

TFM Nuclio Digital School - Proyecto easyMoney

Estudiantes:

- Raphael Cheves
- Juan Manuel Guatta
- Miguel Angel García
- Daniel Guidi

Fecha de entrega y defensa: Jueves 23 de Octubre de 2025

Programa cursado: Máster en Data Science & AI

Tutor: Ignacio Anguita

ÍNDICE

1. INTRODUCCIÓN	3
2. JUSTIFICACIÓN Y CONTEXTUALIZACIÓN	3
3. OBJETIVOS	4
3.1. Objetivo General	4
3.2. Objetivos Específicos	5
4. METODOLOGÍA EMPLEADA	6
4.1. Enfoque general del proyecto	6
4.2. Comprensión inicial y preparación de los datos	6
4.3. Modelado predictivo y analítica supervisada	7
4.4. Segmentación de clientes mediante K-Means	8
4.5. Generación de propuestas de negocio	8
4.6. Organización y metodología de trabajo en equipo	8
5. RESULTADOS Y CONCLUSIONES	9
5.1. Presentación general de resultados	9
5.2. Discusión de resultados por campaña	10
5.3. Evaluación del cumplimiento de objetivos	12
5.4. Escenarios de sensibilidad: pesimista, realista y optimista	13
5.5. Supuestos y racionalidad de costes	13
5.6. Recomendaciones para implementación y mejora continua	14
5.7. Conclusión final	14

1. INTRODUCCIÓN

El presente proyecto se enmarca dentro del Máster de Data Science & AI y busca reproducir un escenario realista de consultoría y desarrollo analítico en una empresa del sector financiero digital. A través del caso easyMoney, se plantea a los estudiantes la necesidad de adoptar el rol de Data Scientist en un contexto empresarial dinámico, con retos estratégicos y comerciales que exigen la aplicación de técnicas avanzadas de analítica de datos, machine learning y visualización.

La situación de partida refleja una problemática habitual en compañías de reciente creación que, tras un período inicial de crecimiento impulsado por la captación de clientes, deben orientar su estrategia hacia la rentabilización de la base de usuarios existente. En este sentido, easyMoney enfrenta presiones tanto internas como externas: las tensiones con su socio estratégico (easyBanking), la necesidad de responder a los inversores con resultados positivos en el corto plazo y las limitaciones tecnológicas derivadas de una rotación elevada en el equipo de IT.

El rol del equipo de Data Science, representado en este proyecto, consiste en aportar soluciones analíticas que faciliten la toma de decisiones comerciales. Concretamente, el trabajo desarrollado se centra en tres grandes ejes:

- a. **Generación de visibilidad de negocio** mediante un dashboard en PowerBI que permita explorar los datos de ventas y clientes, con indicadores clave (KPIs) y análisis detallados de evolución.
- b. **Modelado predictivo de propensión de compra**, con el objetivo de anticipar qué clientes tienen mayor probabilidad de contratar productos financieros específicos y, de esta manera, facilitar campañas comerciales más eficientes.
- c. **Segmentación de clientes** para identificar perfiles diferenciales dentro de la base de datos, posibilitando la personalización de ofertas y la optimización del impacto comercial.

La introducción de estos tres bloques permite abordar de manera integral el desafío planteado por la dirección de la empresa: incrementar la rentabilidad de la cartera actual de clientes con productos existentes, reduciendo la dependencia de costosas campañas de captación masiva.

2. JUSTIFICACIÓN Y CONTEXTUALIZACIÓN

El desarrollo de este proyecto responde a la necesidad de easyMoney de reorientar su estrategia en un momento crítico de su ciclo de vida empresarial. A diferencia de la fase inicial, caracterizada por un fuerte crecimiento en número de clientes gracias

a inversiones en marketing y a la innovación de su primer producto, la compañía se encuentra actualmente en una etapa donde la rentabilidad es la prioridad.

Los principales factores que justifican este estudio son los siguientes:

- a. **Relevancia del problema empresarial:** easyMoney enfrenta la presión de mostrar resultados financieros sostenibles en el corto plazo. Para ello, debe rentabilizar la base de clientes actual y evitar la dependencia excesiva de nuevas rondas de financiación. La analítica avanzada es una herramienta clave para diseñar campañas comerciales basadas en datos y maximizar el retorno de la inversión en marketing.
- b. **Contexto competitivo del sector financiero:** Las fintech compiten en un entorno de elevada innovación y bajo margen de error. La personalización de ofertas y la capacidad de anticipar el comportamiento del cliente son ventajas competitivas críticas. La aplicación de técnicas de machine learning y segmentación en este contexto permite alinear a easyMoney con las mejores prácticas del sector.
- c. **Antecedentes profesionales y académicos:** Este proyecto constituye una oportunidad para poner en práctica conocimientos adquiridos durante el máster en entornos de datos reales, utilizando librerías y modelos de machine learning para aprendizaje supervisado y no supervisado (XGBoost, CatBoost, KMeans, entre otros), así como herramientas de visualización como PowerBI.

En términos de contextualización, el caso se plantea a través de una serie de comunicaciones internas ficticias (emails de la Dirección de IT y de la CEO de la compañía), que introducen de forma narrativa los problemas de negocio. A partir de esas interacciones, el proyecto se estructura en tareas secuenciales: construcción de un dashboard, desarrollo de modelos de propensión, segmentación de clientes y análisis de escenarios de campañas comerciales.

La justificación final, por tanto, radica en la necesidad de generar un puente entre la visión estratégica de la compañía y las capacidades técnicas del Data Science, dotando a easyMoney de herramientas accionables que le permitan tomar decisiones fundamentadas en datos.

3. OBJETIVOS

3.1. Objetivo General

El objetivo general de este proyecto es apoyar a easyMoney en la definición y ejecución de una estrategia comercial basada en datos que incremente la

rentabilidad de la base de clientes actual mediante el uso de analítica avanzada y visualización interactiva.

3.2. Objetivos Específicos

Este objetivo general se desglosa en los siguientes objetivos específicos:

Construir un dashboard de negocio en PowerBI que proporcione al área comercial visibilidad clara y detallada de la situación actual de clientes y ventas. Este dashboard permitirá monitorizar KPIs clave, analizar la evolución temporal de contrataciones de productos y realizar exploraciones profundas sobre características de clientes.

- a. **Desarrollar modelos de propensión de compra** para al menos cinco productos financieros estratégicos de easyMoney (cuentas, cuentas crypto, planes de pensión, depósitos a largo plazo y tarjetas de débito). Estos modelos deben entregar probabilidades individuales por cliente, que puedan ser utilizadas para priorizar campañas y reducir el costo de adquisición por venta.
- b. **Implementar una segmentación de clientes mediante clustering**, con el fin de identificar grupos homogéneos en función de variables demográficas, de comportamiento y de riesgo. La segmentación permitirá comprender mejor la heterogeneidad de la cartera y ajustar tanto el diseño de productos como las campañas comerciales.
- c. **Proponer un caso de uso de campaña comercial**, integrando los resultados de propensión y segmentación. La finalidad de este objetivo es traducir los hallazgos analíticos en una simulación práctica de negocio: Estimar el tamaño de la población objetivo, la tasa de respuesta esperada y los ingresos potenciales que se generarían.
- d. **Validar la utilidad del enfoque analítico en el contexto real de easyMoney**, evaluando si las herramientas desarrolladas permiten responder a las inquietudes planteadas por la dirección: Mejorar la penetración de productos en clientes actuales, incrementar la rentabilidad y apoyar la toma de decisiones estratégicas.

En conjunto, estos objetivos no solo buscan aportar valor inmediato en términos de métricas de negocio, sino también dotar a easyMoney de una cultura de decisión basada en datos, alineada con los estándares de las fintech más avanzadas del sector.

4. METODOLOGÍA EMPLEADA

4.1. Enfoque general del proyecto

El presente proyecto se desarrolló bajo un enfoque teórico–práctico, combinando la aplicación de técnicas científicas de análisis de datos con la resolución de un caso de negocio realista en el ámbito fintech. Si bien la base de datos fue proporcionada por el equipo docente, la interpretación de los requerimientos, el diseño de la propuesta analítica y la toma de decisiones estratégicas se realizaron con una mirada aplicada al negocio, simulando un entorno corporativo real.

La metodología integró dos dimensiones complementarias:

- Una dimensión técnica, basada en la aplicación rigurosa de métodos de data cleaning, modelado predictivo y segmentación estadística.
- Una dimensión organizativa y colaborativa, orientada a la gestión ágil del trabajo en equipo y la coordinación efectiva entre los miembros del grupo para abordar las distintas fases del proyecto.

Este enfoque permitió desarrollar un proyecto analítico completo: Desde la comprensión inicial de los datos hasta la generación de insights accionables y propuestas comerciales concretas, cumpliendo con los objetivos de negocio planteados por la dirección de easyMoney.

4.2. Comprensión inicial y preparación de los datos

El punto de partida del proyecto fue el análisis exploratorio de la base de datos facilitada (EDA). Dado que la recopilación de datos no formaba parte del alcance del caso, el equipo centró sus esfuerzos en la comprensión estructural de la información disponible: Identificación de tablas, relaciones, cardinalidades, campos clave, tipos de datos y presencia de valores nulos o inconsistentes.

Este proceso inicial permitió detectar la necesidad de realizar una limpieza y estandarización. Se llevaron a cabo las siguientes tareas:

- a. Eliminación de duplicados y verificación de consistencia entre tablas.
- b. Tratamiento de valores nulos, empleando estrategias como imputación por media, mediana a través de métodos avanzados por agrupaciones inteligentes basadas en variables relacionadas, de forma que no se alterara la distribución general de los datos.
- c. Definición del concepto de cliente único, elemento fundamental para definición del perímetro de cobertura del proyecto, garantizando la coherencia entre los distintos productos y análisis posteriores.

Una vez depurada la base de datos, el equipo realizó un análisis de penetración y rentabilidad por producto en la cartera actual, con el fin de identificar sobre qué líneas de negocio era más relevante enfocar los modelos predictivos. Este análisis condujo a la selección de cinco productos principales: ***em_acount***, ***emc_account***, ***pension_plan***, ***long_term_deposit*** y ***debit_card***.

4.3. Modelado predictivo y analítica supervisada

Con los productos definidos y la estructura de cliente consolidada, se procedió al desarrollo de los modelos de propensión de compra. El objetivo era estimar, para cada cliente, la probabilidad de adquirir un determinado producto financiero, lo cual permitiría orientar futuras campañas de marketing hacia segmentos de alta conversión.

Se diseñaron cinco modelos de clasificación independientes, uno por producto, aplicando un enfoque basado en machine learning supervisado.

Las etapas seguidas fueron las siguientes:

- a. **Reducción de categorías dispersas y agrupamiento** de valores poco representativos en categorías “Otros”.
- b. **Normalización y codificación de variables** mediante técnicas como One-Hot Encoding y Label Encoding, según el tipo de variable.
- c. **Balanceo de clases** mediante undersampling, para compensar la desproporción entre clientes que compraron y los que no.
- d. **División** del conjunto de datos en **particiones train y test**, garantizando independencia en la evaluación del modelo.
- e. **Entrenamiento competitivo de modelos**: Se probaron diferentes algoritmos (Logistic Regression, Gradient Boosting, CatBoost, XGBoost, etc), seleccionando el modelo con mejor desempeño en AUC sobre el conjunto de prueba.
- f. **Optimización de hiperparámetros** mediante Grid Search para mejorar la precisión y la estabilidad del modelo final.
- g. **Análisis de interpretabilidad usando SHAP** values y feature importance, permitiendo identificar los factores con mayor influencia en la probabilidad de compra.

Este proceso aseguró la rigurosidad estadística y la reproducibilidad de los resultados, cumpliendo con los principios básicos del método científico.

4.4. Segmentación de clientes mediante K-Means

El segundo eje metodológico consistió en la segmentación de la base de clientes mediante un enfoque no supervisado, con el fin de identificar grupos homogéneos de comportamiento y características. Para ello, se seleccionaron únicamente las métricas relevantes para la segmentación en un contexto fintech: variables demográficas, frecuencia de uso, monetarias, de riesgo y de comportamiento.

El procedimiento siguió los pasos siguientes:

- a. Normalización de variables numéricas para evitar sesgos de escala.
- b. Selección de características relevantes (feature engineering) con base en correlaciones y valor informativo con el fin de accionar sobre los clusters posterior a la segmentación.
- c. Análisis del número óptimo de clusters mediante el método del codo y la inspección visual del gráfico.
- d. Aplicación del algoritmo K-Means y validación de estabilidad de los grupos.
- e. Interpretación de segmentos en términos comerciales, asignando nombres descriptivos y fácilmente comunicables a los clusters.

El resultado fue una segmentación de 5 grupos de clientes, cada uno con características socioeconómicas y conductuales diferenciadas, lo que facilitó la creación de estrategias de marketing personalizadas.

4.5. Generación de propuestas de negocio

Una vez obtenidos los resultados de los modelos de propensión y segmentación, el equipo formuló una propuesta de negocio simulada, definiendo escenarios hipotéticos de campañas comerciales. Se estimaron impactos potenciales sobre clientes, tasas de conversión esperadas, productos recomendados y rentabilidad proyectada por campaña, con el fin de traducir los resultados analíticos en valor económico tangible.

Estos escenarios, aunque basados en supuestos razonables y experiencia práctica, no pretenden tener validez científica, sino ilustrar el potencial de la analítica predictiva para orientar decisiones estratégicas en una empresa fintech.

4.6. Organización y metodología de trabajo en equipo

El proyecto fue desarrollado por un equipo de cuatro integrantes, empleando metodologías ágiles y herramientas colaborativas.

La comunicación se gestionó a través de Slack (para la coordinación estructurada de tareas y priorización), WhatsApp (para comunicación inmediata y logística de reuniones) y Trello, donde se implementó un tablero con seis columnas que representaban el flujo de trabajo:

1. **Backlog:** Tareas identificadas pendientes de ejecución.
2. **Prioritized:** Tareas seleccionadas para el siguiente sprint.
3. **In Progress:** Tareas actualmente en desarrollo.
4. **Blocked / Needed Mentoring:** Tareas con dependencias o bloqueos requerían revisión por parte de otro miembro o del mentor.
5. **In Review:** Tareas completadas en proceso de revisión grupal.
6. **Completed:** Tareas finalizadas y validadas.

Los sprints se organizaron cada dos o tres semanas, dependiendo de la disponibilidad del equipo y la carga de otros entregables del máster.

La coordinación presencial se realizó en sesiones de trabajo en coworkings, donde se analizaban avances, comparaban resultados de modelos y se definían prioridades para el siguiente ciclo.

Aunque la mayor parte del trabajo se ejecutó de forma conjunta, se estableció una división temporal de tareas para acelerar el desarrollo: Dos miembros se encargaron del modelo de clustering y dos del dashboard descriptivo en PowerBI, unificando posteriormente los avances para mantener una visión integral del proyecto.

5. RESULTADOS Y CONCLUSIONES

5.1. Presentación general de resultados

A partir de la aplicación de los modelos de propensión de compra y la segmentación de clientes mediante técnicas de machine learning, se diseñó un conjunto de cinco campañas comerciales personalizadas que buscaban equilibrar la reactivación, la rentabilidad y la fidelización de clientes. Cada una fue diseñada para un cluster específico, con un producto, un objetivo comercial diferenciado y un canal de comunicación ajustado a su perfil.

En términos operativos, la estrategia combinó una campaña masiva por canales digitales (email/push) para el segmento de clientes pasivos y una campaña de televenta personalizada para los segmentos más rentables o estratégicos (iniciando por el segmento de Seniors Conservadores). Esta decisión permitió equilibrar

volumen de alcance y profundidad relacional, manteniendo la eficiencia de costes y el foco en el valor generado.

Los resultados se calcularon a partir de los supuestos de close rate (tasas de conversión) definidos en tres escenarios (pesimista, realista y optimista) y utilizando los net margins promedios por producto, junto con las estructuras de coste específicas de cada tipo de campaña. El escenario realista, considerado el más representativo, estima un **ROI total del 440,37%**, con más de **1,68 millones de euros en beneficios netos** generados en tan solo un mes de ejecución, considerando una base de 127.103 clientes contactados y una tasa promedio de conversión del 0,93%.

5.2. Discusión de resultados por campaña

Campaña 1 – “Reactiva tu cuenta” (Pasivos Masivos)

Objetivo: Reincorporar clientes inactivos mediante la oferta de productos base simples y de bajo coste (cuenta estándar, cuenta crypto o tarjeta de débito).

Canal: Email marketing masivo.

Resultados:

- Contactados: 213.203
- Convertidos: 1.706
- Beneficio neto: €81.107,19
- ROI: 415.66%

Pese a la baja rentabilidad individual de los productos, la masividad del segmento permitió generar un impacto positivo y escalable con un coste mínimo (€726,41). Se valida así la hipótesis de que el valor de este cluster radica en el volumen, no en la rentabilidad unitaria. Este segmento representa la puerta de entrada para futuras acciones de upselling, ya que una vez reactivados, los clientes pueden ser impactados con productos de mayor margen como depósitos o planes de pensiones.

Campaña 2 – “Planifica tu futuro hoy” (Vinculados Premium)

Objetivo: Profundizar la relación con clientes altamente vinculados y rentables mediante la oferta de planes de pensiones.

Canal: Televenta personalizada.

Resultados:

- Contactados: 3.884
- Convertidos: 311

- Beneficio neto: €1.293.683
- ROI: 454.48%

Fue la campaña más rentable del conjunto, capturando el 37.7% de los beneficios totales. Este éxito se explica por la alta propensión de compra del cluster, la elevada rentabilidad del producto ofrecido y la eficiencia del canal de venta directa. Reafirma la importancia de invertir esfuerzos personalizados en los segmentos premium, donde el lifetime value justifica los mayores costes de contacto.

Campaña 3 – “Rentabilidad con equilibrio” (Deudores Rentables)

Objetivo: Mantener el valor del segmento reduciendo riesgo de impago mediante la promoción de depósitos a largo plazo.

Canal: Televenta personalizada.

Resultados:

- Contactados: 2.284
- Convertidos: 183
- Beneficio neto: €187.747
- ROI: 419.5%

Los resultados confirman que los clientes con deuda moderada, pero alto engagement responden positivamente a ofertas de ahorro estructurado, lo que reduce su exposición crediticia sin comprometer la rentabilidad. Este segmento, aunque pequeño, es estratégico por su potencial de estabilización financiera y su contribución al equilibrio de riesgo global de cartera.

Campaña 4 – “Tranquilidad Garantizada” (Seniors Conservadores)

Objetivo: Fidelizar a clientes senior con productos previsionales de bajo riesgo, fortaleciendo la confianza y estabilidad a largo plazo.

Canal: Televenta personalizada.

Resultados:

- Contactados: 3.900
- Convertidos: 312
- Beneficio neto: € 1.299.284
- ROI: 455%

Este segmento, pese a su tamaño moderado, aportó un 37.8% del beneficio total, demostrando una fuerte afinidad con los productos de planificación financiera. El éxito se asocia a la coherencia entre perfil demográfico, producto ofrecido y tono

comunicacional de la campaña, validando la estrategia de personalización conductual.

Campaña 5 – “Tu Ahorro Inteligente” (Multivinculados Endeudados)

Objetivo: Mejorar la gestión de liquidez de clientes endeudados mediante depósitos a largo plazo con buena rentabilidad.

Canal: Televenta personalizada.

Resultados:

- Contactados: 6.897
- Convertidos: 552
- Beneficio neto: € 569.739
- ROI: 726.%

Al tener distintos segmentos con campañas variadas para cada uno de ellos se establece la ejecución de dos de ellas como primera etapa comercial para poder concentrar los esfuerzos y el foco en ejecución para garantizar el éxito inicial, para ello se escogieron las siguientes campañas como primera etapa:

- **Campaña 1 - Reactiva tu Cuenta:** Debido a su gran volumen tanto de potenciales nuevos clientes como de retorno generado (principalmente para iniciar la reactivación de la cartera para una posterior rentabilización a partir de distintas acciones de cross-selling en un futuro).
- **Campaña 4 - Tranquilidad Garantizada:** La campaña más rentable entre las campañas de telemarketing, elegida en la primera etapa debido a su alto nivel de rentabilidad y a la alineación del producto ofrecido con el segmento objetivo.

Resultado Total:

- Convertidos: 2.000 clientes
- Beneficio neto: € 1.380.000
- ROI: 452%

Los resultados revelan una respuesta sólida, con el 16.6% del beneficio total. La campaña logró conectar con un segmento sensible al equilibrio entre endeudamiento y rentabilidad, ofreciendo una vía de ahorro que refuerza la relación con la entidad.

5.3. Evaluación del cumplimiento de objetivos

En conjunto, las cinco campañas lograron cumplir y superar los objetivos planteados tanto a nivel de negocio como de validación analítica.

Se comprobó la validez práctica del modelo de propensión, al correlacionar altos scores de probabilidad con tasas efectivas de conversión elevadas.

Se verificó que una estrategia de segmentación dinámica y data-driven permite maximizar el retorno con inversiones relativamente bajas.

Además, se demostró que la combinación de canales (masivo para volumen + televenta para rentabilidad) es una fórmula óptima para equilibrar impacto y eficiencia.

5.4. Escenarios de sensibilidad: pesimista, realista y optimista

El análisis de escenarios permite dimensionar la robustez económica de la estrategia ante diferentes contextos de conversión:

Escenario	Beneficios Netos Totales €	Variación vs. Realista
Pesimista	512.418	-62,9%
Realista	1.380.392	—
Optimista	2.595.585	+88,0%

Las campañas diseñadas muestran un potencial significativo de generación de valor, con beneficios netos proyectados que **representan 17% comparado con** los 8,1 millones de euros obtenidos el mes anterior únicamente en reaprovechamiento de nuevos productos en la base de datos actual. Este incremento refleja la eficacia de una estrategia guiada por modelos de propensión de compra, que priorizan a clientes con alta probabilidad de conversión ($\geq 85\%$) y maximizan el retorno mediante la asignación eficiente de recursos a los segmentos con mayor rentabilidad esperada.

La combinación de productos adecuados por perfil, canales estratégicamente seleccionados y costes variables vinculados al éxito permite anticipar un impacto financiero sólido y sostenible en el tiempo.

5.5. Supuestos y racionalidad de costes

Para garantizar consistencia en las proyecciones, se establecieron supuestos que equilibran rigor analítico y realismo operativo:

Campañas de Email Marketing: Contemplan un cashback de 10 € por conversión, activado únicamente si la venta se concreta. Esto transforma el incentivo en un coste de éxito, reduciendo la exposición al riesgo.

Campañas de Telemarketing: Incluyen una comisión del 15 % sobre el valor de cada venta cerrada, además del coste fijo de personal. Este modelo alinea los incentivos del equipo comercial con los resultados.

Ambos enfoques comparten una lógica de coste variable condicionado, que protege el margen y mantiene la estructura financiera flexible ante escenarios adversos.

5.6. Recomendaciones para implementación y mejora continua

Para sostener la efectividad del modelo y afinar su capacidad predictiva, se recomiendan las siguientes acciones:

- **Monitorear el close rate real de cada campaña:** Es crucial medir en tiempo real la tasa de cierre (contactos efectivos / conversiones), tanto en canales masivos como en televentas. Este indicador permitirá ajustar rápidamente las tasas de propensión utilizadas en los modelos, mejorando la precisión de futuras campañas y evitando la sobreinversión en segmentos de baja respuesta.
- **Revisar la elasticidad del umbral de propensión** (actualmente en 85 %): Si las primeras campañas confirman un escenario optimista (close rates por encima del esperado), podría reducirse el umbral de propensión a 75–80 % para ampliar la base de clientes contactables. Este ajuste permitiría incrementar el volumen absoluto de conversiones y explotar el momentum comercial, sin comprometer significativamente el margen por cliente.
- **Evaluar la capacidad operativa de televentas:** En función de la tasa de contacto y conversión observada, ajustar el número de asesores o su carga mensual para maximizar eficiencia sin sobredimensionar el equipo.
- **Documentar aprendizajes y retroalimentar los modelos predictivos:** Cada campaña debe alimentar una base de datos de resultados reales (conversión, canal, tiempo de contacto, producto adquirido) que permita reentrenar los modelos y refinar la propensión de compra a futuro.

5.7. Conclusión final

La estrategia planteada combina inteligencia analítica, control financiero y ejecución comercial dirigida, demostrando que es posible incrementar el beneficio neto de la entidad sin asumir riesgos desproporcionados.

El modelo propuesto, con costes variables condicionados y decisiones basadas en datos, constituye una base sólida para una gestión de marketing eficiente, escalable y de bajo riesgo operativo.

A medida que se consoliden los primeros resultados, el siguiente paso será institucionalizar este enfoque como práctica recurrente de planificación comercial, integrando los modelos predictivos dentro de la toma de decisiones estratégica de la organización.