos del proyecto.

El proyecto comenzó como un cuadro de mando financiero con aprendizaje automático y evolucionó, durante su desarrollo, hacia un sistema que recomienda decisiones y que corrige un defecto metodológico de fondo en su modelo de riesgo. Esta evolución no estaba contemplada en la planificación original y constituye la principal señal de ambición del trabajo.

7.3 APORTACIONES DEL PROYECTO

Las aportaciones más relevantes de PREFIN son las siguientes.

La primera es el salto de la analítica descriptiva a la prescriptiva en el ámbito de las finanzas personales. Frente a las soluciones del mercado, que en su mayoría se limitan a describir el gasto pasado, PREFIN recomienda activamente cursos de acción y cuantifica su impacto esperado, manteniendo el riesgo bajo control.

La segunda es el tratamiento riguroso de la incertidumbre. Tanto la previsión de gasto, mediante bandas de probabilidad, como el gemelo estocástico, mediante simulación de Monte Carlo, ofrecen al usuario una imagen honesta de los escenarios posibles y de sus probabilidades.

La tercera es de carácter metodológico: la identificación y corrección de la circularidad del modelo de riesgo. Reformular la etiqueta como un evento futuro observado, y validar el modelo sobre usuarios nunca vistos, transforma un clasificador que se limitaba a reproducir

una regla en un modelo predictivo genuino. Esta corrección, además de su valor técnico, refleja una actitud crítica hacia el propio trabajo.

La cuarta es la transparencia: la incorporación de explicaciones en lenguaje natural, fundamentadas en los datos del usuario, acerca la inteligencia del sistema a personas sin formación financiera, que son precisamente su público objetivo.

7.4 TRABAJOS FUTUROS

El análisis crítico de los resultados, recogido en el *Capítulo 6*, orienta con claridad las líneas de mejora prioritarias.

La más importante es la validación con datos bancarios reales. Todo el trabajo se ha desarrollado y evaluado sobre datos sintéticos, lo que, pese al rigor de la validación, limita la extrapolación de las conclusiones. Disponer de datos reales, para lo que sería necesario el marco regulatorio adecuado, permitiría confirmar el comportamiento del sistema en condiciones reales.

En segundo lugar, mejorar la calibración de la incertidumbre, es decir, conseguir que las probabilidades y los rangos que ofrece el sistema se ajusten mejor a lo que ocurre en la realidad. Esto afecta a dos componentes. Por un lado, las bandas de la previsión de gasto son demasiado estrechas: prometen más seguridad de la que realmente tienen. Existen técnicas, como la predicción conforme, que permiten ensancharlas hasta que su fiabilidad coincida con la esperada. Por otro lado, el gemelo de Monte Carlo subestima un poco la probabilidad de quedarse sin dinero; este sesgo podría corregirse haciendo que el modelo contemple con más frecuencia los escenarios extremos, o aplicando un margen de seguridad. Como el motor de recomendación se apoya en el gemelo para evaluar cada plan, mejorar la precisión del gemelo haría automáticamente más fiables las recomendaciones.

En tercer lugar, como ampliación de mayor alcance, el motor prescriptivo podría evolucionar de una búsqueda sobre el simulador hacia un agente de aprendizaje por refuerzo que aprenda políticas de recomendación adaptadas al perfil de cada usuario, empleando el propio gemelo

como entorno de entrenamiento. Asimismo, la conexión efectiva con la banca abierta en un entorno de producción completaría el ciclo previsto en el diseño.

7.5 CONCLUSIÓN FINAL

En conjunto, PREFIN demuestra que es posible construir una herramienta de finanzas personales que no se limite a informar, sino que razone y recomiende, y hacerlo además de forma transparente y accesible. Las limitaciones identificadas, lejos de invalidar el trabajo, delimitan con honestidad su alcance y señalan un camino claro de consolidación. El proyecto cumple sus objetivos, los supera en su dimensión más ambiciosa y sienta una base sólida sobre la que seguir construyendo.

Capítulo 8. BIBLIOGRAFÍA

- Angelopoulos, A. N., & Bates, S. (2023). Conformal prediction: A gentle introduction. *Foundations and Trends in Machine Learning*, 16(4), 494–591. <https://doi.org/10.1561/22000000101>
- Bertsimas, D., & Kallus, N. (2020). From predictive to prescriptive analytics. *Management Science*, 66(3), 1025–1044. <https://doi.org/10.1287/mnsc.2018.3253>
- Deloitte. (2022). Global Mobile Consumer Survey: Financial Services Report. Deloitte Insights. <https://www2.deloitte.com>
- Eurostat. (2023). Inflation and wage dynamics in the EU. Statistical Office of the European Union. <https://ec.europa.eu/eurostat>
- Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica*, 46(1), 33–50. <https://doi.org/10.2307/1913643>
- Lusardi, A. y Mitchell, O. S. (2014). The Economic Importance of Financial Literacy: Theory and Evidence. *Journal of Economic Literature*, 52(1), 5–44. <https://doi.org/10.1257/jel.52.1.5>
- OCDE. (2020). OECD/INFE 2020 International Survey of Adult Financial Literacy. OECD Publishing. <https://www.oecd.org>
- OCDE. (2023). Supporting Informed and Safe Use of Short-Term Online Credit and Buy Now Pay Later through Digital Financial Literacy. OECD Publishing. <https://doi.org/10.1787/finance>
- Banco de España y Comisión Nacional del Mercado de Valores. (2018). *Plan de Educación Financiera 2018-2021*. <https://www.cnmv.es/DocPortal/publicaciones/PlanEducacion/PEF2018-2021.pdf>
- Ramsey, D. (2013). *The Total Money Makeover: A Proven Plan for Financial Fitness*. Thomas Nelson.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>

- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5), 1189–1232. <https://doi.org/10.1214/aos/1013203451>
 - Hardt, D. (Ed.). (2012). The OAuth 2.0 Authorization Framework. RFC 6749. Internet Engineering Task Force. <https://datatracker.ietf.org/doc/html/rfc6749>
 - Hipp, D. R. (2000). SQLite. <https://www.sqlite.org>
 - Negri, E., Fumagalli, L. y Macchi, M. (2017). A review of the roles of digital twin in CPS-based production systems. *Procedia Manufacturing*, 11, 939–948. <https://doi.org/10.1016/j.promfg.2017.07.198>
 - Parlamento Europeo y Consejo de la Unión Europea. (2015). Directiva (UE) 2015/2366 sobre servicios de pago en el mercado interior (PSD2). Diario Oficial de la Unión Europea. <https://eur-lex.europa.eu>
 - Plotly Technologies. (2017). Dash: A Python framework for building analytical web applications. <https://dash.plotly.com>
8. Ramírez, S. (2018). FastAPI. <https://fastapi.tiangolo.com>
- TrueLayer. (2024). TrueLayer Open Banking API documentation. <https://docs.truelayer.com>
 - Emma Technologies. (2024). Emma App. <https://emma-app.com>
9. Fintonic. (2024). Funciones y servicios. <https://www.fintonic.com>
- Monefy. (2024). Monefy Budget App. <https://www.monefy.me>
 - Plum Fintech. (2024). Plum App. <https://withplum.com>
 - YNAB. (2024). You Need A Budget. <https://www.youneedabudget.com>
 - Betterment. (2024). Betterment: The intelligent money manager. <https://www.betterment.com>
10. Finizens. (2024). Finizens: Gestión automatizada de inversiones. <https://finizens.com>
- Indexa Capital. (2024). Indexa Capital: Gestión indexada. <https://indexacapital.com>
 - Wealthfront. (2024). Wealthfront: Automated investing and financial planning. <https://www.wealthfront.com>
 - Cleo. (2024). *Cleo: Your AI financial assistant*. <https://web.meetcleo.com>
 - Plaid. (2024). *Plaid: The network that powers fintech*. <https://plaid.com>

- Naciones Unidas. (2015). Transformar nuestro mundo: la Agenda 2030 para el Desarrollo Sostenible (Resolución A/RES/70/1). Asamblea General de las Naciones Unidas. <https://www.un.org/sustainabledevelopment/es/>

ANEXO I: ALINEACIÓN DEL PROYECTO CON LOS ODS

Los Objetivos de Desarrollo Sostenible (ODS), aprobados por la Asamblea General de las Naciones Unidas en 2015 dentro de la Agenda 2030, constituyen un conjunto de diecisiete metas globales orientadas a erradicar la pobreza, proteger el planeta y garantizar la prosperidad de las personas. Aunque PREFIN es un proyecto de naturaleza tecnológica, su propósito, ayudar a las personas a comprender y mejorar su situación financiera, lo vincula directamente a varios de estos objetivos, especialmente con aquellos relacionados con la reducción de la pobreza, la educación y la igualdad de oportunidades. A continuación, se detallan los ODS con los que el proyecto presenta una alineación más significativa.

8.1 ODS 1. FIN DE LA POBREZA

Uno de los factores que agravan la vulnerabilidad económica de los hogares es la incapacidad de anticipar situaciones de falta de liquidez. PREFIN contribuye directamente a la meta 1.5 de este objetivo, fomentar la resiliencia de las personas en situación de vulnerabilidad frente a perturbaciones económicas, al permitir prever el riesgo de quedarse sin dinero y recomendar medidas concretas para evitarlo. Al transformar la prevención financiera en una herramienta accesible, el proyecto ayuda a que los usuarios eviten caer en descubiertos, deudas o situaciones de impago que pueden desencadenar dificultades económicas mayores.

8.2 ODS 4. EDUCACIÓN DE CALIDAD

La educación financiera es una carencia extendida, especialmente entre los jóvenes y las personas sin formación específica, que son el público objetivo de PREFIN. El proyecto contribuye a la meta 4.4, aumentar las competencias necesarias para la vida, al funcionar no solo como una herramienta de gestión, sino como un instrumento pedagógico: sus

explicaciones en lenguaje natural traducen conceptos financieros complejos a un lenguaje comprensible, ayudando al usuario a entender por qué su situación es la que es y qué consecuencias tienen sus decisiones. De este modo, además de informar, PREFIN forma al usuario en la comprensión de sus propias finanzas.

8.3 ODS 10. REDUCCIÓN DE LAS DESIGUALDADES

El acceso a herramientas avanzadas de análisis y planificación financiera ha estado tradicionalmente reservado a quienes pueden permitirse un asesor o pagar servicios de pago. PREFIN contribuye a la meta 10.2, promover la inclusión de todas las personas con independencia de su situación económica, al ofrecer de forma gratuita capacidades de predicción, simulación y recomendación que en el mercado tienen un coste mensual. Al democratizar el acceso a estas herramientas, el proyecto reduce la brecha entre quienes disponen de recursos para gestionar bien su dinero y quienes no, contribuyendo a una mayor igualdad de oportunidades económicas.

8.4 ODS 8. TRABAJO DECENTE Y CRECIMIENTO ECONÓMICO

De forma complementaria, PREFIN se alinea con la meta 8.10, fortalecer la capacidad de las personas para acceder a los servicios financieros y gestionarlos. Al fomentar hábitos de ahorro, la planificación a medio plazo y la gestión responsable de los ingresos y gastos, el proyecto promueve la salud financiera individual, que constituye una base para la estabilidad económica personal y, de forma agregada, para un crecimiento económico más sólido e inclusivo.

ANEXO II

ANEXO II.1 — INTERFAZ DE LA APLICACIÓN

Este anexo recoge las capturas de las principales vistas de la interfaz de PREFIN que no se han incluido en el cuerpo de la memoria, con el fin de ofrecer una visión completa de la aplicación. Las vistas correspondientes a la capa de inteligencia (riesgo, gemelo digital, plan prescriptivo y previsión) se muestran en el *Capítulo 5*.

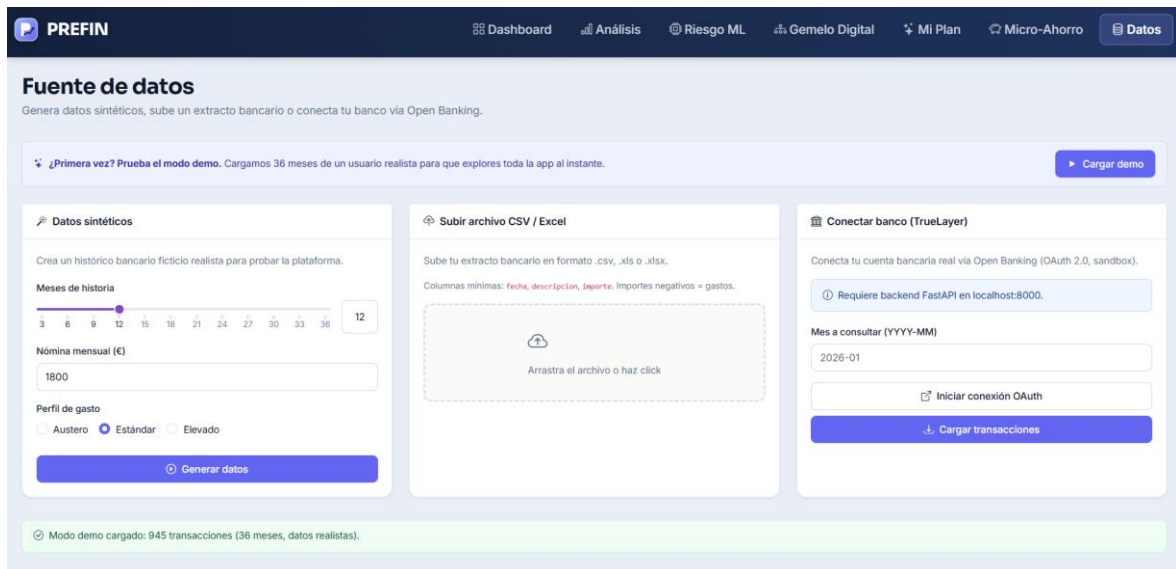


Figura A.1. Pestaña de gestión de datos con las tres fuentes: generación sintética, carga de extractos y conexión bancaria.

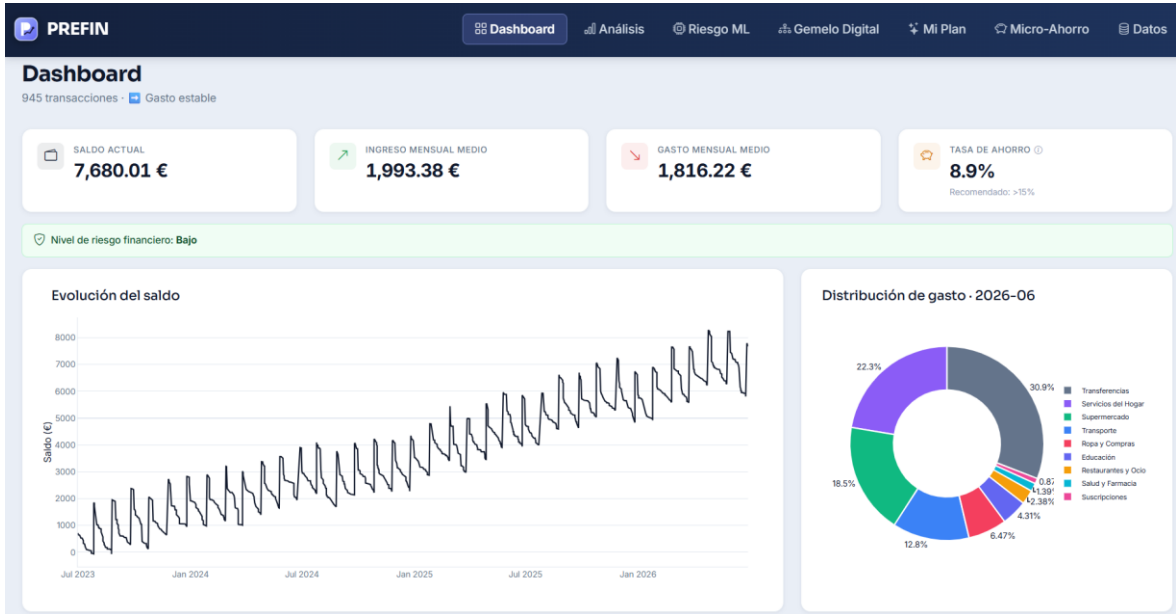


Figura A.2. Cuadro de mando principal con los indicadores clave y la evolución del saldo.

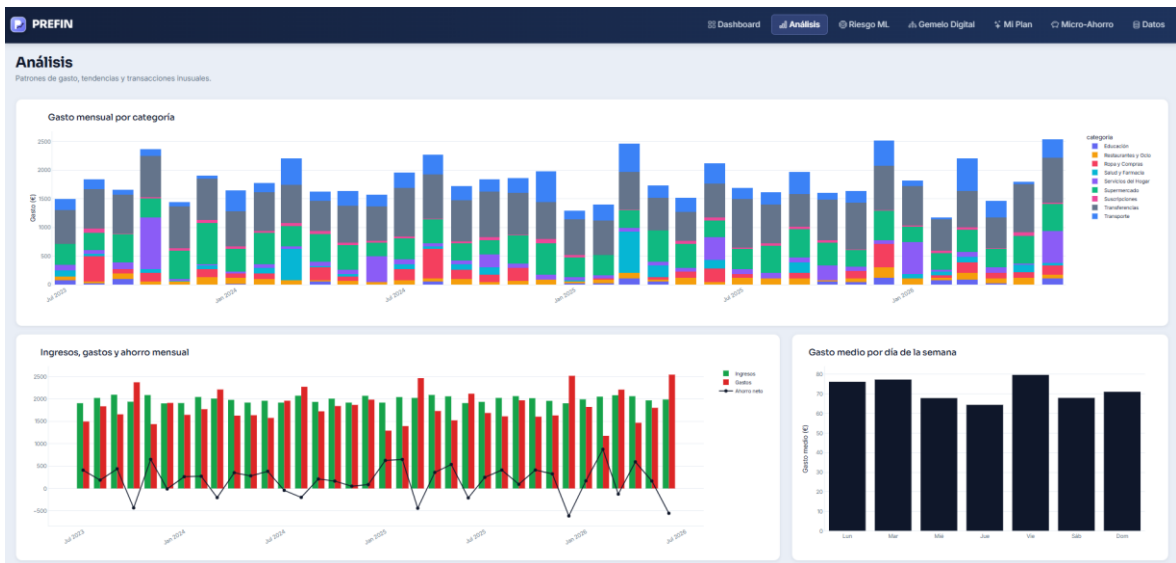


Figura A.3. Pestaña de análisis: gasto por categoría, balance mensual y patrón por día de la semana.

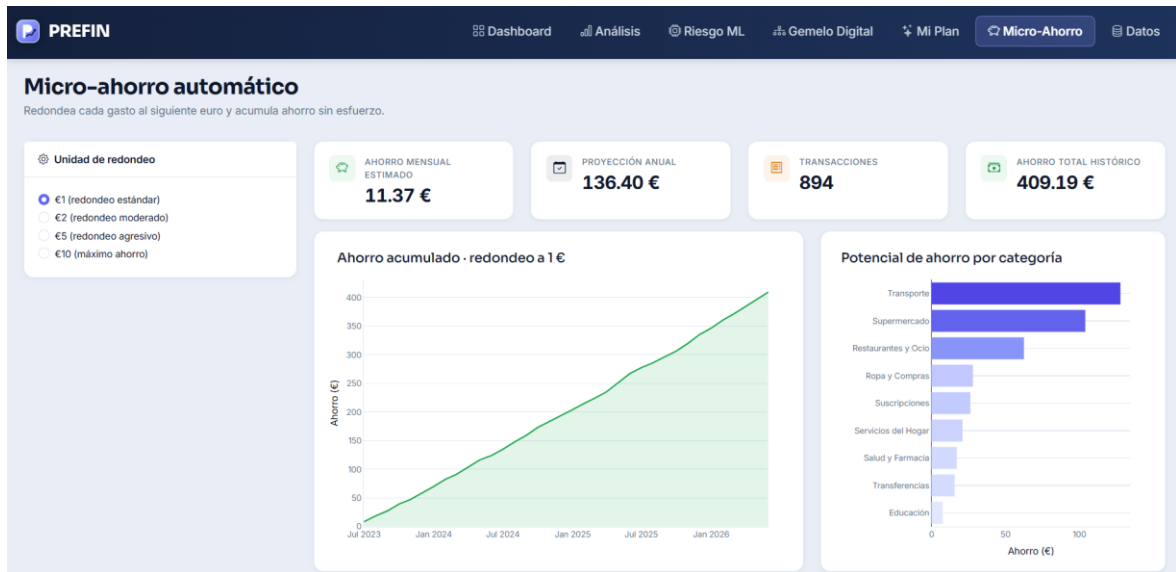


Figura A.4. Pestaña de micro-ahorro: ahorro acumulado por redondeo y potencial por categoría.

ANEXO II.2 — ACCESO AL CÓDIGO FUENTE

El código fuente completo de PREFIN se encuentra disponible en el repositorio del proyecto <https://github.com/paofdezcheca/PREFIN>. El sistema está organizado en módulos independientes de Python (ingesta, categorización, análisis, previsión, gemelo digital, motor prescriptivo, detección y explicabilidad), cuya función se detalla en el *Capítulo 5*. Por su extensión, no se reproduce aquí, en línea con las recomendaciones de la normativa sobre la elaboración de anexos.