Universidad de los Andes

Facultad de Ingeniería y Ciencias Aplicadas

Aprendizaje automático causal para analítica de negocios: Métodos y aplicaciones novedosas

Catalina Belén Sánchez Cortés

Tesis para optar al grado de
doctorado en ciencias de la ingeniería

Profesor Guía: Carla Vairetti

Santiago, Mayo de 2026

Dedicatoria

Este trabajo está dedicado a mi familia, especialmente a mi marido, Tito, y a mi hijo que nació mientras cursaba este doctorado y que me han motivado para lograr todo lo que me propongo. A mis papás y hermanos que me han acompañado desde el inicio y a cada uno que no nombro pero que han estado ahí para mi.

Resumen

En la actualidad, las empresas enfrentan el desafío de transformar grandes volúmenes de datos en decisiones estratégicas que sean precisas desde el punto de vista estadístico y también rentables desde una perspectiva financiera. A pesar de los avances en el aprendizaje automático tradicional, existe una brecha significativa entre la capacidad predictiva y la optimización de acciones e intervenciones. Esta investigación aborda esta problemática mediante el desarrollo de un marco que integra la minería de datos, la inferencia causal y las métricas orientadas al beneficio, aplicadas en la gestión de cobranza y la venta cruzada.

En primera instancia se analiza la influencia de variables de *contact center* en la gestión de cobranza mediante la aplicación de modelos de aprendizaje automático tradicional. Se demuestra la aplicación de modelos de clasificación combinando la información financiera y de *contact center* alcanza un mejor rendimiento en términos de AUC y permite predecir comportamientos críticos como el contacto exitoso y la promesa de pago.

Posteriormente, se extiende este análisis a la aplicación de aprendizaje automático causal en cobranzas. A diferencia del enfoque tradicional, este permite identificar el efecto de la gestión de cobro sobre el deudor. Sobre esta base, se introducen marcos de evaluación basados en el beneficio (profit), tanto para modelos tradicionales como causales. A partir de esta aplicación se determina que mediante la implementación de una estrategia de profit multi-segmento, se puede gestionar la heterogeneidad de los montos adeudados y los costos de contacto, garantizando que la asignación de recursos maximice el retorno financiero de la empresa.

El marco de modelos de aprendizaje causal basado en beneficio también se aplica al contexto de venta cruzada, donde se adaptan las métricas de beneficio para modelar la compensación de los agentes de ventas. Los experimentos demuestran que la aplicación de clasificadores orientados al beneficio, en lugar de métricas estadísticas, permite alcanzar incrementos de hasta un 31.6 % en la utilidad promedio por intervención.

Finalmente, se propone una contribución metodológica, la estimación del efecto del tratamiento individual denominado Bounded ITE en modelos causales. Esta métrica introduce límites probabilísticos que permiten diferenciar a los clientes persuasibles para excluir a aquellos que la intervención no genera ningún impacto en su decisión. Este enfoque logra un rendimiento superior al enfoque ITE y la clasificación tradicional demostrando su eficiencia al centrarse en los casos donde se puede generar un impacto.

En base a estos análisis, esta investigación demuestra la aplicación de tecnologías de aprendizaje automático para la predicción, causalidad y el beneficio constituye un marco estratégico para la toma de decisiones empresarial.

Abstract

Companies face the challenge of transforming large volumes of data into strategic decisions that are statistically accurate and financially profitable. Despite advances in traditional machine learning, a significant gap remains between predictive capability and the optimization of actions and interventions. This research addresses this problem by developing a framework that integrates data mining, causal inference, and profit-driven metrics, applied to debt collection management and cross-selling.

First, the influence of contact center variables on debt collection management is analyzed using traditional machine learning models. The results demonstrate that the application of classification models combining financial and contact center information achieves superior performance in terms of AUC and enables the prediction of critical behaviors, such as successful contact and promise of payment.

Subsequently, this analysis is extended to the application of uplift modeling in debt collection. Unlike traditional approaches, this method allows for the identification of the specific effect of collection management on the debtor. Based on this, profit-driven evaluation frameworks are introduced for traditional and causal models. This application determines that by implementing a multi-segment profit strategy, it is possible to manage the heterogeneity of debt amounts and contact costs, ensuring that resource allocation maximizes the company's financial return.

The profit-driven uplift modeling framework is also applied to a cross-selling context, where profit metrics are adapted to model sales agent compensation. Experiments demonstrate that the application of profit-driven classifiers, rather than standard statistical metrics, leads to increases of up to 31.6 % in the average utility per intervention.

Finally, a methodological contribution is proposed: the estimation of the individual treatment effect termed "Bounded ITE" for uplift models. This metric introduces probabilistic bounds that allow for the identification of "persuadable" customers while excluding those for whom the intervention generates no impact on their decision. This approach achieves superior performance compared to the standard ITE approach and traditional classification, demonstrating its efficiency by focusing on cases where an actual impact can be generated.

Based on these analyses, this research demonstrates that the application of machine learning technologies for prediction, causality, and profit constitutes a strategic framework for business decision-making.

Agradecimientos

En primer lugar, quiero agradecer a ANID y a la Universidad de los Andes, por las becas que me entregaron para financiar mis estudios de doctorado.
Además, esta tesis contó con financiamiento parcial del Instituto de Sistemas Complejos de Ingeniería (ISCI) y del proyecto FONDECYT N.° 1250045 de ANID.
También quiero agradecer a los investigadores que colaboraron en esta investigación, especialmente a mi tutor, Carla Vairetti y a Sebastián Maldonado.
Finalmente, quiero agradecer a las empresas Solutions Now S.A. y CrossNet S.A., por la confianza depositada en esta investigación, y por permitirnos utilizar sus datos para generar estudios de casos reales.

Índice general

Índice de figuras

Índice de cuadros

Capítulo 1

Introducción

En la industria de servicios existe una alta competencia y los mercados han alcanzado una alta madurez, es por esto que las empresas se ven obligadas a fortalecer la relación con sus clientes. Utilizar los datos para mejorar los contactos con los clientes se ha vuelto cada vez más importante para las empresas que buscan mejorar las experiencias de los clientes y generar lealtad a largo plazo. Al aplicar técnicas basadas en datos, como los modelos predictivos, las empresas pueden obtener información valiosa sobre el comportamiento, las preferencias y las necesidades de los clientes. Estos modelos ayudan a identificar patrones y tendencias en las interacciones con los clientes, encontrar áreas de mejora y optimizar las estrategias de participación del cliente.

Una técnica importante de aprendizaje automático que se ha utilizado son los modelos de aprendizaje causal (modelos uplift), estos son herramientas estadísticas y analíticas que buscan comprender las relaciones de causa y efecto entre variables o eventos. Implica examinar datos para obtener conclusiones del impacto de una intervención o tratamiento específico en un resultado de interés. El avance de la inteligencia artificial y el aprendizaje profundo ha facilitado en gran medida la predicción de los efectos del tratamiento, ampliando el campo de la inferencia causal más allá del ámbito tradicional de los estudios econométricos.

En esta tesis se busca reunir una serie de avances metodológicos y aplicaciones novedosas del modelado causal y el aprendizaje automático. Como investigación aplicada, se presenta un nuevo enfoque que no ha sido discutido en la literatura de aprendizaje causal: la cobranza. Se diseña una estrategia para definir el subconjunto de deudores

cuya probabilidad de pago se ve influenciada por la intervención de cobro.

La gestión de la deuda morosa representa un desafío importante para las instituciones financieras a nivel mundial debido al aumento de clientes morosos en los últimos años. Esto genera la necesidad para las empresas de recuperar la deuda de manera eficiente, convirtiéndose en una tarea relevante.

En base a lo anterior, se identifica una importante laguna en la literatura. Si bien el problema de la calificación crediticia (riesgo previo al préstamo) ha sido ampliamente estudiado, la gestión de cobros permanece relativamente desatendida, recurriendo con frecuencia a estrategias rígidas y heurísticas. En este contexto, los modelos de aprendizaje automático podrían generar un aporte significativo en esta área.

Un elemento clave en esta tesis son los modelos basados en beneficio (modelos profit). El objetivo es estimar la ganancia real de la clasificación causal y tradicional, incorporando los beneficios y costos que resultan de su implementación. Un ejemplo común y ampliamente aplicado en la literatura es la predicción de abandono impulsada por las ganancias, que estima la ganancia esperada de una campaña de retención como una métrica del rendimiento del modelo. Se busca extender este modelo impulsado por las ganancias a la cobranza.

Al igual que en cobranza, la aplicación de modelos de aprendizaje causal basados en beneficio no se ha aplicado para el contexto de venta cruzada. En esta tesis se expande el enfoque aplicado a cobranzas para este contexto, donde se busca maximizar los ingresos por cliente pero también optimizar la rentabilidad de las campañas al considerar los esquemas de incentivos de la fuerza de ventas, transformando la recomendación de productos en una decisión estratégica orientada al valor económico.

Por último, se propone un nuevo enfoque para la estimación del modelo causal llamado efecto de tratamiento individual acotado (BITE, de acuerdo con sus siglas en inglés), que se personaliza para tareas de análisis empresarial. El objetivo es priorizar a los clientes que no están completamente convencidos de las acciones filtrando a los individuos en función de sus probabilidades de control.

La contribución científica que se busca con esta tesis incorpora diferentes ámbitos, incluye avances metodológicos como la estimación de una nueva métrica para los modelos causales y aplicaciones novedosas en el contexto del aprendizaje tradicional y causal basado en beneficios en modelos de cobranza y venta cruzada.

1.1. Hipótesis

Cuatro hipótesis fueron consideradas en el desarrollo de esta investigación:

I. Se puede obtener información valiosa sobre los deudores a partir de los *contact center*, que pueden complementar la información financiera de los clientes y mejorar las capacidades de clasificación de los modelos de cobranza.

II. La aplicación de modelos de cobranza puede beneficiarse de los modelos de aprendizaje causal. Las tareas de análisis predictivo que no se han analizado desde una perspectiva causal antes en estudios científicos se pueden abordar con éxito a través de modelos de aprendizaje causal.

III. Los modelos causales que se adaptan especialmente a las aplicaciones de análisis empresarial, como la predicción del cobro de deudas, conducen a un mejor rendimiento predictivo que las técnicas tradicionales.

IV. La aplicación de modelos causales basados en métricas de rentabilidad a estrategias de venta cruzada, permiten mejorar la efectividad de las recomendaciones mediante una asignación más adecuada de incentivos.

V. Las métricas para modelos de aprendizaje causal que no consideran los clientes completamente convencidos de un determinado resultado entregan resultados más certeros que las que sí los incluyen debido a que estos pueden ser excluidos de recibir el tratamiento independientemente del efecto de tratamiento individual (ITE).

1.2. Objetivos

El objetivo general de la investigación es reunir una serie de avances metodológicos y aplicaciones novedosas en los modelos causales y el aprendizaje automático. Estos avances permitirán el desarrollo de modelos predictivos más precisos y rentables en el contexto del análisis empresarial.
Los objetivos específicos de esta investigación son:

a) Estudiar la influencia de las variables de *contact center* en modelos predictivos de cobro de deuda, las cuales están combinadas con información financiera de los clientes deudores

b) Estudiar la diferencia entre la aplicación de modelos causales y modelos predictivos en los modelos de cobro de deudas.

c) Estudiar y comparar la rentabilidad de las técnicas tradicionales y de aprendizaje causal en modelos de cobro de deudas que no han sido analizadas desde una perspectiva causal antes en estudios científicos, proporcionando un marco adecuado orientado a la rentabilidad.

d) Evaluar la aplicación de modelos causales basados en métricas de rentabilidad para la generación de recomendaciones de venta cruzada en el sector bancario, analizando su capacidad para identificar esquemas de incentivos que optimicen la asignación de recursos comerciales.

e) Estudiar y comparar el rendimiento del efecto del tratamiento individual (ITE) y efecto de tratamiento individual acotado (BITE) de modelos causales aplicados en análisis empresariales.

1.3. Estructura del informe

El presente informe se estructura como sigue:

a) Capítulo 1: Introducción.

b) Capítulo 2: Revisión bibliográfica general y específica para esta investigación.

c) Capítulo 3: Metodología propuesta para esta investigación. En este capítulo se presenta la metodología para la aplicación de modelos de aprendizaje tradicional y causal en cobranza, para su posterior análisis en un marco basado en beneficio. En esa misma línea, se extiende el marco de modelos causales basado en beneficio para el ámbito empresarial de la venta cruzada. Finalmente, se presenta una versión acotada del efecto de tratamiento individual para la evaluación de los modelos causales.

d) Capítulo 4: Presentación y análisis de los principales resultados de esta investigación, para cada uno de los métodos presentados en la sección anterior.

e) Capítulo 5: Principales conclusiones de esta investigación, y propuestas de trabajo futuro para su continuidad.

Capítulo 2

Revisión Bibliográfica

En este capítulo se presenta la revisión bibliográfica que sustenta el desarrollo de esta tesis. En primer lugar, se introduce el proceso de descubrimiento de conocimiento en bases de datos (KDD), el cual proporciona el marco general para el desarrollo de proyectos de analytics y dentro del cual se desarrollan todos los experimentos considerados en este trabajo.

Sobre esta base, se revisa la evolución de los modelos aplicados al contexto de cobranza, desde enfoques estadísticos tradicionales hasta la incorporación de técnicas de aprendizaje automático, destacando los principales algoritmos utilizados en la literatura. A continuación, se amplía este marco hacia modelos de aprendizaje causal, enfatizando su creciente relevancia en aplicaciones de negocio y, en particular, en problemas de cobranza.

Posteriormente, se abordan los enfoques basados en costos y beneficios, integrando la literatura sobre modelos orientados a maximizar el valor económico. Finalmente, se incorpora una revisión de la literatura en gestión de relaciones con clientes y venta cruzada, con el objetivo de también medir el rendimiento de la aplicación de modelos causales basados en beneficio en el contexto de venta cruzada. De esta manera, la revisión bibliográfica configura un marco unificado que conecta modelos predictivos, causales y basados en beneficios en distintos contextos de negocio, estableciendo el fundamento teórico para los experimentos desarrollados a lo largo de la tesis.

2.1. Proceso de descubrimiento de conocimiento en bases de datos

La generación de predicciones mediante técnicas de minería de datos requiere de un conjunto limpio, es decir, sin valores faltantes y datos inconsistentes. La calidad de la información es importante, ya que la presencia de valores nulos, datos erróneos o inconsistencias puede afectar significativamente el desempeño de los modelos. Por esta razón, antes de aplicar cualquier técnica analítica se debe realizar un proceso de preparación de los datos.

Este conjunto de actividades se enmarca dentro del proceso conocido como descubrimiento de conocimiento en bases de datos, más conocido como KDD por sus siglas en inglés (Fayyad, 1996).

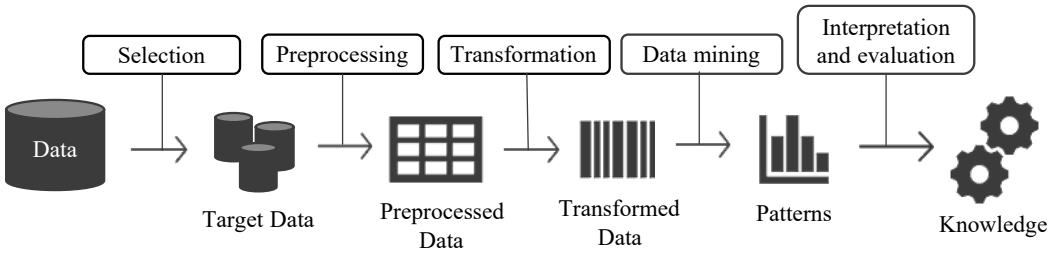


Figura 2.1: Proceso KDD

Este, es un procedimiento iterativo que permite extraer conocimiento útil a partir de grandes volúmenes de información. En las diferentes etapas, el usuario tiene un papel clave, ya que participa activamente en la toma de decisiones relacionadas con la selección de técnicas, la interpretación de resultados y la validación del conocimiento obtenido.

El proceso KDD se compone de cinco etapas principales. La primera es la selección, donde se determina el conjunto de datos que será analizado. En esta fase se busca que los datos sean representativos.

La segunda etapa es el preprocesamiento, que busca mejorar la calidad de los datos. Aquí se identifican y corrigen problemas como valores atípicos, datos faltantes, inconsistencias y registros duplicados. Para ello, se emplean diversas técnicas estadísticas y métodos de limpieza que permiten obtener un conjunto de datos más confiable y coherente. La adecuada ejecución de esta fase es fundamental ya que es la base de las fases posteriores.

Posteriormente, la fase de transformación donde el objetivo principal es enriquecer la

base de datos. Entre las técnicas empleadas se encuentran la agregación, la discretización, la segmentación y el análisis de correlación (Han, Kamber & Pei, 2012). Estas transformaciones facilitan la construcción de modelos más eficientes y evitan redundancias o información irrelevante.

La cuarta etapa corresponde a minería de datos, se utilizan técnicas de aprendizaje automático para la construcción de modelos y patrones. Se pueden aplicar métodos de aprendizaje supervisado, que utilizan datos previamente etiquetados, o aprendizaje no supervisado, que busca estructuras internas sin contar con información previa.

Finalmente, la etapa de interpretación y evaluación, que corresponde a la fase más importante del proceso, la cual consiste en analizar el desempeño de los modelos, validar los resultados obtenidos y extraer conclusiones. Sin esta fase no podemos obtener conclusiones y por ende tomar buenas decisiones en base a los resultados obtenidos.

2.2. Modelos de cobranza

A pesar de la gran cantidad de modelos predictivos diseñados para la gestión de riesgos y, en particular, para la calificación crediticia (Maldonado, Bravo, López & Pérez, 2017), la cobranza ha recibido poca atención en la comunidad académica. Los modelos de cobranza no solo permiten hacer proyecciones precisas de los flujos de caja y cuentas por cobrar futuros, sino que también mejoran el éxito de las operaciones de cobranza (Crespo & Govindarajan, 2018). Por ejemplo, se puede redirigir la fuerza de trabajo hacia aquellos clientes que es más probable que no paguen sus obligaciones (Kim & Kang, 2016), o se pueden diseñar nuevas políticas de cobranza distinguiendo entre los deudores que tienen una probabilidad de que paguen o los que no (Zurada & Lonial, 2005).

Este problema fue considerado por primera vez por Mitchner & Peterson (1957), quienes optimizaron el momento en que un agente debe contactar a un moroso para cobrar su deuda en base al costo que genera.

La investigación sobre análisis de cobranzas se enfoca principalmente en predecir si los clientes pagarán su deuda, para optimizar el proceso de cobro, generar nuevas políticas y hacer proyecciones en los flujos de efectivo, entre otros (Bahrami, Bozkaya & Balcisoy, 2020). Lo más utilizado para este tipo de modelos han sido las técnicas estadísticas, sin

embargo, el auge del aprendizaje automático y la inteligencia artificial ha proporcionado un nuevo conjunto de herramientas que ha demostrado su eficacia en los modelos de cobranza (Bahrami, Bozkaya & Balcisoy, 2020; van der Geer, Wang & Bhulai, 2018).

Zurada & Lonial (2005) pronosticaron el pago de la deuda en los sistemas de salud, donde existe un alto número de clientes morosos; ellos concluyeron que las redes neuronales, la regresión logística y los métodos de conjunto eran los mejores modelos para la predicción del pago de la deuda. Abe, Melville, Pendus, Reddy, Jensen, Thomas, Bennett, Anderson, Cooley, Kowalczyk, Domick & Gardinier (2010) presentó un proceso de optimización del cobro de deudas que apunta a encontrar las mejores acciones que conduzcan a la máxima recuperación. Para ello, propusieron un enfoque basado en el proceso de decisión de Markov restringido (MDP) a través del aprendizaje por refuerzo restringido.

En la misma dirección, Chehrazi, Glynn & Weber (2019) propusieron un enfoque de programación estocástica para la optimización dinámica de cobros de crédito. Si bien este enfoque no considera el modelado basado en datos, muestra en ejemplos simulados que su modelo de control estocástico óptimo es capaz de reducir las pérdidas, concentrando las políticas de cobranza en el mejor momento posible.

En los últimos años, se han incorporado técnicas de inteligencia artificial y aprendizaje automático. Tran & Tran (2025) utilizan algoritmos de aprendizaje automático basados en datos, combinados con investigación operativa para maximizar la recuperación de deudas en bancos e instituciones de crédito con un enfoque principal en las comisiones de recuperación. Estos modelos fueron probados en una *credit union*, demostrando mejores resultados que otras técnicas convencionales.

Por su parte, Przybyłek, Przybyłek, Wojciechowski & Shkroba (2025) presenta el desarrollo de un sistema inteligente de cobro de deudas (SDCS), que automatiza la toma de decisiones y se adapta dinámicamente al comportamiento del deudor. Este sistema combina procesos de Markov, redes neuronales profundas y un lenguaje específico del dominio.

Otros enfoques se han centrado en clasificar a los clientes según su nivel de riesgo. Sivamayilvelan, Rajasekar, Vairavasundaram, Balachandran & Suresh (2024) desarrollaron un algoritmo híbrido actor-crítico con optimización de política mediante región de confianza adaptativa para clasificar clientes y recomendar acciones que reduzcan la

morosidad. Como resultado, mejoraron la tasa de recompensa con y sin aprendizaje de transferencia a 73.22 % y 62.44 % y aumentaron la cobranza en un 20.34 %, reduciendo en un 16.30 % las visitas realizadas por los agentes de cobranza.
Finalmente, Sivamayilvelan, Rajasekar, Vairavasundaram, Balachandran & Suresh (2025) desarrollaron un sistema de recomendaciones basado en aprendizaje por refuerzo, generando acciones personalizadas según el riesgo del cliente y el desempeño del cobrador. Para hacer el modelo explicable, usaron inteligencia artificial de Google Vertex. El enfoque mejoró la tasa de cobro en clientes de riesgo muy bajo, bajo y medio.
Se identifican brechas importantes en la literatura, las cuales se busca abordar en esta tesis. En primer lugar, el estado del arte no considera estimar la probabilidad de un contacto exitoso con el cliente, que es clave para definir las mejores políticas de contacto y canales de comunicación, ni la probabilidad de que un cliente se comprometa a pagar su deuda pendiente en el debido tiempo. Este compromiso es importante en el proceso de cobranza ya que aumenta considerablemente la probabilidad de que un cliente pague su deuda pendiente, que es el objetivo principal de la tarea. Además, tener en consideración que los estudios existentes se centran solo en esta tarea de pago.

2.3. Modelos de aprendizaje automático

Para predecir el comportamiento del cliente, existen diversos modelos de minería de datos y estadística que permiten realizar predicciones basadas en sus datos. En esta línea, la inteligencia artificial ha tomado relevancia en los últimos años, donde el aprendizaje automático constituye una de las ramas principales porque incluye diversas técnicas que permiten a los sistemas computacionales "aprender" a partir de los datos (Wolff, Erichsen, Kevrekidis & Rotermund, 2000). Existen distintos tipos de aprendizaje para el entrenamiento de los modelos, entre los que destacan el aprendizaje supervisado, el aprendizaje no supervisado y el aprendizaje por refuerzo. En esta tesis se estudian diversos problemas enmarcados en el aprendizaje supervisado, donde los modelos se entrenan utilizando datos históricos previamente etiquetados, de modo que el conocimiento previo sirva como base para realizar predicciones sobre nuevas observaciones.
Como se mencionó anteriormente, en el área de cobranza y centros de contacto, se han probado diversas aplicaciones en diferentes escenarios, identificando los modelos más

efectivos para cada caso. En este trabajo, se mide el rendimiento de diferentes modelos de clasificación para determinar cuál es el que mejor se adecúa a los datos y el caso de uso que se está aplicando.

- Regresión logística: Consiste en crear una ecuación que, para cada variable, asigna un coeficiente, $\boldsymbol{\beta}$, que indica su influencia en la predicción de la variable objetivo. La ecuación será la suma de todas las variables multiplicada por su coeficiente, lo que da como resultado la variable objetivo.

$$\mathbf{y} = \beta_0 + \beta_1 \mathbf{x}_1 + ... + \beta_n \mathbf{x}_n. \tag{2.1}$$

- Árbol de decisión: Utiliza un algoritmo de partición basado en el índice Gini o la entropía para seleccionar la mejor variable en cada ramificación del árbol, generando una estructura jerárquica hasta alcanzar la pureza del nodo (Baesens, 2014). Esto permite analizar la relevancia de las variables según los criterios de división. Random forest extiende este enfoque mediante la creación de múltiples árboles de decisión y combina sus resultados para obtener la predicción final (Hastie, Tibshirani & Friedman, 2009).

- Extremely Randomized Trees (Extra-T): Usa una técnica bastante similar a Random Forest, sin embargo, introduce un mayor grado de aleatoriedad en el proceso de construcción de los árboles al hacer las ramificaciones de manera aleatoria. Esta estrategia incrementa la diversidad entre los árboles del modelo, lo que puede ayudar a reducir la varianza y mejorar la capacidad de generalización.

- K-vecinos más cercanos: Es un algoritmo de clasificación que, basándose en los atributos de cada objeto, entrena el modelo comparándolo con otros objetos con atributos similares, ubicándolos en un plano de n dimensiones, donde n representa el número de atributos (Han, Kamber & Pei, 2012). Posteriormente, los nuevos objetos se colocan en el plano según sus atributos y se clasifican en relación con los k vecinos más cercanos.

- Red neuronal: Consiste en una red de nodos (neuronas), donde cada nodo es una unidad de procesamiento. La estructura de estas redes se define por capas: la

capa inicial contiene las variables de la base de datos, la capa final representa la salida del modelo y las capas intermedias se encargan del procesamiento de los datos. Estas capas se conectan mediante neuronas, cuyas conexiones tienen un peso asignado que determina su relevancia.

- Gradient Boosting (GBoost): Es otro método de ensamble basado en aprendices débiles, como los árboles de decisión. Este enfoque utiliza el principio de *boosting*, en lugar de *bagging*, con el objetivo de mejorar progresivamente el aprendizaje del modelo. Cada nuevo modelo se entrena para corregir los errores cometidos por los anteriores, prestando especial atención a las muestras que resultan más difíciles de clasificar.

- LightGBM (LGBM): Es una versión optimizada y más rápida de Gradient Boosting. LightGBM hace crecer los árboles siguiendo un enfoque *leaf-wise*, es decir, expandiendo primero las hojas que generan mayor ganancia. Este procedimiento suele permitir modelos más eficientes y precisos.

- eXtreme Gradient Boosting (XGBoost): Es una variante avanzada de Gradient Boosting que optimiza el proceso de minimización de la función de pérdida utilizando el método de Newton-Raphson, en lugar del descenso por gradiente tradicional.

2.4. Aprendizaje automático causal en analítica de negocios

A lo largo de los años, se han llevado a cabo investigaciones para entender los efectos causales de diversas acciones, como el impacto de los programas de formación en los ingresos de los trabajadores, el papel de la escolarización católica en el aprendizaje, y la influencia de los antecedentes familiares y la capacidad mental en el desempeño educativo (Lalonde, 1986; Coleman & Hoffa, 1987; Reyna, 1968). También, se ha utilizado ampliamente el análisis estadístico para entender las relaciones causales entre diferentes eventos o características de las empresas y su impacto en las ventas y marketing (Friberg & Sanctuary, 2017; Seiler, Yao & Wang, 2017).

Descubrir relaciones causales es difícil ya que no se puede aplicar más de un tratamiento a un individuo al mismo tiempo. Sin embargo, los avances recientes en aprendizaje automático se han adaptado para inferir relaciones causales (Pearl, 2009; Peters, Janzing & Scholkopf, 2017; Malinsky & Danks, 2018; Glymour, Zhang & Spirtes, 2019) o predecir los efectos del tratamiento a partir de datos (Athey & Imbens, 2015; Künzel, Sekhon, Bickel & Yu, 2019a; Athey S, 2019).

El aprendizaje automático causal integra los campos de la inferencia causal y el aprendizaje automático para evaluar el impacto de las decisiones sobre variables de interés para las organizaciones. Su objetivo es traducir los datos a percepciones razonables teniendo en cuenta que ciertas acciones afectan a cada individuo de manera diferente. A diferencia de los modelos de clasificación tradicionales que predicen el valor futuro de un resultado, los modelos causales estiman la magnitud del cambio en el valor de un resultado debido a la aplicación de un tratamiento. Esto permite a los tomadores de decisiones desarrollar campañas personalizadas donde se aplica la acción más favorable a cada individuo.

Otra diferencia entre los métodos convencionales de aprendizaje automático y el aprendizaje causal es que este último no se centra en predecir el valor futuro de los resultados, sino en la predicción del efecto de los tratamientos a nivel individual (ITE). La acción generalmente se aplica a personas para quienes el ITE es mayor que un umbral determinado.

El aprendizaje causal se ha abordado en la literatura a través de varios enfoques, incluidos los modelos causales (Devriendt, Moldovan & Verbeke, 2018a), la estimación del efecto del tratamiento heterogéneo (Wager & Athey, 2017), el aprendizaje de reglas de tratamiento individualizado (Zhou, Mayer-Hamblett, Khan & Kosorok, 2015) y la estimación del efecto del tratamiento promedio condicional (Fan, Hsu, Lieli & Zhang, 2021). Estos enfoques difieren en su énfasis en predecir el tratamiento óptimo a nivel individual o en sacar conclusiones de inferencia causal.

Un ejemplo de cómo el aprendizaje automático causal puede mejorar significativamente la rentabilidad de las campañas de orientación se presenta en la metodología proporcionada por Gubela, Lessmann & Jaroszewicz (2020). Los experimentos con datos de comercio electrónico de una tienda europea indican que los métodos basados en conjuntos diseñados para el aprendizaje causal superan en rendimiento a los métodos de

regresión.

Además, varios estudios en diferentes dominios han llegado a conclusiones similares en la predicción de abandono (Devriendt, Berrevoets & Verbeke, 2021a; Olaya, Vásquez, Maldonado, Miranda & Verbeke, 2020; Vafeiadis, Diamantaras, Sarigiannidis & Chatzisavvas, 2015). Teniendo en cuenta que centrarse en clientes de alto riesgo puede ser ineficaz para mejorar la retención (Ascarza, 2017), estos estudios tienen como objetivo identificar a aquellos clientes que tienen más probabilidades de permanecer en la empresa solo si se les hace una oferta.

La literatura sobre modelos causales distingue entre métodos de preprocesamiento que ajustan los datos antes de la aplicación de los modelos y de procesamiento de datos que estiman los ITE de forma directa o indirecta, es decir, usando modelos separados para tratados y no tratados, o adaptando algoritmos de aprendizaje automático (Devriendt, Moldovan & Verbeke, 2018a). Los métodos para estimar el ITE se organizan en dos grandes categorías: los enfoques de *meta-learners* y los modelos de aprendizaje automático diseñados específicamente para inferencia causal. Los *meta-learners* combinan algoritmos tradicionales de clasificación o regresión para estimar los ITE sin modificar el modelo base (Künzel, Sekhon, Bickel & Yu, 2019b), mientras que los algoritmos causales incorporan directamente la estructura del problema causal en su diseño. En este trabajo se adopta el enfoque de *meta-learners*, aprovechando su flexibilidad y capacidad para integrarse con modelos predictivos estándar. Estas metodologías se describen a continuación:

- Enfoque de dos modelos (TMA) o T-learner: Este método implica la construcción de dos clasificadores distintos, uno para el grupo de tratamiento y otro para el grupo de control.

- Enfoque de Covariables Modificado (MCA) o S-learner: Se entrena un único clasificador utilizando una variable indicadora de tratamiento/control. La matriz de covariables se amplía para incluir las interacciones entre los atributos originales y la variable causal.

- Enfoque de Resultados Modificado (MOA): Se define una única tarea predictiva, pero el vector objetivo se modifica en función de cuatro posibles resultados, entre

otras alternativas: el tratado positivo (TP), el control positivo (CP), el tratado negativo (TN) y el control negativo (CN).

- Causal forest (CF): Se construye un conjunto de árboles de decisión u otros clasificadores, donde cada modelo se entrena con un subconjunto diferente de los datos con el propósito específico de estimar efectos de tratamiento heterogéneos. El promedio de las estimaciones del efecto del tratamiento de todos los modelos proporciona el ITE final (Athey S, 2019).

La validez en la estimación del efecto individual del tratamiento (ITE) depende del mecanismo de asignación, es decir, de cómo se determinan los tratamientos. En la literatura se distinguen tres tipos de diseños: experimentos aleatorios, experimentos naturales y estudios observacionales. En los dos primeros, el investigador conoce o controla la asignación, mientras que en los estudios observacionales este proceso no es observable ni controlable (Morgan & Winship, 2014)

La estimación de efectos de tratamiento individual (ITE) en estudios observacionales se ve afectada por la asignación no aleatoria (NRA) y sesgos de confusión. Para corregirlos se utilizan métodos como *propensity score matching* (PSM), ponderación y controles sintéticos, aunque en escenarios desbalanceados el PSM puede causar pérdida de información, motivando extensiones como *caliper matching*, funciones kernel y ponderación (Nyberg & Klami, 2023; Vairetti, Gennaro & Maldonado, 2024).

Además del sesgo de asignación, en los modelos causales puede existir desbalance de clases y de tratamiento, problemas que se agravan en estudios observacionales donde métodos como PSM pueden eliminar ejemplos críticos por no tener en cuenta las clases (Nyberg & Klami, 2023). Avances recientes como ProSOM aplican técnicas de sobremuestreo para manejar el desbalance de tratamiento bajo NRA, reduciendo la pérdida de información (Vairetti, Gennaro & Maldonado, 2024).

Como se puede ver, existen varias aplicaciones de modelos de aprendizaje causal enfocados en el rendimiento de las campañas, pero en la literatura no se ha utilizado para mejorar los modelos de cobranza. En este trabajo, se busca determinar qué clientes se debe contactar debido a que esto llevará al pago de la deuda y de esta manera, enfocarse en los que se obtendrá un beneficio de esta llamada.

2.5. Análisis de costo en modelos de aprendizaje causal

La decisión de ofrecer o aplicar un tratamiento significa un beneficio o costo para la empresa, por lo que es importante evaluar la ganancia de aplicarlo. Los modelos de aprendizaje causal permiten determinar a quién aplicar un tratamiento para que se logre un objetivo, esto nos permite medir los costos de estas acciones de forma de maximizar el beneficio para la empresa, reduciendo costos de campaña y aumentando el retorno. Por ejemplo, el beneficio obtenido al implementar una campaña de retención puede ser una métrica de desempeño más adecuada que cualquier medida estadística tradicional (Jiang, Liu, Abedin, Wang, Yang & Dong, 2024; Maldonado, Domínguez, Olaya & Verbeke, 2021).

Para medir los beneficios y costos de los modelos de aprendizaje automático se han desarrollado los modelos basados en beneficio o modelos profit. En la historia han existido pocos estudios que analicen los costos y beneficios, Hansotia (2002) calcula el rendimiento incremental de la inversión al nivel del margen bruto. Radcliffe & Surry (2012) calcula la ganancia incremental multiplicando la tasa de respuesta incremental por la ganancia total.

Recientemente, Devriendt, Berrevoets & Verbeke (2021a) analizaron el problema de costos y beneficios para problemas causales de abandono de clientes donde proponen como medida el aumento de la ganancia máxima que calcula el beneficio que se obtiene al aplicar modelos de aprendizaje causal. También, se han propuesto enfoques novedosos donde se integra la clasificación causal y sensible a los costos mediante la elaboración de un marco de evaluación unificado (Verbeke, Olaya, Berrevoets, Verboven & Maldonado, 2020). A partir de esto, demuestran que la clasificación convencional es un caso específico de clasificación causal en términos de un rango de medidas de desempeño cuando el número de acciones es igual a uno.

En los últimos años se ha avanzado en el estudio de los beneficios de los modelos de aprendizaje causal, enfocado principalmente en la retención de clientes, aprendizaje causal de ingresos y nuevas formas de calcular este beneficio. A pesar de esto, no existe literatura que analice el beneficio para los modelos causales de cobranza, que permitan identificar a qué clientes contactar para que paguen su deuda con la entidad financiera.

Otro aspecto relevante que se aborda es el aprendizaje automático con fines de lucro. A diferencia de las métricas de evaluación estadística como la precisión, las métricas de ganancias calculan los beneficios y costos reales de implementar un modelo predictivo (Höppner, Stripling, Baesens, vanden Broucke & Verdonck, 2017; Zhang, Wang & Liu, 2023). Por ejemplo, un enfoque de predicción de abandono impulsado por las ganancias consiste en encontrar el clasificador que maximiza la ganancia promedio de una campaña de retención (Maldonado, López & Vairetti, 2020; Maldonado, Álvaro Flores, Verbraken, Baesens & Weber, 2015). Se han dedicado varios estudios al uso de métricas de ganancias en clasificación, pero solo hay unos pocos dedicados al aprendizaje causal impulsado por las ganancias (Petrides, Moldovan, Voets, Guns & Verbeke, 2022).

2.5.1. Métricas para modelos basados en beneficio

En Verbeke, Olaya, Guerry & Van Belle (2023) se presenta un marco unificado que combina la clasificación convencional, la clasificación orientada al beneficio y el enfoque causal. La clasificación orientada en beneficio tiene como objetivo determinar el umbral óptimo ϕ^* que maximiza el Beneficio Máximo (MP):

$$MP := \max_{\forall \phi} \big(P(\phi; \mathbf{CB})\big) = P(\phi^*; \mathbf{CB}), \tag{2.2}$$

donde P corresponde al profit, obtenido como la suma de los elementos de la matriz de confusión **CF** y la matriz de costos-beneficios **CB**:

$$P := \Sigma_{i,j}(\mathbf{CF}_{ij} \circ \mathbf{CB}_{ij}), \tag{2.3}$$

siendo $\Sigma_{i,j}(\mathbf{A}_{ij})$ la suma de todos los elementos de la matriz **A** y $\circ$ el producto de Hadamard. Para un umbral ϕ, la matriz de confusión **CF** entrega las proporciones de verdaderos negativos, falsos positivos, falsos negativos y verdaderos positivos. Se estructura de la siguiente forma:

$$\mathbf{CF} := \begin{array}{c} \\ Y=0 \\ \\ Y=1 \end{array} \begin{array}{c} \begin{array}{cc} \hat{Y}=0 & \hat{Y}=1 \end{array} \\ \begin{bmatrix} \pi_0 F_0(\phi) & \pi_0\big(1-F_0(\phi)\big) \\ \text{(Verdaderos negativos)} & \text{(Falso positivos)} \\ \pi_1 F_1(\phi) & \pi_1\big(1-F_1(\phi)\big) \\ \text{(Falso negativos)} & \text{(Verdadero positivos),} \end{bmatrix} \end{array} \tag{2.4}$$

donde $\hat{Y}$ representa la predicción del modelo, π_1 (π_0) es la probabilidad de la clase positiva (negativa) y $F_1(\phi)$ ($F_0(\phi)$) es la función de distribución acumulada de los valores predichos para la clase positiva (negativa). Cada uno de estos resultados puede tener impactos económicos distintos según la aplicación, la matriz **CB** indica los costos y beneficios asociados a cada combinación posible:

$$\mathbf{CB} := \begin{array}{c} \\ Y=0 \\ Y=1 \end{array} \begin{array}{c} \begin{array}{cc} \hat{Y}=0 & \hat{Y}=1 \end{array} \\ \begin{bmatrix} cb_{00} & cb_{01} \\ cb_{10} & cb_{11} \end{bmatrix} \end{array} \tag{2.5}$$

Las métricas basadas en beneficio podrían ser útiles para seleccionar clientes en campañas de cobranza, pero existe el desafío de la heterogeneidad de los montos adeudados que genera distribuciones de beneficio asimétricas. Maldonado, Domínguez, Olaya & Verbeke (2021) demuestran que en sectores con valores heterogéneos, como los fondos mutuos, un único umbral de decisión es subóptimo y demuestran que el enfoque de profit multisegmento (MSP) mejora la rentabilidad al ajustar los umbrales según cuantiles de la distribución.

Verbeke, Olaya, Guerry & Van Belle (2023) formalizan el beneficio causal por instancia ($\dot{P}$) como la diferencia entre los beneficios derivados de los resultados y los costos asociados a los tratamientos aplicados. Para ello, se define la matriz de confusión causal $\dot{\mathbf{CF}}$, que recoge las proporciones de resultados positivos y negativos en los grupos de control y tratamiento para un umbral $\dot{\phi}$, y la matriz causal de costos-beneficios $\dot{\mathbf{CB}}$, que establece el beneficio neto ($b_{ij} - c_{ij}$) según el tratamiento aplicado.

El beneficio del modelo causal se calcula como:

$$\dot{P} = \Sigma_{i,j}(\dot{\mathbf{CF}}_{ij} \circ \dot{\mathbf{CB}}_{ij}), \tag{2.6}$$

mientras que el Beneficio Causal Máximo ($\dot{MP}$) se define como:

$$\dot{MP} := \max_{\phi} \left(\dot{P}(\phi; \dot{\mathbf{CB}})\right) = P(\phi^*; \dot{\mathbf{CB}}). \tag{2.7}$$

Este enfoque fue desarrollado originalmente para la predicción de abandono y la optimización de campañas de retención (Verbeke, Olaya, Guerry & Van Belle, 2023). No obstante, los modelos causales sensible a costos no se ha utilizado en escenarios distintos del marketing. En esta tesis, se adapta al ámbito de cobranza para determinar castigos óptimos de deuda mediante una estrategia multi-segmento.

2.6. Gestión de relaciones con el cliente y venta cruzada

El análisis de datos en CRM desempeña un papel importante en la optimización de las interacciones con los clientes y la mejora del rendimiento empresarial general (Baesens & de Caigny, 2022; Baier & Stöcker, 2022). Dentro de esto, la integración de 3 conceptos permite que las empresas expandan su base de clientes y al mismo tiempo creen y mantengan relaciones valiosas con los clientes, estos conceptos son modelos de adquisición de clientes, predicción del abandono e iniciativas de venta cruzada (Baesens & de Caigny, 2022; Vu, 2024).

Los primeros dos conceptos, adquisición de clientes y predicción de abandono se han abordado varias veces en la literatura del aprendizaje automático para la toma de decisiones (Baesens & de Caigny, 2022). Esto les ha permitido a las empresas identificar y atraer a potenciales clientes. Además, los modelos de predicción de abandono que utilizan aprendizaje automático permiten identificar señales tempranas de desvinculación para que las empresas puedan implementar estrategias de retención a tiempo fomentando la fidelización del cliente y manteniendo relaciones a largo plazo (Verbeke, Olaya, Guerry & Van Belle, 2023)

La clasificación basada en beneficios toma relevancia en el análisis de CRM, donde los costos y beneficios se incorporan al entrenamiento de los modelos de clasificación (Devriendt, Berrevoets & Verbeke, 2021b). En este contexto, el beneficio de implementar una campaña de retención puede ser una mejor métrica de rendimiento que una medida estadística (Lemmens & Gupta, 2020; Maldonado, López & Vairetti, 2020; Petrides, Moldovan, Voets, Guns & Verbeke, 2022). Este enfoque se ha aplicado para la predicción de abandono de clientes y la adquisición de clientes (Baesens & de Caigny, 2022), pero no se ha utilizado en el contexto de venta cruzada. Esto implica analizar el comportamiento y las preferencias de los clientes para recomendar productos o servicios complementarios, maximizando así los ingresos de los clientes existentes (Baesens & de Caigny, 2022).

La venta cruzada puede abordarse mediante distintos enfoques predictivos. El modelo de "próximo producto a comprar" (NPTB) busca anticipar qué adquirirá un cliente según su historial y comportamiento (Kalkan & Şahin, 2023), mientras que los sistemas de recomendación ofrecen sugerencias personalizadas basadas en datos de transacciones (Ghoshal, Mookerjee & Sarkar, 2021). Además, también puede abordarse como un problema no supervisado mediante la segmentación de clientes, asumiendo que los productos comprados por usuarios similares son buenas recomendaciones (Zhang, Priestley, DeMaio, Ni & Tian, 2021).

En relación con el enfoque que se busca en esta tesis, estudios sobre la gestión de la fuerza de ventas han analizado cómo los esquemas de compensación influyen en el desempeño de los vendedores (Mantrala, Albers, Caldieraro, Jensen, Joseph, Krafft, Narasimhan, Gopalakrishna, Zoltners, Lal & others, 2010). Otros estudios muestran que mecanismos como cuotas, bonos o límites pueden afectar de manera distinta la motivación, incluso desincentivando a los vendedores de alto rendimiento (Misra & Nair, 2011; Chung, Steenburgh & Sudhir, 2014). Daljord, Misra & Nair (2016) exploraron cómo la heterogeneidad entre vendedores influye en la efectividad de incentivos homogéneos, sugiriendo que esquemas más personalizados podrían ser más rentables, aunque los enfoques uniformes pueden funcionar si se optimiza la composición del equipo con la compensación. Por último, se ha estudiado la asignación de incentivos por producto, evidenciando que estos afectan tanto el esfuerzo de los vendedores como las decisiones y preferencias de los consumidores. No obstante, estos enfoques buscan mejorar las estrategias de venta desde distintas perspectivas, pero ninguno utiliza el análisis pre-

dictivo basado en aprendizaje automático que es lo que se busca implementar en esta tesis. Ademas, se busca incorporar la rentabilidad a través de la aplicación de modelos basados en beneficios que en la literatura no se ha abordado para el contexto de venta cruzada.

Capítulo 3

Metodología

Este capítulo presenta la metodología utilizada en esta tesis, orientada a la optimización de decisiones en contextos empresariales, bajo un enfoque predictivo y basado en beneficios. En primer lugar, se desarrolla un modelo de clasificación en el ámbito de cobranza, cuyo objetivo es estimar la probabilidad de pago de los clientes, incorporando métricas de evaluación alineadas con el impacto económico de las decisiones. Luego, se introduce el enfoque de modelos de aprendizaje causal basados en beneficio, el cual permite estimar el efecto de diferentes intervenciones sobre el comportamiento de pago, manteniendo como criterio central la maximización del beneficio esperado.
Posteriormente, este marco metodológico se extiende al contexto de la venta cruzada, donde los modelos causales se utilizan para identificar la influencia de los incentivos dentro de las recomendaciones de venta entre unidades de negocio, bajo una perspectiva de rentabilidad. Finalmente, se abordan métricas de evaluación específicas para modelos causales que buscan enfocarse en los clientes que la acción tiene mayor impacto.

3.1. Modelos predictivos en cobranzas

Primero, se define un marco de análisis predictivo para los procesos de cobranza a través de la aplicación de modelos de aprendizaje automático con información del *contact center* y la entidad financiera de los clientes morosos. En primera instancia se integran estas dos fuentes para mejorar el rendimiento predictivo de modelos y apoyar la toma de decisiones en la cobranza. Además, se proponen nuevas tareas predictivas que complementan la toma de decisiones no solo determinando si pagar la deuda o no.

También se aborda la optimización de las ganancias por recuperación de deudas, donde se plantea un marco predictivo basado en beneficios. El enfoque incluye la creación de umbrales multisegmento con un modelo sensible a los costos para gestionar la heterogeneidad de los saldos de deuda.

3.1.1. Conjunto de datos

Para los modelos de cobranza se utilizan los datos de una entidad financiera chilena, donde se considera la integración de datos a partir de dos fuentes: información que tiene la entidad financiera de los clientes y datos del *contact center* con el que trabaja dicha entidad para el cobro de la deuda a los clientes morosos.

El conjunto de datos consta de un total de 1,023,418 llamadas con la información financiera del cliente, recopilada entre noviembre de 2019 y enero de 2020. Los datos se presentan en la tabla 3.1. Del total de llamadas, se observa que 75,595 son contactos exitosos, 26,466 de ellas resultan en una promesa de pago y, en 54,349 el deudor paga la deuda morosa, independientemente de si se realizó o no una promesa.

Cuadro 3.1: Resumen de los datos para tres tareas predictivas.

Tarea predictiva	Clase	Cantidad de registros
Contacto exitoso	True	75,595 (7.387 %)
	False	947,823 (92.613 %)
Promesa de pago	True	26,466 (2.586 %)
	False	996,952 (97.414 %)
Pago de la deuda	True	54,349 (5.311 %)
	False	969,069 (94.689 %)

Se define el conjunto con información financiera del banco como *datos financieros* (FD) y para aquellas construidas a partir de la información histórica del centro de contacto como *datos del contact center* (CCD). A partir de estas dos fuentes de datos, se dividen en 5 subconjuntos, 2 de la fuente de datos financiera: FD1.Debt que incluye las variables

que describen la deuda del cliente y FD2.Debtor con las variables que describen al deudor. Por el otro lado, 3 de la fuente de datos del *contact center*: CCD1.BTTC que incluye las variables de cuándo es el mejor momento para contactar al cliente, CCD2.Attempts con los intentos de llamada y CCD3.Contacts que incluyen la información del historial de contactos de los clientes. Las variables de cada conjunto se describen en el anexo A.

3.1.2. Modelos de aprendizaje tradicional

En la literatura se ha abordado el problema de cobranzas, pero existen pocas aplicaciones con modelos de aprendizaje automático que utilicen la información de los *contact center* para predecir el desempeño de los clientes.
En esta sección, se plantea un análisis predictivo para extraer conocimiento de los datos del *contact center* y la información financiera de clientes morosos. El objetivo es fusionar estas dos fuentes para crear clasificadores con mejores resultados que permitan apoyar la toma de decisiones en el cobro de deudas. También se plantean dos nuevas tareas predictivas que van más allá de predecir si un cliente pagará la deuda. A partir de esto, se abordan las 3 tareas predictivas planteadas a continuación:

- *Contacto Exitoso*: Busca predecir si un deudor responderá a la llamada realizada a través del *contact center* para el cobro de su deuda. Un contacto exitoso se define cuando el deudor responde a la llamada y no un tercero.

- *Promesa de pago atrasado*: Predice si un deudor se comprometerá a pagar su deuda al ser contactado para el cobro de esta.

- *Pago de la deuda*: Busca predecir si un deudor pagará su deuda. Se considera como caso exitoso cuando el pago se realiza entre cero y cinco días después de la llamada. En este caso, no es necesario que el cliente responda la llamada o que se comprometa a pagar.

Para aplicar estas 3 tareas predictivas se utilizan modelos de clasificación. En relación a la revisión bibliográfica, ningún método supera a otros en esta aplicación, por lo tanto, se mide el rendimiento de varios clasificadores. Se aplican cinco modelos de clasificación tradicional, regresión logística, árboles de decisión, random forest, k-vecinos más cercanos y redes neuronales descritos en el capítulo 2. Para la evaluación de estos

modelos, se aplica la métrica de área bajo la curva ROC (AUROC). Esta medida entrega un equilibrio entre sensibilidad y especificidad, lo que la hace adecuada en condiciones de desequilibrio de clases (Hastie, Tibshirani & Friedman, 2009).

El desbalance de clases constituye una dificultad clave en el análisis predictivo ya que afecta el desempeño de los modelos, que tienden a favorecer la clase mayoritaria y presentan un bajo rendimiento al identificar correctamente los casos de la clase minoritaria (Simsek, Dag, Tiahrt & Oztekin, 2020). Con el objetivo de mitigar este problema, se aplican dos técnicas de remuestreo: el submuestreo aleatorio (RUS) y el método de sobremuestreo sintético para la clase minoritaria (SMOTE). Se evalúan diferentes combinaciones para incrementar el tamaño de la clase minoritaria y posteriormente, se aplica RUS sobre la clase mayoritaria para reducir su tamaño hasta igualarlo con el de la clase minoritaria. Esta metodología ha demostrado ser efectiva en diversas aplicaciones de clasificación (por ejemplo, (Simsek, Dag, Tiahrt & Oztekin, 2020)).

Finalmente, se analiza la influencia de las distintas variables en la solución final con el fin de obtener información adicional sobre el comportamiento del modelo. Para ello, se emplea el método TreeSHAP (Lundberg, Erion, Chen, DeGrave, Prutkin, Nair, Katz, Himmelfarb, Bansal & Lee, 2020), integrado en el clasificador mediante un procedimiento de eliminación regresiva, lo que permite tanto la selección como la interpretación de las características, además de la construcción de representaciones visuales que facilitan el análisis. Este método calcula los valores de Shapley, una medida ampliamente reconocida en la teoría de juegos cooperativos, como una estimación de la contribución de cada atributo (Molnar, 2020). Así, el rendimiento del modelo se descompone en aportes individuales que explican la predicción, y los valores de Shapley proporcionan un criterio fundamentado para distribuir dichas contribuciones entre las variables consideradas.

3.1.3. Modelos de aprendizaje tradicional basados en beneficio

Luego, se aplican modelos basados en beneficio en los resultados obtenidos de los modelos de aprendizaje tradicional, con el fin de evaluar los costos y beneficios de la empresa para los modelos e identificar técnicas para contactar a los clientes que permitan recuperar la mayor cantidad de deuda. Además, se aplica un enfoque multisegmento para gestionar la heterogeneidad inherente de los saldos de deuda.

Para los parámetros de entrada de la matriz **CB** se tiene el beneficio (b) que viene dado

por la cantidad de deuda que recuperará la empresa cuando un cliente paga y se obtiene cuando se identifica de manera correcta a los deudores que pagarán la deuda. Este dato sería el equivalente al valor de CLV en la predicción de abandono.

El beneficio no es exactamente la deuda ya que no se sabe cuánto pagará el cliente de la deuda pendiente con la entidad financiera. A partir de esto, el beneficio se calcula como la deuda de cada cliente, multiplicada por el promedio de la deuda pagada por todos los clientes de la base de datos.

Los costos de la campaña de cobranza son en primer lugar el costo de condonación de la deuda que viene dado por el descuento en el pago de la deuda, definido como d, que es asumido por el cliente en su versión más general, y el costo de contactar al cliente, que es igual para todos $\mathbf{f} = f \cdot \mathbf{1}$. De esta forma, la matriz **CB** queda de la siguiente forma:

$$\mathbf{CB} := \begin{array}{c} \begin{array}{cc} \hat{Y} = 0 & \hat{Y} = 1 \end{array} \\ \left[\begin{array}{cc} 0 & -\mathbf{f} \\ 0 & \mathbf{b} - (\mathbf{d} + \mathbf{f}) \end{array}\right] \end{array} \begin{array}{l} Y = 0 \\ Y = 1 \end{array} \tag{3.1}$$

De acuerdo con estudios previos sobre marcos sensibles al costo para la clasificación tradicional (por ejemplo, (Maldonado, Domínguez, Olaya & Verbeke, 2021)), nos centramos en el beneficio incremental de la acción de cobro. Si se incluye **b** en la ecuación (3.1) para los falsos negativos (FN), el modelo de clasificación se ve "recompensado" por no identificar a los deudores que habrían pagado independientemente de la intervención. El marco estándar basado en el beneficio presenta varias limitaciones en su aplicación a la cobranza: por un lado, el uso de un descuento uniforme ignora la heterogeneidad de las deudas, lo que puede generar pérdidas en segmentos con montos bajos debido a que la deuda del cliente seria menor al incentivo. Para resolverlo, este enfoque propone un esquema segmentado con descuentos flexibles que se ajustan a los distintos niveles de deuda.

Por otro lado, asumir una probabilidad de aceptación constante (γ) no se ajusta al comportamiento observado en la realidad, ya que los deudores presentan distintos grados de sensibilidad y capacidad de respuesta ante los incentivos, en este caso, el descuento en su deuda. El enfoque basado en la segmentación permite atacar parcialmente esta

limitación, debido a que las deudas de mayor monto suelen requerir descuentos mayores para ser aceptadas, de esta forma, se aplicaría un γ homogéneo entre los deudores de cada segmento. Sin embargo, una solución más robusta seria los modelos causales orientados a las ganancias, el cual permite capturar la heterogeneidad en la respuesta de los deudores frente a las distintas estrategias de cobranza y orientar las decisiones hacia la maximización de la rentabilidad esperada.

Finalmente, dado que los costos de contacto en cobranza suelen ser más altos que en otros sectores porque se requiere mayor cantidad de intentos de contacto para lograr un contacto exitoso, el enfoque incorpora valores más elevados para f que reflejan estos mayores costos.

Clasificación tradicional multisegmento orientada a las ganancias

Siguiendo el razonamiento de la métrica MSP (Maldonado, Domínguez, Olaya & Verbeke, 2021), la población se divide en q segmentos (deciles) según el saldo:

Definición 3.1.1 (Parámetros de rentabilidad por segmento)**.** *Sea $\bar{b}_i$ el saldo promedio de la deuda para cada segmento $i \in \{1, \ldots, q\}$. Definimos el incentivo normalizado (fracción de cancelación) como $\delta_i = d_i/\bar{b}_i$ y el costo de contacto normalizado como $\gamma_i = f/\bar{b}_i$.*

La rentabilidad del segmento i se calcula en función del rendimiento de un clasificador en un umbral dado ϕ_i. La ganancia por segmento es:

$$P_i(\phi_i) = \bar{b}_i \left[\pi_{1,i}(1 - \delta_i - \gamma_i)F_{1,i}(\phi_i) - \pi_{0,i}(\delta_i + \gamma_i)F_{0,i}(\phi_i)\right], \tag{3.2}$$

donde $\pi_{1,i}$ y $\pi_{0,i}$ son las probabilidades de pago y no pago en el segmento i. Los términos $F_{1,i}(\phi_i)$ y $F_{0,i}(\phi_i)$ representan las funciones de distribución acumulada de los valores predichos, que indican la fracción de clientes que paga la deuda y el modelo predijo que la pagarían (Verdaderos Positivos) y de clientes que no pagan y el modelo predijo que la pagarían (Falsos Positivos). De esta forma, el MSP es la suma de las ganancias maximizadas en todos los segmentos:

$$MSP = \sum_{i=1}^{q} \max_{\phi_i} P_i(\phi_i). \tag{3.3}$$

3.2. Modelos causales en cobranzas

Esta sección extiende el enfoque previamente presentado para modelos predictivos en cobranza hacia un marco de modelos de aprendizaje causal, con el objetivo de estimar el efecto de las acciones para cobrar la deuda sobre el comportamiento de pago de los clientes. En primer lugar, se describen los modelos de aprendizaje automático causal, luego, se incorporan modelos basados en beneficios (modelos profit), los cuales permiten alinear la estimación del efecto de una acción con la maximización del beneficio para la empresa. Finalmente, se presenta un enfoque multisegmento que busca capturar diferencias estructurales entre grupos de clientes, con el fin de mejorar la asignación de estrategias de cobranza. Esta estructura metodológica mantiene coherencia con la sección anterior, integrando consideraciones causales y de rentabilidad en el proceso de modelación.

3.2.1. Modelos de aprendizaje causal

Los modelos causales tienen el objetivo de estimar el efecto neto de aplicar un tratamiento a un resultado, para esto, calculan el efecto del tratamiento individual (ITE). En esta tesis, se considera un caso exitoso como la variable objetivo del marco causal, es decir, Y=1 si el cliente pagó la deuda y Y=0 en caso contrario, mientras que la técnica de contacto para el cobro de la deuda define los grupos de tratamiento y control, W=1 si se le ofrece al deudor un incentivo para pagar su deuda y W=0 en caso contrario.
De esta forma, a partir del modelo causal se obtienen dos valores:

- ITE control: Probabilidad de que se cumpla la variable objetivo considerando que el individuo no se somete a tratamiento. Para el caso de este estudio, será la probabilidad de pago, en caso de no ser expuesto al tratamiento (no se contactó para cobrar la deuda).

$$P(Y = 1|x, W = 0) \tag{3.4}$$

- ITE Tratamiento: Probabilidad de que se cumpla la variable objetivo considerando que el individuo se somete a tratamiento. Para el caso de este estudio, será la probabilidad de pago, en caso de ser expuesto al tratamiento (se contactó para

cobrar la deuda).

$$P(Y = 1|x, W = 1) \tag{3.5}$$

Dado $X = \{x_1, ..., x_n\}$ las variables independientes, $Y \in \{0, 1\}$ la variable objetivo, y $W \in \{0, 1\}$ la variable de tratamiento.

A partir de estas dos probabilidades se calcula la medida Uplift del modelo, que consiste en la resta de las dos probabilidades definidas anteriormente y se define a través de la función $\tau(x)$ que se muestra a continuación:

$$\tau(x) = P(Y = 1|x, W = 1) - P(Y = 1|x, W = 0)$$

Para esto, se aplican modelos de aprendizaje automático causal que clasifican las muestras según su probabilidad de respuesta a un tratamiento. En este trabajo se aplican 4 clasificadores descritos en el capítulo anterior, estos son los árboles de decisión, extremely randomized trees, regresión logística y LightGBM (LGBM). Para la implementación de estos modelos se utiliza el método de *meta-leaners*, específicamente los enfoques T-learner, S-learner y MOA descritos en el Capítulo 2.

Al igual que en la aplicación de modelos de aprendizaje tradicional, se divide el conjunto de datos en dos, entrenamiento y testeo. Al tener una variable objetivo y otra variable de tratamiento, se utiliza una técnica estratificada para que la proporción de muestras sin tratamiento/con tratamiento y pagan/no pagan sea la misma en ambos conjuntos. Luego, se realiza una validación cruzada de 10 particiones dentro del conjunto de entrenamiento para explorar diferentes configuraciones de parámetros. Una vez definida la mejor configuración de parámetros, se construyen los clasificadores causales con todo el conjunto de entrenamiento y, posteriormente, se aplican al conjunto de prueba.

Este trabajo se aplica sobre datos observacionales donde la falta de control sobre la asignación puede generar sesgos debido a diferencias sistemáticas entre los grupos tratado y de control. La falta de control sobre la asignación puede generar sesgos debido a diferencias sistemáticas entre los grupos tratado y de control. Para mitigar estos problemas, se aplican métodos que buscan equilibrar las características observables entre ambos grupos y el desequilibrio de clases; en este trabajo se aplican 3 estrategias de muestreo para el sesgo de NRA: sin remuestreo (NoRS), Pareamiento por puntaje de propensión (PSM) y ProSOM descritos en la revisión bibliográfica.

Por último, las métricas que se calculan para la evaluación y comparación de modelos son (Devriendt, Moldovan & Verbeke, 2018b):

- Uplift al 10 % y 30 % : Medida de uplift por segmento calculada para el 10 % y 30 % superior.
- El coeficiente de Qini (Qini): Esta métrica agrega la información de la curva de Qini, que estima la diferencia acumulada en términos de tasa de respuesta entre los dos grupos. A diferencia de la AUUC, esta medida tiene en cuenta la diagonal de un modelo aleatorio.

3.2.2. Modelos de aprendizaje causal basados en beneficio

Siguiendo la misma lógica de los modelos tradicionales, se aplican modelos de aprendizaje causal basados en beneficio. Esto con el fin de medir los costos y beneficios de estos modelos para identificar el que mayor retorno le entrega a la empresa.
Para los parámetros de entrada de la matriz $\dot{\mathbf{CB}}$ tenemos los mismos que en los modelos tradicionales, donde el beneficio viene dado por la cantidad de deuda que recuperará la empresa cuando un cliente paga. De esta forma, la matriz $\dot{\mathbf{CB}}$ queda de la siguiente forma:

$$\dot{\mathbf{CB}} := \begin{array}{c} \begin{array}{cc} W=0 & W=1 \end{array} \\ \left[\begin{array}{cc} 0 & -\mathbf{f} \\ \mathbf{b} & \mathbf{b}-(\mathbf{d}+\mathbf{f}) \end{array}\right] \end{array} \begin{array}{l} Y=0 \\ Y=1 \end{array} \tag{3.6}$$

De acuerdo con estudios previos sobre marcos sensibles al costo para la clasificación tradicional (por ejemplo, (Maldonado, Domínguez, Olaya & Verbeke, 2021)), nos centramos en el beneficio incremental de la acción de cobro. Si se incluye $\mathbf{b}$ en la ecuación (3.1) para los falsos negativos (FN), el modelo de clasificación se ve "recompensado" por no identificar a los deudores que habrían pagado independientemente de la intervención. En contraste, el marco sensible al costo para modelos causales considera el beneficio de los deudores que se autocorrigen en la ecuación (3.6), siguiendo las directrices para la predicción de la retención de clientes propuestas en (Verbeke, Olaya, Berrevoets, Verboven & Maldonado, 2020).

Clasificación causal multisegmento orientada a las ganancias

A diferencia de la clasificación tradicional, que predice el estado Y, el marco MSP para los modelos causales ($M\dot{S}P$) estima el cambio de comportamiento causado por el tratamiento W (la oferta). Definimos $W = 1$ si se contacta a un deudor con el incentivo d_i, y $W = 0$ en caso contrario.

Definición 3.2.1 (Beneficio Causal Individual)**.** *El beneficio incremental $\dot{P}$ para un deudor k en el segmento i es la diferencia entre el beneficio esperado bajo tratamiento y el valor de referencia sin contacto ($E[P|W = 0] = 0$):*

$$\dot{P}_k = \underbrace{P(Y = 1|X_k, W = 1)(\bar{b}_i - d_i)}_{\textit{Recuperación con Tratamiento}} - \underbrace{P(Y = 1|X_k, W = 0)\bar{b}_i}_{\textit{Recuperación Espontánea}} - f \tag{3.7}$$

Reordenando los términos, se puede derivar el beneficio en términos del efecto del tratamiento individual (incremento): $\tau(X_k) = P(Y = 1|X_k, W = 1) - P(Y = 1|X_k, W = 0)$.

Teorema 3.2.2 (Límite de decisión de beneficio causal)**.** *Una estrategia óptima de cobro de deudas se dirige al deudor k en el segmento i si y solo si su incremento estimado $\tau(X_k)$ satisface:*

$$\tau(X_k) > \frac{f}{\bar{b}_i} + \frac{d_i}{\bar{b}_i} P(Y = 1|X_k, W = 1) \tag{3.8}$$

Demostración. Sea $\dot{P}_k > 0$. De la definición: $\tau(X_k)\bar{b}_i - P(Y = 1|W = 1)d_i - f > 0$. Dividiendo por $\bar{b}_i$ y despejando $\tau(X_k)$ se obtiene el límite. Este límite es una adaptación específica de la regla de decisión de Verbeke, Olaya, Guerry & Van Belle (2023), donde el incremento necesario para justificar el tratamiento aumenta linealmente con la probabilidad de reembolso bajo tratamiento, ya que los casos seguros aumentan el costo de oportunidad del incentivo d_i, dado que su condonación genera pérdidas de margen. □

La regla de decisión es la siguiente: Un deudor del segmento i solo se selecciona si el beneficio incremental esperado es positivo ($\dot{P}_k > 0$). Esto garantiza que la campaña solo se dirija a individuos donde el incremento estimado τ sea lo suficientemente alto como para compensar tanto el costo operativo f como los ingresos perdidos por el incentivo d_i. Al sumar estas ganancias incrementales en todos los segmentos $i \in \{1, \ldots, q\}$, el marco $M\dot{S}P$ identifica la estrategia de tratamiento óptima que maximiza el valor adicional total generado por el departamento de cobranza en comparación con un escenario pasivo de no acción.

Sea n_i el número de instancias en el segmento i, y h_i la fracción del segmento objetivo del modelo (el umbral). El beneficio causal promedio para el segmento en el umbral h_i es:

$$\dot{P}_i(h_i) = \frac{1}{n_i} \sum_{k=1}^{n_i} h_i \dot{P}_k \tag{3.9}$$

La métrica de evaluación final, $M\dot{S}P$, es la suma de los beneficios causales promedio máximos encontrados en todos los segmentos:

$$M\dot{S}P = \sum_{i=1}^{q} \max_{h_i} \dot{P}_i(h_i). \tag{3.10}$$

Esta formulación ofrece una ventaja significativa: $P(Y = 1|W = 1)$ encapsula de forma natural la respuesta del deudor al incentivo. A diferencia del marco estándar, que requiere un parámetro externo de probabilidad de aceptación, $M\dot{S}P$ permite calcular la elasticidad a partir de los datos causales, identificando donde el saldo $\bar{b}_i$ es lo suficientemente alto como para justificar la pérdida en incentivos de aquellos que podrían haber pagado de todos modos.

3.3. Extensión de modelos causales basados en beneficio para venta cruzada

En la misma línea que los modelos aplicados en la sección anterior se realizó otra aplicación de modelos de aprendizaje causal basados en beneficio a una base de datos de venta cruzada. A continuación se detalla la base de datos y el diseño experimental aplicado.

3.3.1. Conjunto de datos

Se utiliza un conjunto de datos proporcionado por un grupo financiero chileno que contiene 61,363 recomendaciones de clientes realizadas por vendedores pertenecientes a distintas unidades de negocio con el objetivo de promover la venta cruzada. La información fue recopilada entre 2005 y 2011 e incluye datos relevantes sobre los vendedores y las unidades de negocio a las que pertenecen.

En este análisis, los vendedores son capaces de recomendar clientes de una unidad de negocio a otra. Una recomendación corresponde a un mensaje escrito enviado por un vendedor (emisor) a otro (receptor), en el cual el primero anticipa una posible oportunidad de venta dentro de la unidad de negocio del segundo. Para fomentar este tipo de interacciones, el grupo financiero implementó un sistema de incentivos que otorga recompensas monetarias a los emisores cuando sus recomendaciones derivan en la venta de un nuevo producto por parte del receptor. A este resultado se le denomina recomendación exitosa. Adicionalmente, la empresa establece periodos de campaña entre unidades de negocio específicas, cuyo objetivo es estimular el envío de recomendaciones mediante incentivos no monetarios, tales como dispositivos electrónicos o viajes.
En el anexo B se presenta la estadística descriptiva de nuestros datos brutos. Entre 2005 y 2011, el 19.94 % de las recomendaciones dentro de la empresa tuvieron éxito, y solo el 7.41 % se incluyeron en la campaña. Además, en el anexo C se detalla una serie de variables que fueron creadas para el modelo predictivo y reflejan la relación entre el emisor y el receptor, así como factores adicionales que pueden ser relevantes según estudios de marketing.

3.3.2. Propuesta metodológica

Los modelos de aprendizaje causal permiten estimar el Efecto del Tratamiento Individual (ITE) con el objetivo de ordenar a los individuos según el impacto que una intervención produce sobre ellos. En este contexto, la intervención corresponde al incentivo para generar una recomendación entre unidades de negocio. En este trabajo, se define como variable objetivo una recomendación exitosa, de modo que $Y = 1$ cuando una recomendación se traduce en una venta e $Y = 0$ en caso contrario. A su vez, los distintos periodos de campaña permiten establecer la asignación al tratamiento: $W = 1$ cuando la recomendación se realiza durante un periodo de campaña y $W = 0$ cuando ocurre fuera de este.
Se aplican 6 modelos de clasificación para determinar qué derivaciones deberían incentivarse, estos modelos corresponden a árboles de decisión, random forest, extremely randomized trees, gradient boosting, LGBM, extreme gradient boosting. Se aplican 4 enfoques causales: MOA, S-learner, T-learner y Causal forest. Al igual que en cobranza se divide el conjunto de datos en entrenamiento y validación, luego se aplica valida-

ción cruzada dividiendo en 5 subconjuntos los datos de entrenamiento. Por último, se calculan las siguientes métricas: coeficiente Qini y el profit promedio de la solución.
Se define un marco orientado a maximizar el beneficio. En el contexto de venta cruzada, una acción se considera exitosa cuando el cliente recomendado acepta la propuesta, es decir, cuando y=1, independientemente de si la recomendación fue incentivada (w=1) o no (w=0). El beneficio obtenido en este caso se define como b_{cs}, el cual corresponde al valor del ciclo de vida del cliente (CLV) asociado al nuevo producto. La Matriz OB_{cs} presenta los beneficios asociados al resultado; esta matriz es similar a la utilizada en la predicción de abandono de clientes, ya que en ambos casos el cliente puede decidir permanecer en la empresa; por lo tanto, el beneficio corresponde al CLV y no depende del tratamiento aplicado.

$$\mathbf{OB_{cs}} := \begin{array}{c} \begin{matrix} W=0 & W=1 \end{matrix} \\ \begin{bmatrix} 0 & 0 \\ b_{cs} & b_{cs} \end{bmatrix} \end{array} \begin{matrix} Y=0 \\ Y=1 \end{matrix} \tag{3.11}$$

En relación a los costos asociados a la campaña, los incentivos solo se pagan cuando una recomendación termina en una venta. Se define como C_{cs} el costo que asume la empresa al entregar un incentivo a los vendedores cuando una recomendación resulta exitosa, independientemente de si esta fue parte de la campaña o no. Adicionalmente, se define C_{camp} como el bono adicional que la empresa entrega cuando la recomendación forma parte de la campaña de venta cruzada (w=1). Estos se traducen a la matriz TC_{cs} que representa los costos del tratamiento.

$$\mathbf{TC_{cs}} := \begin{array}{c} \begin{matrix} W=0 & W=1 \end{matrix} \\ \begin{bmatrix} 0 & 0 \\ C_{cs} & C_{camp}+C_{cs} \end{bmatrix} \end{array} \begin{matrix} Y=0 \\ Y=1 \end{matrix} \tag{3.12}$$

Finalmente, la fórmula para calcular el beneficio causal para el caso de venta cruzada quedaría de la siguiente forma:

$$\dot{P}_{cs} = \pi_1^C F_1^C(\dot{\phi})(b_{cs} - C_{cs}) + \pi_1^T \left(1 - F_1^T(\dot{\phi})\right)(b_{cs} - C_{camp} - C_{cs}) \tag{3.13}$$

La empresa estimó una rentabilidad promedio de €239 por referencia exitosa, con costos sugeridos de C_{cs} =€10 y C_{camp} =€5. No obstante, se plantea optimizar estos valores mediante un análisis de sensibilidad.
Para determinar el umbral óptimo, se utilizan dos enfoques: uno insensible al costo, basado en la curva Qini y enfocado en maximizar ganancias acumuladas, y otro sensible al costo, que maximiza el beneficio causal promedio. Ambos métodos ordenan a los clientes de forma distinta generando diferentes curvas Qini.

3.4. Bounded ITE

En la siguiente sección se describe la metodología utilizada para la propuesta de un nuevo enfoque, Bounded ITE, en los modelos de aprendizaje causal. En este contexto, la medida uplift ha sido ampliamente utilizada, debido a que, a través de este valor se puede organizar a los individuos dentro de cuatro categorías:

- Asegurados: Son aquellos individuos que, independientemente de ser expuestos al tratamiento, van a tener una respuesta positiva, por lo tanto, no tiene sentido aplicar un tratamiento a estos individuos ya que sería una pérdida de tiempo y recursos.

- Persuasibles: Son aquellos individuos que tienen un impacto positivo, al ser expuestos al tratamiento y que, en caso de no ser tratados, es posible que tengan una respuesta adversa. Este es el grupo de interés, en ellos conviene invertir recursos.

- No molestar: Son los individuos que tienen una reacción negativa frente al tratamiento, es decir, en caso de ser expuestos al tratamiento, no pagarán o viceversa. Debido a esto, no hay que aplicar tratamiento a estos individuos, ya que tiene un impacto negativo.

- Causas perdidas: Son aquellas observaciones que, independientemente de ser expuestas al tratamiento o no, no van a cambiar de opinión. Claramente, no es necesario exponer a estos individuos al tratamiento, ya que la decisión, por parte del cliente, ya está tomada.

El objetivo es filtrar los individuos que probablemente sean causas perdidas o asegurados definiendo límites para algunas probabilidades. Las instancias con $\tau(x)$ más grandes son el objetivo. La acción generalmente se aplica a personas para quienes el ITE es mayor que un umbral dado.

Los individuos catalogados como no molestar tendrán un $\tau(x)$ negativo y, por lo tanto, el enfoque tiene éxito en filtrar a esos individuos. Sin embargo, afirmamos en este estudio que la ITE no es tan efectiva para distinguir entre persuasibles (nuestro objetivo), causas perdidas y asegurados.

Por ejemplo, $\tau(x) = 0.15$ significa que el tratamiento mejora la probabilidad de un resultado positivo en 0.15. Sin embargo, no puede distinguir entre los siguientes escenarios: $P(Y = 1|x, W = 1) = 0.2$ $P(Y = 1|x, W = 0) = 0.05$ donde el individuo es una causa perdida, $P(Y = 1|x, W = 1) = 0.95$ y $P(Y = 1|x, W = 0) = 0.8$ donde el individuo es asegurado y $P(Y = 1|x, W = 1) = 0.55$ y $P(Y = 1|x, W = 0) = 0.4$ donde el individuo es persuasible.

Teniendo esto en cuenta, se plantea una versión acotada del ITE (BITE), de la siguiente manera:

$$\tau_{BITE}(x) := \begin{cases} 0 & , \text{ if } P_t < \beta_t \\ 0 & , \text{ if } P_c > \beta_c \\ \tau_{ITE}(x) & , \text{ otherwise,} \end{cases} \tag{3.14}$$

En base a la revisión bibliográfica analizada (Devriendt, Moldovan & Verbeke, 2018b), se utilizan seis clasificadores con sus configuraciones predeterminadas de hiperparámetros para analizar el desempeño de los enfoques ITE y BITE. Los Modelos utilizados corresponden a árboles de decisión, gradient boosting, random forest, LightGBM, eXtreme gradient boosting y extremely randomized trees.

Como punto de comparación, también se consideró un enfoque de clasificación tradicional. En este caso, los modelos predicen la probabilidad de un resultado positivo ($Y = 1$) para clasificar a los clientes. Estos modelos se entrenan combinando los datos de los grupos de tratamiento y control.

Para la implementación de los experimentos se utilizó el paquete *scikit-learn*. Los valores de ITE y BITE se calcularon mediante el enfoque *S-learner*. La evaluación de los métodos se realizó utilizando un procedimiento de validación cruzada de diez particio-

nes. A partir de este proceso se calcularon métricas de uplift al 5 %, 10 %, 15 %, 20 %, 25 % y 30 %.

No se calculan métricas como AUUC o el coeficiente Qini, ya que el propósito de la métrica BITE es mejorar la selección de clientes mediante el ajuste de los valores de ITE a partir de la incorporación de límites. Para determinar estos límites dentro del enfoque BITE, se exploraron diferentes combinaciones de parámetros utilizando una validación cruzada de diez particiones dentro del conjunto de entrenamiento: $\beta_c \in \{0.75,0.8,0.85,0.9,0.95\}$ y $\beta_t \in \{0.05,0.1,0.15,0.2,0.25\}$.

3.4.1. Conjuntos de datos

Para este análisis se consideraron 7 conjuntos de datos, los cuales se encuentran descritos en el anexo D. En la Tabla 3.2 se muestra un resumen de los datos donde se incluye el número de observaciones (*# muestras*) y predictores (*# variables*), así como las proporciones de la variable causal (*%(W=1, W=0)*) y objetivo (*%(Y=1, Y=0)*), respectivamente.

Cuadro 3.2: Estadísticas descriptivas de los conjuntos de datos

nombre	# muestras	# variables	%(W=1, W=0)	%(Y=1, Y=0)
Crosssell	20,297	8	(5.87, 94.13)	(8.55, 61.45)
HillAll	28,375	20	(56.78, 43.22)	(14.08, 85.92)
HillMen	26,951	19	(44.60, 55.40)	(13.97, 86.03)
HillWomen	26,936	19	(44.75, 55.25)	(12.68, 87.32)
Informs	112,104	31	(44.66, 55.34)	(3.55, 96.45)
Insurance	12,620	68	(44.71, 55.29)	(20.10, 79.90)
Udemy	56,506	16	(11.51, 88.49)	(3.18, 96.82)

Capítulo 4

Resultados y Análisis

En este capítulo se presentan los resultados obtenidos a partir de la implementación de los modelos propuestos, organizados bajo un marco común de aplicación de modelos de aprendizaje automático y su evaluación basado en beneficios. En primer lugar, se reportan los resultados de los modelos de aprendizaje tradicional en cobranza. A continuación, se presentan los resultados de la incorporación de modelos causales en el mismo contexto, evaluando su capacidad predictiva y su impacto en términos de rentabilidad. Posteriormente, este análisis se extiende al caso de venta cruzada, donde se examina el desempeño de los modelos causales bajo el mismo criterio de evaluación. Finalmente, se presentan los resultados asociados a la propuesta de definir limites al efecto del tratamiento individual, analizando su incidencia en la estimación y desempeño de los modelos. En conjunto, los resultados permiten comparar de manera consistente distintos enfoques de modelación y evaluar su contribución a la optimización de decisiones.

4.1. Mejora de la gestión de cobranza mediante información del *contact center*

En esta sección se presentan los resultados obtenidos al aplicar modelos de clasificación tradicional en cobranzas, además, se analiza la repercusión de utilizar diferentes conjuntos de datos en los modelos.

A continuación, se describen los resultados estadísticos 4.1.1 y el análisis de variables mas relevantes 4.1.2 para las 3 tareas predictivas:

4.1.1. Resultados de los modelos predictivos para las 3 tareas

La Tabla 4.1 presenta los resultados según la métrica AUC para cada tarea predictiva, modelo de clasificación y conjunto de datos. Se puede observar que el uso exclusivo de variables financieras genera un desempeño inferior en comparación con otras fuentes de información, en cambio, el conjunto Unificado muestra un rendimiento ligeramente superior al conjunto *contact center*. A partir de esto, se puede inferir que el éxito de un contacto y promesa de pago depende principalmente de variables relacionadas con el momento óptimo para llamar y el historial de interacciones con el deudor, más que de sus características financieras.

Para la tarea de predicción de pago, el mejor rendimiento también se logra con el conjunto de datos unificado, pero a diferencia de los casos anteriores, el rendimiento de los conjuntos de datos financiera y de *contact center* en muy similar. En relación a lo anterior, podemos determinar que para la predicción de pago las variables financieras toman mayor relevancia que para las otras tareas predictivas.

El mejor desempeño se alcanza con el clasificador de redes neuronales integrando ambas fuentes en las 3 tareas predictivas, logrando resultados que pueden considerarse muy buenos en el contexto del análisis de negocio (Baesens, 2014). Esto respalda tanto la relevancia de la integración de datos como la eficacia de las redes neuronales.

Además, la Figura 4.1 resume el desempeño predictivo del mejor clasificador para cada tarea y fuente de datos, permitiendo comparar el impacto de cada una. La predicción de pago destaca como la tarea con mejores resultados; sin embargo, las tres tareas presentan desempeños similares y elevados cuando se combinan ambas fuentes. Esto indica que pueden implementarse de manera complementaria en las políticas de cobranza y confirma que la integración de datos conduce sistemáticamente a mejores resultados en todas las tareas analizadas.

4.1.2. Resultados del análisis de variables para las 3 tareas predictivas

Por otro lado, se pueden obtener conocimientos adicionales analizando la influencia de las variables en el clasificador final mediante el método TreeSHAP. A continuación se presenta el análisis de las variables para las 3 tareas predictivas:

Cuadro 4.1: Resultados predictivos para la tarea de predicción de contacto exitoso.

Tarea predictiva	Modelo	Conjunto de datos		
		Unificado	Financiero	Contact Center
Contacto exitoso	Regresión logística	0.770	0.607	0.762
	Random Forest	0.776	0.598	0.777
	Árbol de decisión	0.788	0.589	0.785
	K-Vecinos	0.768	0.624	0.747
	Red neuronal	**0.802**	0.607	0.780
Promesa de pago	Regresión Logística	0.746	0.629	0.721
	Random Forest	0.729	0.621	0.736
	Árbol de decisión	0.773	0.613	0.755
	K-Vecinos	0.737	0.599	0.697
	Red neuronal	**0.781**	0.632	0.76
Pago	Regresión logística	0.795	0.708	0.742
	Random Forest	0.802	0.696	0.763
	Árbol de decisión	0.811	0.707	0.769
	K-Vecinos	0.788	0.686	0.685
	Red neuronal	**0.826**	0.718	0.766

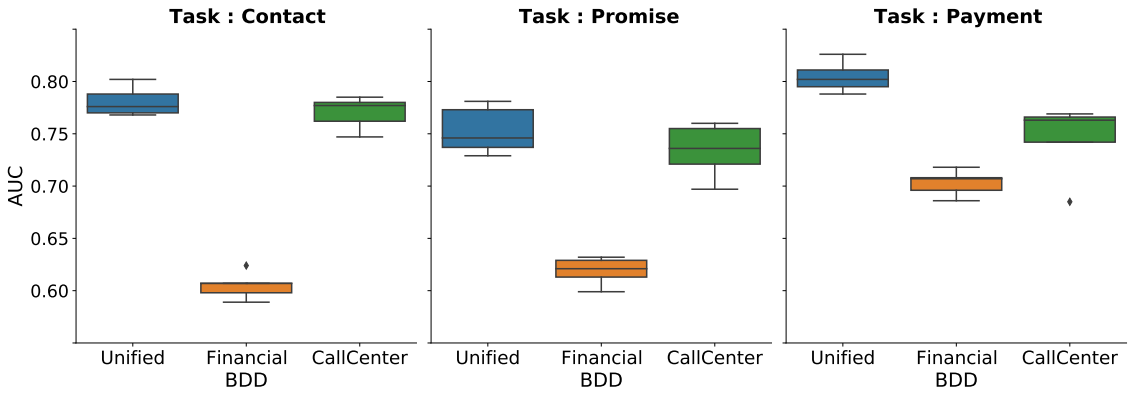


Figura 4.1: Resumen del rendimiento para cada tarea de predicción y fuente de datos.

Contacto Exitoso En la Figura 4.2 se ven los resultados obtenidos del método TreeSHAP para la tarea predictiva de contacto exitoso con el conjunto de datos unificado. El gráfico muestra la relevancia de cada una de las variables, ordenando de más a menos relevante. A partir de este análisis, se puede concluir que las variables más relevantes corresponden a distintos conjuntos. Las primeras son del conjunto de datos de *Contact center*, número de llamadas y contactos pasados y si la llamada se hace los primeros 10 días del mes. A pesar de que en la sección anterior se detectó que las variables financieras obtienen

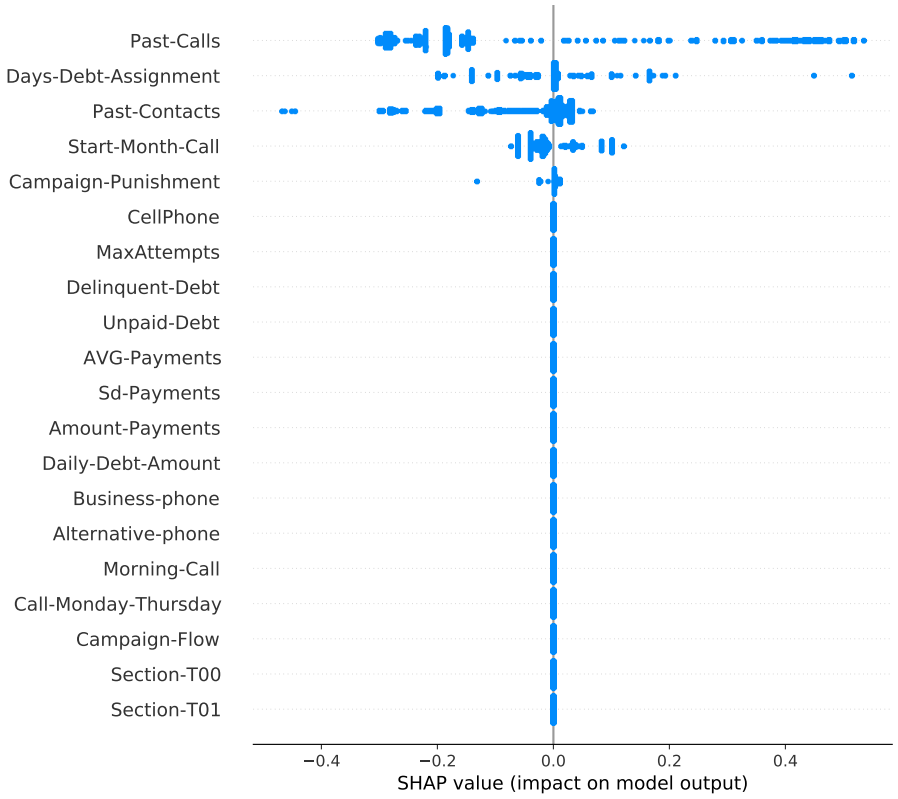


Figura 4.2: Resultado de análisis de variables para predicción de contacto exitoso.

un bajo rendimiento en esta tarea predictiva, igual se destacan dos dentro de las más relevantes, el número de días que tiene en mora (*Days-Debt-Assignment*) y si el deudor tiene un atraso de más de 30 días en el pago de su deuda (*Campaign-Punishment*). A partir de esto, se puede concluir que la información de la deuda pendiente influye en si el contacto es exitoso o no.

Promesa de pago La Figura 4.3 representa el análisis de variables con el método TreeSHAP para la tarea predictiva de promesa de pago con el conjunto de datos unificado. Al igual que la tarea de predicción de contacto, variables de contactos y llamadas pasadas con relevantes y se le incorporan las promesas pasadas. También se repite si la llamada se hizo los primeros 10 días del mes, los días que tienen en mora y si el deudor tiene mas de 30 días en mora. Además, del conjunto de datos financiero se incluye el monto de deuda morosa.

A partir de esto, se determina que la información de la deuda influye más en la tarea de predicción de promesa de pago ya que 3 de las 7 variables más relevantes corresponden al conjunto de datos financiero. También existe coherencia entre las dos tareas predictivas

ya que las 5 variables relevantes para predecir el contacto exitoso también lo son para predecir una promesa de pago

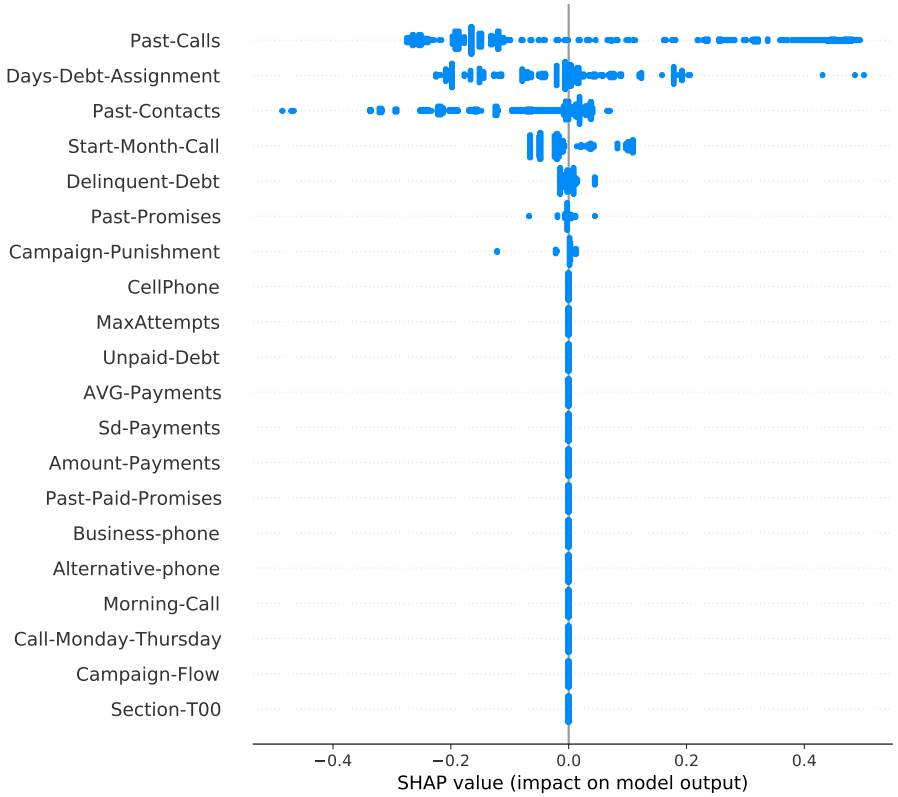


Figura 4.3: Resultado de análisis de variables para predicción de promesa de pago.

Pago de la deuda La Figura 4.4 representa el análisis de variables para la tarea de predicción de pago. De las 20 variables 8 son relevantes para esta tarea predictiva, 3 son del conjunto de datos de *contact center*, donde se repiten las llamadas y contactos pasados y se incorpora si la llamada se hace en la mitad del mes (dentro de los días 10 y 20). Por el lado de las variables financieras, toman relevancia el número de días que está en mora, la cantidad de deuda, la cantidad de pagos realizados anteriormente y si el nivel de riesgo es alto según la clasificación de riesgo del banco basado en los pagos pasados.

A diferencia del análisis realizado para las dos tareas anteriores, hay mas variables relevantes del conjunto de datos financiero que del de *contact center*, esto era esperable dado que el rendimiento predictivo con el conjunto de datos financiero y *contact center* era muy similar. Además, toma relevancia la variable que indica si la llamada se realizó a mitad de mes, esto difiere de las tareas anteriores donde la variable relevante era si

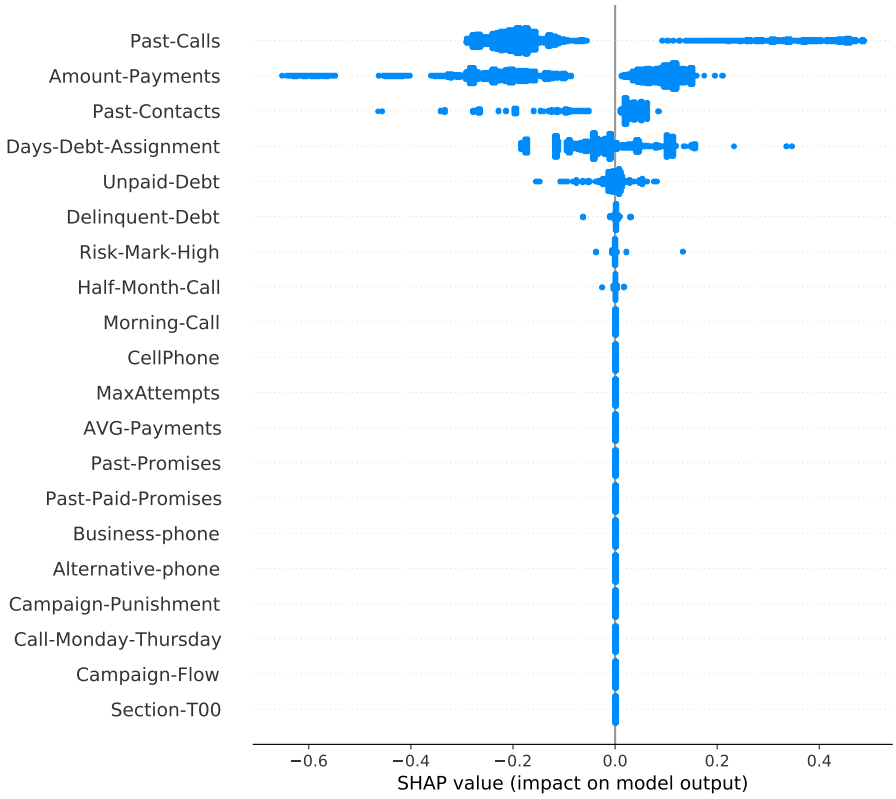


Figura 4.4: Resultado de análisis de variables para predicción de pago de deuda

la llamada se realizaba los primeros 10 días del mes.

En base a estos dos análisis, por un lado los resultados predictivos y por otro las variables más relevantes, se concluye que la predicción de contacto exitoso constituye una herramienta útil desde el punto de vista operativo, ya que permite identificar los momentos y canales más adecuados para contactar a los clientes, apoyándose además en modelos interpretables que facilitan la extracción de reglas prácticas. Por su parte, la predicción de pago, aunque ampliamente utilizada en contextos tradicionales de cobranza, presenta limitaciones al incluir clientes cuya conducta de pago no depende de la acción de contacto.

En este contexto, la predicción de promesa de pago se posiciona como la tarea más relevante, dado que se enfoca en identificar a aquellos clientes que al contactarlos se obtiene un beneficio, como lo es la promesa de pago. Este enfoque, alineado con los principios de la inferencia causal, permite distinguir entre clientes persuasibles y aquellos en los que la acción no genera impacto.

4.2. Aplicación de enfoque orientado a las ganancias para la clasificación tradicional y causal

Se compararon las siguientes estrategias para evaluar la efectividad de nuestro marco propuesto en el capítulo anterior:

- Clasificación tradicional insensible al costo (CI): Siguiendo (Verbeke, Olaya, Guerry & Van Belle, 2023), consideramos el enfoque óptimo insensible al costo basado en un umbral $\phi^* = 0.5$ (probabilístico) para la segmentación de clientes, con precisión y AUROC para la selección de hiperparámetros.

- Modelos causales insensibles al costo: Nos dirigimos al 10 % y al 50 % de los clientes en el conjunto de prueba clasificados mediante ITE para la campaña de cobro de deudas, utilizando el coeficiente Qini para la selección de hiperparámetros.

- Clasificación tradicional y causal basada en beneficios (profit_s): Tanto la selección de hiperparámetros como los umbrales de segmentación se obtienen mediante el criterio $MP/\dot{MP}$ (Verbeke, Olaya, Guerry & Van Belle, 2023).

- El marco multisegmento y multiumbral propuesto para la clasificación tradicional y modelos causales (profit_m): Tanto la selección de hiperparámetros como los umbrales de segmentación se obtienen mediante el criterio $MSP/\dot{MSP}$.

Se adoptó la función profit `empChurn` para implementar los experimentos. Para esta función, se utilizan los hiperparámetros predeterminados para `gamma`, la probabilidad de que un deudor acepte el descuento en la deuda, modelada como una distribución beta con $\alpha = 6$ y $\beta = 14$; (por ejemplo, (Maldonado, Domínguez, Olaya & Verbeke, 2021)). El beneficio (a nivel de cliente) $\mathbf{b}$ se calcula a partir de los datos como el saldo pendiente de cada deudor multiplicado por el porcentaje promedio de la deuda pagada de los clientes.

Para los enfoques no segmentados, se exploraron los siguientes valores para el descuento en la deuda d y el costo de contacto f: $d \in \{\bar{b}/20, \bar{b}/10, \bar{b}/8, \bar{b}/5, \bar{b}/3, \bar{b}/2\}$, $f \in \{1, 2, 5\}$. El objetivo con este entorno experimental es extender los análisis basados en la rentabilidad ($d = \bar{b}/20$, $f = 1$) a un escenario más completo que considere costos de contacto y descuentos en la deuda más altos, ya que la dinámica de cobro de deudas puede diferir

en comparación con la predicción de abandono de clientes, la principal aplicación para la que se ha desarrollado este enfoque.

El marco de umbrales múltiples considera que ofrecer un descuento único, por ejemplo, $d = \bar{b}/20$, a todos los deudores no es realista, al igual que la suposición de que la probabilidad de aceptación se mantiene constante en los diferentes niveles de oferta. En cambio, mantener una proporción consistente dentro de cada segmento para la relación entre la deuda b y el incentivo d: $d_i \in \{\bar{b}_i/20, \bar{b}_i/10, \bar{b}_i/8, \bar{b}_i/5, \bar{b}_i/3, \bar{b}_i/2\}$. En consecuencia, esta escala proporcional garantiza que la elasticidad de respuesta al incentivo se conserve en relación con el saldo pendiente promedio dentro del segmento $\bar{b}_i$.

En base a las estrategias nombradas, primero se presentan los resultados para la clasificación tradicional y luego para la clasificación causal.

4.2.1. Resumen de resultados modelos de aprendizaje tradicional basados en beneficio

Primero se analiza si el uso de métricas basadas en ganancias afecta negativamente el rendimiento estadístico en comparación con un enfoque insensible al costo. La Tabla 4.2 presenta los valores de accuracy y AUROC obtenidos por el clasificador de mejor desempeño (LGBM) bajo tres estrategias de evaluación: insensible al costo (CI), profit_s y profit_m, considerando distintos niveles de incentivo d y costo de contacto f. Para cada combinación de métrica y parámetros, el mejor resultado se resalta en negrita. Profit_s y profit_m arrojan resultados idénticos en estas métricas, ya que únicamente difieren en el orden de los deudores para el cálculo de las ganancias.

En términos generales, se observa que el enfoque insensible al costo tiende a presentar un desempeño ligeramente superior en métricas estadísticas como AUROC y accuracy frente a los métodos basados en ganancias. No obstante, estas diferencias son marginales y los resultados se mantienen comparables en la mayoría de los escenarios, lo que indica que la incorporación de métricas de ganancias no disminuye la capacidad predictiva de los modelos.

Los resultados en términos de beneficio (MP/MSP) se presentan en la Tabla 4.3. A diferencia de lo observado con las métricas estadísticas, al analizar la rentabilidad se identifican diferencias entre los enfoques. En particular, la métrica de beneficio multi-

Cuadro 4.2: Resumen de resultados para la comparación del rendimiento estadístico entre métricas de ganancias y enfoque insensible al costo.

			d (Incentivo/Descuento deuda. €)						
f(€)	Estrategia	Clasificador	$\bar{b}/2$	$\bar{b}/3$	$\bar{b}/5$	$\bar{b}/8$	$\bar{b}/10$	$\bar{b}/15$	$\bar{b}/20$
Accuracy como métrica de rendimiento									
$f=1$	CI	LGBM	80.2	**80.2**	80.2	80.2	80.2	**80.2**	**80.2**
	$\text{profit}_{s,m}$	LGBM	**80.8**	**80.2**	**80.6**	**80.4**	**80.4**	79.8	**80.2**
$f=2$	CI	LGBM	**80.2**	80.2	80.2	80.2	**80.2**	**80.2**	80.2
	$\text{profit}_{s,m}$	LGBM	80.0	**81.0**	**80.3**	**80.9**	79.8	**80.2**	**80.5**
$f=5$	CI	LGBM	80.2	80.2	80.2	80.2	**80.2**	**80.2**	**80.2**
	$\text{profit}_{s,m}$	LGBM	**80.4**	**80.6**	**80.5**	**80.3**	**80.2**	80.1	80.1
AUROC como métrica de rendimiento									
$f=1$	CI	LGBM	**84.3**	**84.3**	**84.3**	**84.3**	**84.3**	**84.3**	**84.3**
	$\text{profit}_{s,m}$	LGBM	84.1	84.2	84.0	**84.3**	84.1	84.1	**84.3**
$f=2$	CI	LGBM	**84.3**	**84.3**	**84.3**	**84.3**	**84.3**	**84.3**	**84.3**
	$\text{profit}_{s,m}$	LGBM	**84.3**	84.2	84.1	84.2	**84.3**	**84.3**	**84.3**
$f=5$	CI	LGBM	**84.3**	**84.3**	**84.3**	**84.3**	**84.3**	**84.3**	**84.3**
	$\text{profit}_{s,m}$	LGBM	**84.3**	84.1	**84.3**	**84.3**	**84.3**	**84.3**	84.1

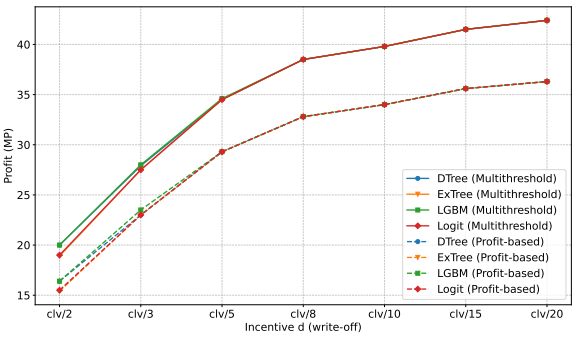
segmento (profit_m) muestra mejoras significativas frente a la clasificación insensible al costo (CI) y al método basado en un umbral único de beneficio (profit_s). Este mejor desempeño se explica porque el enfoque multisegmento permite distinguir segmentos con deudas de bajo monto que no resultan rentables de intervenir, al mismo tiempo que optimiza aquellos grupos con deudas más elevadas, donde el potencial de ganancia es mayor.

Cuadro 4.3: Resumen de resultados en base a métricas de ganancias.

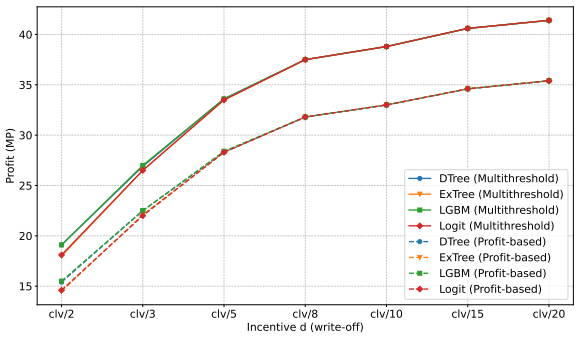
			d (Incentivo/Descuento deuda. €)						
f(€)	Estrategia	Clasificador	$\bar{b}/2$	$\bar{b}/3$	$\bar{b}/5$	$\bar{b}/8$	$\bar{b}/10$	$\bar{b}/15$	$\bar{b}/20$
$f=1$	CI	LGBM	9.6	13.2	16.0	17.6	18.1	18.9	19.2
	profit_s	LGBM	16.4	23.5	29.3	32.8	34.0	35.6	36.3
	profit_m	LGBM	**20.0**	**28.0**	**34.6**	**38.5**	**39.8**	**41.5**	**42.4**
$f=2$	CI	LGBM	9.1	12.7	15.5	17.1	17.7	18.4	18.7
	profit_s	LGBM	15.5	22.5	28.4	31.8	33.0	34.6	35.4
	profit_m	LGBM	**19.1**	**27.0**	**33.6**	**37.5**	**38.8**	**40.6**	**41.4**
$f=5$	CI	LGBM	7.7	11.2	14.1	15.7	16.2	16.9	17.3
	profit_s	LGBM	12.8	19.7	25.5	28.9	30.1	31.6	32.4
	profit_m	LGBM	**16.4**	**24.2**	**30.7**	**34.6**	**35.9**	**37.6**	**38.5**

Se presenta el rendimiento en términos de beneficio máximo para los distintos clasi-

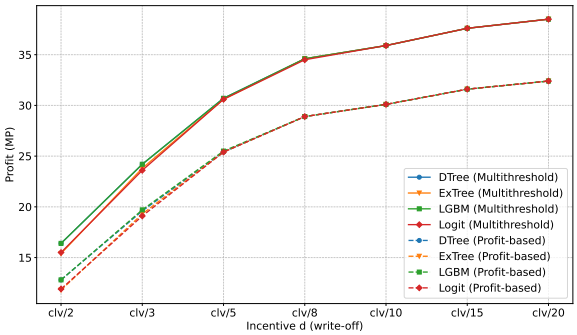
ficadores (DTree, ExTree, LGBM y Logit) en la Figura 4.5 para diferentes valores de incentivo (eje X) y costos de contacto (Figuras 4.5a, 4.5b y 4.5c). Además, se destaca con línea discontinua los clasificadores basados en beneficio con un único umbral y con línea continua los con multiumbral.



(a) Costo de contacto $f = 1$.



(b) Costo de contacto $f = 2$.



(c) costo de contacto $f = 5$.

Figura 4.5: Resumen de profit para la clasificación tradicional estándar y la estrategia multiumbral.

Como se muestra en la Figura 4.5, el enfoque multiumbral supera considerablemente a la clasificación basada en ganancias con un único umbral en todos los valores de d y f. El rendimiento entre los clasificadores base es bastante similar para cada estrategia, lo que se traduce en líneas superpuestas en los gráficos. Este análisis confirma las ventajas del enfoque multisegmento para lograr soluciones más rentables mediante la toma de decisiones por segmentos.

4.2.2. Resumen de resultados para la clasificación causal

Primero, se evalúa el desempeño del método propuesto en términos del coeficiente Qini y del beneficio máximo, comparándolo con un enfoque insensible al costo en el contexto de

modelos causales. La Tabla 4.4 muestra los resultados del coeficiente Qini y los valores de $\dot{MP}/\dot{MSP}$ (en euros) obtenidos por el clasificador causal con mejor rendimiento bajo cuatro estrategias de evaluación: enfoque insensible al costo (CI) dirigido al 10 % y 50 % del conjunto de prueba, y las métricas profit_s y profit_m para modelos causales, considerando distintos incentivos d y costo de contacto f. Para cada combinación de métrica y parámetros, el mejor resultado se destaca en negrita. Además, se presentan los resultados para la mejor combinación de *meta-learners* (S-Learner, T-Learner, MOA), clasificadores base (DTree, ExTree, LGBM y Logit) y estrategias de muestreo para el sesgo NRA (sin remuestreo o NoRS, PSM, ProSOM).

Cuadro 4.4: Resumen de resultados en coeficiente Qini y beneficio máximo en modelos causales

				d (Incentivo/descuento deuda. €)				
f(€)	Estrategia	Clasificador	Qini	$\bar{b}/5$	$\bar{b}/8$	$\bar{b}/10$	$\bar{b}/15$	$\bar{b}/20$
$f=1$	U. CI 10 %	T-Logit ProSOM	**5.4**	2.9	3.4	3.6	3.9	4.0
	U. CI 50 %	T-Logit ProSOM	**5.4**	4.4	7.1	8.0	9.2	9.8
	U. profit_s	S-LGBM PSM	3.6	**11.3**	16.0	17.5	19.6	20.7
	U. profit_m	S-ExTree NoRS	3.1	**11.3**	**16.6**	**18.4**	**20.7**	**21.9**
$f=2$	U. CI 10 %	T-Logit ProSOM	**5.4**	2.8	3.4	3.6	3.8	4.0
	U. CI 50 %	T-Logit ProSOM	**5.4**	4.3	7.0	7.9	9.1	9.7
	U. profit_s	S-LGBM NoRS	3.6	11.1	15.8	17.3	19.4	20.5
	U. profit_m	S-ExTree NoRS	3.1	**11.2**	**16.5**	**18.2**	**20.6**	**21.7**
$f=5$	CI 10 %	T.Logit ProSOM	**5.4**	2.7	3.3	3.5	3.7	3.9
	U. CI 50 %	T-Logit ProSOM	**5.4**	4.0	6.7	7.7	8.9	9.5
	U. profit_s	S-LGBM NoRS	3.6	10.5	15.2	16.8	18.9	19.9
	U. profit_m	S-ExTree NoRS	3.1	**10.7**	**16.0**	**17.7**	**20.1**	**21.2**

A diferencia de lo observado en la clasificación tradicional, en la Tabla 4.4 se evidencian diferencias importantes en relación con el coeficiente Qini, donde el enfoque causal insensible al costo muestra un mejor desempeño que los métodos sensibles al costo. Este comportamiento es esperado, ya que los enfoques orientados a maximizar beneficios no están diseñados para optimizar métricas independientes del umbral, como el coeficiente Qini. No obstante, la métrica de beneficio multisegmento (profit_m) obtiene mejores resultados en términos de rentabilidad en comparación con el método insensible al costo y con el enfoque de umbral único (profit_s). En particular, profit_m alcanza el mayor beneficio en los 15 escenarios experimentales considerados, para las distintas configuraciones de d y f. También, la rentabilidad disminuye a medida que aumentan

el descuento en la deuda d y el costo de contacto f.

Finalmente, el desempeño detallado en términos de beneficios ($\dot{MP}/\dot{MSP}$) para los distintos modelos predictivos (DTree, ExTree, LGBM y Logit) se reporta en las Figuras E.1, E.2 y E.3 del anexo E, correspondientes a las estrategias MOA, S-Learner y T-Learner, respectivamente. En cada una de estas figuras, el eje X representa los distintos valores de incentivo d (descuento en la deuda), mientras que los diferentes costos de contacto f se muestran en las subfiguras. Los métodos basados en beneficio con umbral único y las estrategias multiumbral se diferencian mediante líneas discontinuas y continuas, respectivamente.

Como se observa en las figuras, las estrategias multiumbral superan a los modelos de umbral único en todas las metodologías y configuraciones de hiperparámetros. De acuerdo con el análisis para la clasificación tradicional (Figura 4.5), solo se observan pequeñas diferencias en el rendimiento entre los distintos métodos de predicción base.

Los resultados experimentales presentados muestran que el enfoque propuesto basado en ganancias y múltiples umbrales supera tanto a la estrategia de umbral único y a los enfoques estadísticos clásicos, estos son la clasificación tradicional y los modelos causales. También, los resultados sugieren que la elección del marco de decisión orientado a ganancias tiene un impacto más determinante en el beneficio final que la selección del algoritmo de aprendizaje automático subyacente.

Impacto económico de la metodología propuesta

Los resultados obtenidos evidencian que la incorporación de modelos de aprendizaje causal basados en beneficio, junto con su extensión multisegmento, constituye una alternativa efectiva para optimizar las estrategias de cobranza. A diferencia de los enfoques tradicionales, centrados principalmente en estimar la probabilidad de pago de un cliente, la metodología propuesta permite identificar aquellos casos en los que una intervención específica, como una llamada acompañada de un incentivo, genera un cambio real en el comportamiento del deudor. Estimar el efecto causal de la intervención permite seleccionar de manera más precisa a los clientes sobre los cuales la gestión de cobranza generara un impacto mayor. Desde una perspectiva económica, este enfoque permite asignar los recursos de forma más eficiente, concentrando las acciones de cobranza únicamente en los clientes que presentan una respuesta favorable a la intervención y evitando destinar esfuerzos a casos donde la gestión no modifica el resultado

esperado. En consecuencia, las empresas pueden reducir costos operacionales asociados a llamadas, incentivos y recursos humanos, al mismo tiempo que incrementan la recuperación de deuda mediante campañas más focalizadas y personalizadas. Este aspecto resulta especialmente relevante en escenarios donde el volumen de clientes morosos es alto y existen pocos recursos. Asimismo, aplicar modelos basados en beneficio incorpora una ventaja adicional al integrar el impacto económico dentro del proceso de optimización. En lugar de maximizar únicamente métricas predictivas, como la exactitud o el AUC, estos modelos buscan maximizar el beneficio esperado para la empresa, alineando el proceso de aprendizaje con los objetivos del negocio. En conjunto, los resultados presentados respaldan el potencial de esta metodología como una herramienta para mejorar la eficiencia de las campañas de cobranza, proporcionando una base robusta para su evaluación en entornos empresariales reales, donde se podrá medir con mayor precisión el impacto sobre la recuperación de deuda, la reducción de costos y el beneficio económico generado por su implementación.

4.3. Influencia de los incentivos en la venta cruzada

En esta sección se presentan los resultados obtenidos al aplicar modelos causales enfocados en las ganancias en la base de datos de venta cruzada. Para esta aplicación se utilizaron 7 clasificadores, los 4 enfoques causales y 4 métricas de rendimiento: Qini con umbral óptimo insensible al costo (Qini_{ins}), Qini con umbral óptimo sensible al costo (Qini_{sens}), profit promedio con umbral óptimo insensible al costo (Profitins) y el Profit promedio con umbral óptimo sensible al costo (Profit$sens$). Para el cálculo de la métrica profit en los modelos, se considera b_{cs} =€239, C_{cs} =€10 y C_{camp} =€5.

Los resultados se presentan en la Tabla 4.5 donde se puede ver que las estrategias insensibles (ins) y sensibles ($sens$) al costo generan valores muy similares tanto en Qini como en profit, lo que significa que los enfoques para ordenar las muestras (ECP e ITEs) tienen rendimientos similares y que los umbrales que maximizan el uplift y el profit son muy cercanos. Sin embargo, la principal conclusión es que, cuando se dispone de una matriz costo-beneficio causal, la selección del modelo debe priorizar la maximización del beneficio.

Desde el punto de vista estadístico, el mejor desempeño se obtiene con XGBoost bajo

Cuadro 4.5: Resumen de los resultados obtenidos para la base de datos de venta cruzada

Enfoque Uplift	Modelo	Qini_{ins}	QINI_{sens}	Profit_{ins} (€)	Profit_{sens} (€)
CF	DT	0.057	0.048	5.416	4.605
	Extra-T	0.142	0.137	6.247	5.679
	GBoost	0.062	0.056	-2.467	-3.033
	LGBM	0.008	0.002	-0.089	-0.569
	Logit	0.149	0.145	11.252	10.826
	RF	0.142	0.138	3.087	2.602
	XBGoost	0.033	0.028	-2.449	-2.995
MOA	DT	0.253	0.253	15.364	15.364
	Extra-T	0.285	0.285	17.963	17.963
	GBoost	0.220	0.220	6.075	6.075
	LGBM	0.293	0.293	15.065	15.065
	Logit	-0.032	-0.032	-1.771	-1.771
	RF	0.325	0.325	19.816	19.816
	XBGoost	**0.337**	**0.337**	20.134	20.134
S-learner	DT	-0.092	0.065	4.195	7.779
	Extra-T	0.254	0.249	23.342	22.917
	GBoost	0.233	0.220	11.360	10.760
	LGBM	0.199	0.183	15.068	14.633
	Logit	0.091	0.114	-8.491	-0.273
	RF	0.260	0.254	22.588	22.196
	XBGoost	0.152	0.143	16.498	15.883
T-learner	DT	0.144	0.142	16.106	15.449
	Extra-T	0.250	0.246	22.414	22.030
	GBoost	0.237	0.235	**26.497**	**26.394**
	LGBM	0.217	0.215	24.164	23.997
	Logit	0.158	0.156	13.105	13.070
	RF	0.250	0.245	22.081	21.623
	XBGoost	0.192	0.190	21.251	20.947

la estrategia MOA (Qini_{ins}=Qini_{sens}=0.337). Sin embargo, al considerar métricas de negocio, el enfoque T-learner con gradient boosting incrementa el profit promedio de 20.134 a 26.497, lo que representa una mejora del 31.6 %, a pesar de una disminución en el valor de Qini. Este resultado evidencia la relevancia de emplear métricas basadas en beneficio en contextos de analítica empresarial para mejorar la toma de decisiones. Finalmente, la presencia de valores negativos de Qini indica que el modelo tiene un desempeño inferior a hacerlo de manera aleatoria, lo que ocurre cuando existe ruido en el problema o sobreajuste. Por otro lado, se presentan beneficios negativos cuando los costos superan los retornos esperados, esto se puede presentar cuando los modelos se comportan de forma aleatoria o cuando hay una estrecha relación entre el costo y

beneficio, que no sería este caso dado que los modelos que tienen un valor de Qini alto tienen una buena rentabilidad. En particular, el enfoque MOA-Logit presenta un rendimiento cercano al azar y genera rentabilidad negativa dado que todos sus resultados son negativos.

Luego, se analizan las curvas Qini (Figura 4.6) y profit (Figura 4.7) para la mejor configuración del clasificador en términos de beneficio (enfoque T-learner con modelo gradient boosting). Ambas figuras ilustran las soluciones insensibles y sensibles al costo (líneas verde y azul, respectivamente), y el intervalo de confianza obtenido durante el proceso de validación cruzada.

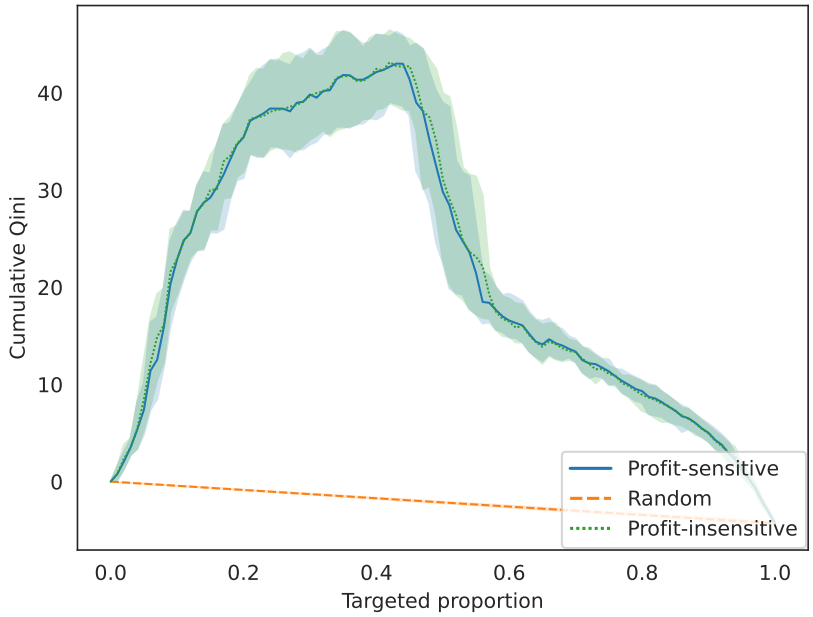


Figura 4.6: Curva Qini

A partir de las figuras, se observa que los enfoques sensibles e insensibles se superponen. Esto podría explicarse por la estructura de costos y beneficios definida para la tarea de venta cruzada enfocada en empleados. En ambos casos, las curvas alcanzan su máximo cuando se dirige aproximadamente al 40 % de la población. La forma de las curvas indica que es posible obtener mejoras importantes en ambas métricas en comparación con una asignación aleatoria, lo que sugiere que la estrategia actual de la institución financiera no es óptima y podría optimizarse de manera significativa.

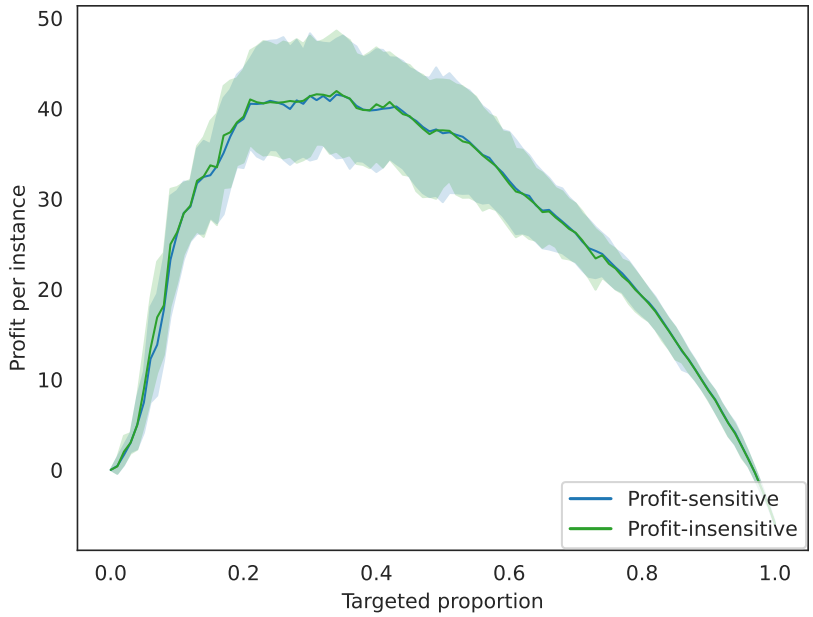


Figura 4.7: Curva profit

Por otro lado, se analiza la relevancia de las variables a través del método TreeSHAP. La Figura 4.8 presenta los resultados, siendo los atributos en la parte superior del gráfico los más relevantes. A partir de esto podemos concluir que las variables de *network* (*Tie trength* y *indegree*) son las mas relevantes, que hacen referencia a la intensidad de la relación entre el receptor y el emisor. Esto destaca la importancia de fomentar relaciones entre vendedores de diferentes unidades, así como de apoyar y preparar a los agentes de ventas que reciben un gran número de referencias.

Finalmente, se analiza el efecto de variar los costos. Suponiendo que b_{cs} =€239 es fijo, se exploran los siguientes valores para la campaña y las recompensas: $C_{cs} \in \{5, 10, 25\}$ y $C_{camp} \in \{0, 1, 5, 10\}$. Los resultados para la métrica Qini y Profit se presentan en la Tabla 4.6 y 4.7, también se calcula el porcentaje promedio de muestras objetivo en el conjunto de prueba (*Targeted*) y el porcentaje promedio de muestras que tienen el mismo resultado (objetivo o no) que el esquema de recompensa actual (*Shared*).

La solución que mejor refleja el esquema de recompensa actual (C_{cs} =€10 y C_{camp} =€5) aparece subrayada en la Tabla 4.7, y la mejor en términos de Qini_{sens} y Profit_{sens} (C_{cs} =€1 y C_{camp} =€0) está resaltada en negrita en la Tabla 4.6.

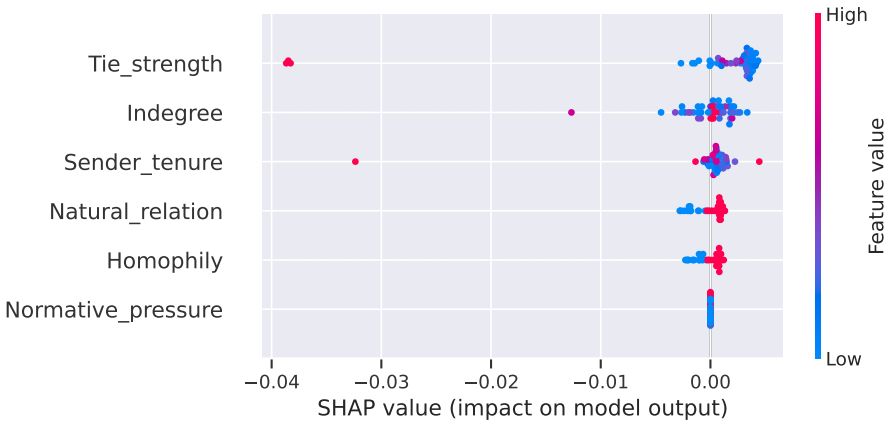


Figura 4.8: Relevancia de las variables para la base de datos de venta cruzada.

Cuadro 4.6: Resultados de la variación de los costos de campaña y retención de clientes ($C_{cs} \in \{5, 10, 25\}$ y $C_{camp} \in \{0, 1\}$).

Métrica $\setminus$ $C_{camp} - C_{cs}$	0-5	0-10	0-25	1-5	1-10	1-25
Qini_{sens}	**0.237**	0.236	0.235	0.233	0.237	0.236
Profit_{sens}(€)	**28.53**	28.22	26.99	25.51	27.92	27.61
Targeted	**41.31 %**	41.31 %	41.31 %	41.20 %	41.20 %	41.19 %
Shared	**99.28 %**	99.28 %	99.28 %	99.38 %	99.38 %	99.39 %

Cuadro 4.7: Resultados de la variación de los costos de campaña y retención de clientes ($C_{cs} \in \{5, 10, 25\}$ y $C_{camp} \in \{5, 10\}$).

Métrica $\setminus$ $C_{camp} - C_{cs}$	5-5	5-10	5-25	10-5	10-10	10-25
Qini_{sens}	0.235	0.233	0.237	0.236	0.234	0.233
Profit_{sens}(€)	26.39	24.90	26.09	25.79	24.58	23.07
Targeted	40.60 %	40.59 %	40.56 %	39.68 %	39.66 %	39.41 %
Shared	99.99 %	-	99.98 %	99.09 %	99.07 %	98.83 %

Las Tablas 4.6 y 4.7 muestran que las métricas presentan cambios marginales en comparación con la solución que representa el esquema de recompensas vigente, lo que indica que los modelos mantienen un comportamiento estable y consistente frente a variaciones en los costos utilizados para determinar el umbral óptimo.

En cuanto a la rentabilidad promedio ($Profit_{sens}$), esta tiende a disminuir conforme aumentan los incentivos, un resultado esperado dado que el enfoque sensible al costo no considera un efecto positivo de campaña. A pesar de ello, los niveles de beneficio siguen siendo elevados incluso en escenarios con incentivos altos, y los resultados se

mantienen estables ante distintas configuraciones de C_{cs} y C_{camp}. Este tipo de análisis resulta útil para la definición de costos de campaña y evidencia patrones similares tanto en la proporción de muestras seleccionadas como en aquellas alineadas con el esquema de recompensas actual.

4.4. Efecto del tratamiento individual acotado

En esta sección se presentan los resultados de la aplicación del nuevo enfoque para analizar modelos causales propuestos en el capítulo 3, efecto del tratamiento individual acotado (BITE). Para medir su rendimiento, se compara con los modelos de clasificación tradicional y el enfoque de efecto de tratamiento individual (ITE) en modelos causales. Además se mide el rendimiento de estos 3 enfoques en 7 bases de datos en términos del uplift al 5 %, 10 %, 15 %, 20 % , 25 % y 30 %. Estos resultados se presentan en las Tablas 4.8 y 4.9.

Cuadro 4.8: Resumen de resultados de las métricas uplift al 5 %, 10 % y 15 % para los 3 enfoques de clasificación

	Uplift al 5 %			Uplift al 10 %			Uplift al 15 %		
Dataset	PRED	ITE	BITE	PRED	ITE	BITE	PRED	ITE	BITE
Crosssell	0.068	**0.723**	**0.723**	0.074	**0.551**	**0.551**	0.107	**0.491**	**0.491**
HillAll	**0.179**	0.165	0.166	0.129	**0.131**	**0.131**	**0.122**	0.121	0.121
HillMen	0.171	**0.221**	**0.221**	0.175	**0.178**	**0.178**	0.123	0.124	**0.139**
HillWomen	0.104	**0.155**	**0.155**	0.076	**0.118**	**0.118**	0.085	0.093	**0.102**
Informs	**0.045**	0.043	0.0423	**0.039**	0.036	0.036	**0.032**	0.024	0.024
Insurance	-0.018	**0.226**	**0.226**	-0.087	**0.164**	**0.164**	-0.073	0.106	**0.115**
Udemy	**0.056**	0.023	0.032	**0.034**	0.028	0.028	**0.037**	0.0274	0.0274

A partir de los resultados, se puede concluir que el mejor rendimiento para BITE se obtiene con uplift al 20 %, donde el uplfit más alto es obtenido en seis de las siete bases de datos. También, se puede ver que BITE del modelo causal tiene mejor rendimiento en general, alcanzando el mejor uplift at k % el 66.7 % de las veces compartido con ITE y 33.3 % de las veces en general.

En la Tabla 4.8 se puede ver que ITE y BITE se comportan de manera similar, obteniendo mejores resultados que los modelos de clasificación tradicional, pero en la Tabla 4.9 se observan diferencias entre ITE y BITE esto debido a que la mayoría de las muestras

Cuadro 4.9: Resumen de resultados de las métricas uplift al 20 %, 25 % y 30 % para los 3 enfoques de clasificación.

	Uplift al 20 %			Uplift al 25 %			Uplift al 30 %		
Dataset	PRED	ITE	BITE	PRED	ITE	BITE	PRED	ITE	BITE
Crosssell	0.121	0.474	**0.487**	0.134	0.421	**0.431**	0.131	**0.426**	0.359
HillAll	0.112	0.113	**0.120**	0.106	**0.111**	**0.111**	0.098	**0.107**	**0.107**
HillMen	0.119	0.109	**0.128**	0.108	0.108	**0.124**	0.101	0.101	**0.119**
HillWomen	0.072	0.088	**0.099**	0.065	**0.099**	**0.099**	0.065	**0.087**	**0.087**
Informs	**0.024**	0.018	0.015	0.018	**0.014**	0.013	0.016	**0.013**	0.011
Insurance	-0.038	0.063	**0.079**	-0.002	0.062	**0.075**	-0.002	0.061	**0.062**
Udemy	0.026	0.027	**0.028**	**0.026**	0.022	0.022	0.019	**0.025**	0.021

con un aumento significativo no son ni casos perdidos ni casos seguros. Por lo tanto, las medidas ITE y BITE arrojan resultados similares cuando se selecciona un pequeño porcentaje de individuos (5 % o 10 %). Sin embargo, algunos individuos que se encuentran dentro de los límites presentan valores ITE positivos, esto sugiere que el método BITE filtra varios ejemplos, y los deciles superiores definidos por ITE y BITE podrían diferir porque este último solo considera los que no son ni casos perdidos ni seguros como se puede ver en la Figura 4.9. Las causas perdidas ($P(Y = 1|x, W = 1) < \beta_t$) y las causas seguras ($P(Y = 1|x, W = 0) > \beta_c$) definidas por el marco BITE se resaltan en azul claro y azul, respectivamente.

Para analizar los resultados de las Tablas 4.8 y 4.9, se clasifican las estrategias según su desempeño en términos de uplift en distintos valores de k %. Luego, se promedian estos rankings para obtener una medida global de rendimiento, junto con el uplift promedio. Estos análisis se presentan en el Tabla 4.10

Cuadro 4.10: Resultados promedio y prueba de Holm para comparaciones por pares.

Método	Rank	Uplift	p-value	$\frac{\alpha_Z}{(k-1)}$	Resultado	W/T/L
Bounded ITE	1.6429	0.1521	-	-	-	-
Estandar ITE	1.9762	0.1499	0.1266	0.05	not reject	5/20/17
Predictivo	2.3810	0.0667	0.0007	0.025	reject	12/0/31

Se aplica el test de Friedman (con corrección de Iman-Davenport) el cual muestra diferencias estadísticamente significativas entre los métodos ($p < 0.001$), rechazando la hipótesis de igualdad de desempeño. Por otro lado, se aplica el test de Holm, mediante comparaciones pareadas, que indica que el enfoque Bounded ITE supera significativa-

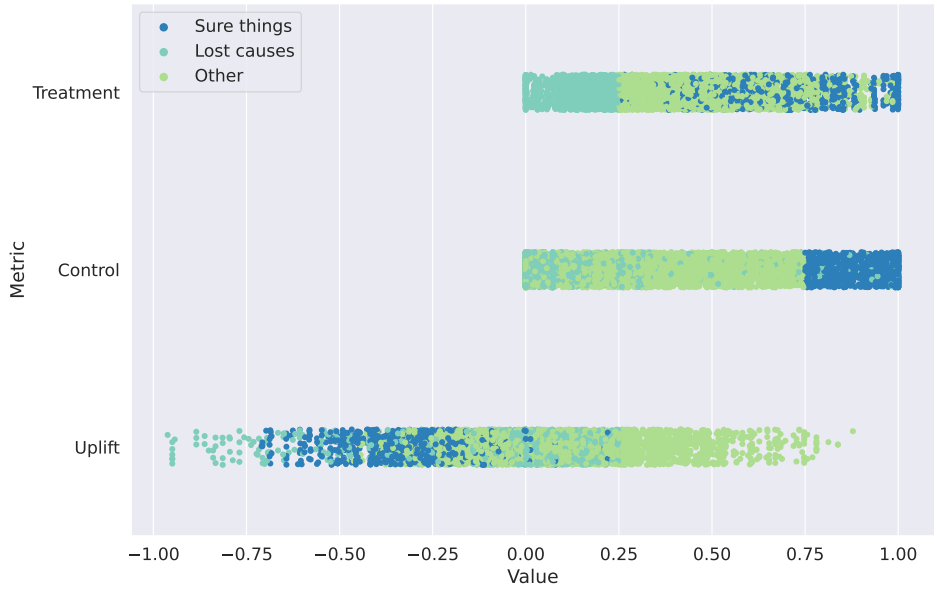


Figura 4.9: Distribución de datos en términos de probabilidades de tratamiento, control y uplift (ITE).

mente al método predictivo tradicional, mientras que no hay diferencias significativas frente al enfoque ITE estándar.

En términos de resultados, Bounded ITE obtiene el mejor ranking global y el mayor uplift promedio, lo que confirma su superioridad en este contexto.

Capítulo 5

Conclusiones

5.1. Conclusiones teóricas y prácticas

Esta tesis tuvo como objetivo principal reunir una serie de avances metodológicos y aplicaciones novedosas en modelos causales y aprendizaje automático, abordando el problema desde un enfoque integrado que combina modelos predictivos, modelos causales y criterios de optimización basados en beneficios, en los contextos de cobranza y venta cruzada.

En primer lugar, se optimizó la gestión de cobranza mediante la integración de variables de *contact centers* e información financiera, introduciendo además dos nuevas tareas predictivas: el contacto exitoso y la promesa de pago. A partir de los resultados obtenidos mediante la aplicación de modelos de aprendizaje automático, evaluados a través del área bajo la curva (AUC), se demuestra que las variables provenientes del *contact center* poseen, en promedio, una mayor capacidad predictiva que la información financiera disponible. Sin embargo, la combinación de ambas fuentes de datos permitió maximizar el rendimiento en todas las tareas analizadas, destacando la predicción de pago de deuda como la de mayor precisión. Estos resultados establecen una línea base relevante para evaluar la contribución de enfoques más avanzados.

A continuación, la incorporación de modelos causales en cobranza permitió validar su eficacia para estimar el efecto del tratamiento a nivel individual, logrando identificar con precisión a los deudores cuya probabilidad de pago aumenta específicamente ante una gestión de cobro. En comparación con los modelos tradicionales, se observó que,

mientras los enfoques predictivos convencionales se centran principalmente en estimar el riesgo de impago, los modelos causales permiten optimizar la eficiencia operativa al evitar contactos redundantes.

De esta forma, la integración de un enfoque basado en beneficios demostró que el uso del Beneficio Causal Máximo ($\dot{MP}$) resulta superior a las métricas estadísticas tradicionales para la toma de decisiones financieras, permitiendo identificar a los clientes sobre los cuales la gestión de cobro genera una influencia efectiva en su decisión de pago y evitando costos innecesarios asociados a contactos redundantes. Asimismo, la estrategia multi-segmento (MSP) resultó clave en este contexto, ya que permite maximizar la rentabilidad económica mediante el ajuste de las políticas de cobranza según la heterogeneidad de los montos adeudados y los costos de contacto. En lugar de aplicar una política uniforme, este enfoque permite implementar incentivos más agresivos en segmentos de alto valor, donde la recuperación de la deuda justifica mayores descuentos, manteniendo al mismo tiempo márgenes más altos en segmentos de menor valor. Los resultados experimentales muestran que el enfoque multiumbral basado en rentabilidad supera tanto al enfoque de umbral único como a los enfoques estadísticos clásicos, tanto para clasificación tradicional como para modelos causales. Además, los resultados sugieren que la elección del marco orientado a la rentabilidad tiene una mayor influencia sobre el desempeño económico que la elección específica del algoritmo de aprendizaje automático subyacente.

Posteriormente, la aplicación de modelos de aprendizaje causal basados en beneficio en el contexto de venta cruzada mostró que estos enfoques superan la estrategia actualmente utilizada por la empresa, optimizando el esfuerzo comercial y alcanzando un valor Qini de 0.337. Asimismo, la aplicación de clasificadores orientados al beneficio, en lugar de métricas estadísticas tradicionales, incrementó la rentabilidad promedio por muestra en un 31.6 %, validando la importancia de integrar estructuras de costo-beneficio en el diseño de modelos de decisión orientados a maximizar el impacto económico de las campañas. Estos resultados confirman la capacidad del enfoque propuesto para extenderse a distintos problemas de negocio manteniendo coherencia metodológica.

Otro aporte relevante de esta tesis corresponde a la propuesta de una nueva metodología de evaluación para modelos causales, denominada Bounded ITE (BITE), la cual mejora la precisión en la identificación de clientes "persuasibles" mediante la incorporación de

límites basados en las probabilidades de tratamiento y control. Este enfoque otorga mayor relevancia a las probabilidades asociadas al grupo de control, con el objetivo de mitigar el ruido y la complejidad inherentes a la predicción del comportamiento humano. Los resultados demuestran que el método BITE supera tanto a los modelos causales estándar como a la clasificación convencional en métricas de Uplift at k %, confirmando que la exclusión de casos extremos (clientes seguros o perdidos) permite optimizar la captura del efecto causal.

Finalmente, los resultados de esta tesis demuestran que la integración de modelos predictivos, aprendizaje causal y enfoques orientados al beneficio constituye un marco robusto para apoyar la toma de decisiones empresariales. En conjunto, las contribuciones metodológicas y aplicadas desarrolladas permiten avanzar hacia estrategias más precisas, eficientes y rentables en contextos de negocio complejos, aportando además nuevas líneas de investigación para el desarrollo de sistemas de decisión basados en inteligencia artificial y causalidad.

5.2. Investigación futura

Este trabajo resuelve varias interrogantes, pero también nos lleva a nuevas preguntas, que requieren estudios posteriores para poder responderse. En esa línea, esta sección presenta una serie de propuestas de investigación futura, que podrían complementar, completar, o bien enriquecer esta investigación.

Con respecto a la aplicación de modelos causales basado en beneficio para cobranzas, una limitación del marco propuesto es que solo mitiga parcialmente el supuesto de probabilidades de aceptación uniformes para las cancelaciones de deuda al dividir a los deudores en segmentos. Un enfoque más sofisticado implicaría el diseño de una función explícita para modelar esta relación, en línea con Latorre, Meza, Lopez-Ospina, Verbeke & Pérez (2025) para la predicción de la deserción de clientes. Sin embargo, dicho análisis requeriría datos detallados de experimentos aleatorios con diferentes niveles de condonación de la deuda y sus correspondientes tasas de aceptación. Si bien este estudio utiliza un conjunto de datos observacionales a gran escala, una validación futura mediante un ensayo controlado aleatorio (ECA) consolidaría aún más las estimaciones de elasticidad causal.

Otro enfoque que se podría explorar es la aplicación de estos modelos causales orientados a las ganancias a datos aún más detallados, como las transcripciones de voz a texto de las llamadas de cobranza, para refinar aún más las características de comportamiento utilizadas en la estimación de uplift. Además, el marco podría extenderse a un entorno de optimización multiperíodo donde el incentivo para el pago de la deuda se ajusta dinámicamente a lo largo del ciclo de cobranza para tener en cuenta el valor temporal del dinero y los rendimientos decrecientes del contacto repetido. En definitiva, esta investigación transforma la gestión de cobranzas, pasando de ser una carga administrativa reactiva a una estrategia proactiva basada en la evidencia, que maximiza la resiliencia institucional y la eficiencia en la recuperación.

Otro tema interesante a investigar y que actualmente se encuentra en desarrollo, pero quedo fuera del alcance de esta tesis, es la extensión del enfoque Bounded ITE (BITE). Este enfoque aborda el problema de que el efecto de tratamiento individual (ITE) no logra distinguir entre clientes que son casos seguros o perdidos de los persuasibles ya que tienen igual uplift a partir de sus ITE. Para esto, BITE aplica umbrales de probabilidades estrictos para filtrar las causas perdidas y los casos seguros. Sin embargo, la dependencia de BITE de las restricciones binarias introduce discontinuidades en el espacio de decisión y no logra capturar la naturaleza continua del comportamiento del cliente y el riesgo empresarial.

Una posible extensión consiste en reemplazar los límites rígidos de BITE por mecanismos de ponderación continua derivados de operadores *Ordered Weighted Averaging* (OWA) y cuantificadores *Regular Increasing Monotone* (RIM). Basándose en los fundamentos teóricos de medidas de consenso en toma de decisiones grupales y en el marco estratégico de BITE, este enfoque daría origen al denominado Efecto de Tratamiento Individualizado Difuso (FITE). Bajo esta formulación, la identificación de clientes objetivo dejaría de abordarse como un problema de pertenencia a conjuntos nítidos, pasando a modelarse como un problema de pertenencia difusa. Mediante el uso de cuantificadores RIM para transformar probabilidades de referencia en ponderaciones continuas, FITE permitiría asignar relevancia parcial a candidatos subóptimos de manera más flexible y matizada. Además, este enfoque preservaría la estructura de orden del estimador causal, incorporando simultáneamente la actitud frente al riesgo del tomador de decisiones, optimismo o pesimismo, mediante el concepto de *Orness*.

Nomenclatura

GCD: greatest common divisor

AI: Inteligencia Artificial (Artificial Intelligence)

ML: Aprendizaje Automático (Machine Learning)

DL: Aprendizaje Profundo (Deep Learning)

XAI: Inteligencia Artificial Explicable (Explainable Artificial Intelligence)

CRM: Gestión de Relaciones con Clientes (Customer Relationship Management)

ROC: Curva Característica Operativa del Receptor (Receiver Operating Characteristic)

AUC: Área Bajo la Curva (Area Under the Curve)

AUUC: Área Bajo la Curva Uplift (Area Under the Uplift Curve)

Qini: Coeficiente Qini

ITE: Efecto de Tratamiento Individual (Individual Treatment Effect)

CATE: Efecto Promedio Condicional del Tratamiento (Conditional Average Treatment Effect)

ATE: Efecto Promedio del Tratamiento (Average Treatment Effect)

BITE: Bounded Individual Treatment Effect

FITE: Fuzzy Individualized Treatment Effect

MSP: Multi-Segment Policy

OWA: Ordered Weighted Averaging

RIM: Regular Increasing Monotone

ECA: Ensayo Controlado Aleatorio

SVM: Máquina de Vectores de Soporte (Support Vector Machine)

RF: Random Forest

XGBoost: Extreme Gradient Boosting

LightGBM: Light Gradient Boosting Machine

CatBoost: Categorical Boosting

ANN: Red Neuronal Artificial (Artificial Neural Network)

SHAP: SHapley Additive exPlanations

KPI: Indicador Clave de Desempeño (Key Performance Indicator)

Referencias

Abe, N., Melville, P., Pendus, C., Reddy, C., Jensen, D., Thomas, V., Bennett, J., Anderson, G., Cooley, B., Kowalczyk, M., Domick, M., & Gardinier, T. (2010). Optimizing debt collections using constrained reinforcement learning. In *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, (pp. 75–84). Association for Computing Machinery.

Ascarza, E. (2017). Retention futility: Targeting high risk customers might be ineffective. *Journal of Marketing Research*, *55*.

Athey, S. & Imbens, G. (2015). Machine learning for estimating heterogeneous causal effects. Research papers, Stanford University, Graduate School of Business.

Athey S, W. S. (2019). Estimating treatment effects with causal forests: an application. *Observational studies*.

Baesens, B. (2014). *Analytics in a Big Data World: The Essential Guide to Data Science and its Applications* (1st ed.). Wiley Publishing.

Baesens, B. & de Caigny, A. (2022). Customer Lifetime Value Modeling with Applications in Python and R: Lessons and Experiences from Industry and Research on how to Become a Customer-Centric. Post-Print hal-03982860, HAL.

Baesens, B., Roesch, D., & Scheule, H. (2016). *Credit Risk Analytics: Measurement Techniques, Applications, and Examples in SAS*. John Wiley & Sons.

Bahrami, M., Bozkaya, B., & Balcisoy, S. (2020). Using behavioral analytics to predict customer invoice payment. *Big data*, *8*(1), 25–37.

Baier, D. & Stöcker, B. (2022). Profit uplift modeling for direct marketing campaigns: approaches and applications for online shops. *Journal of Business Economics*, *92*(4), 645–673.

Bellotti, A., Brigo, D., Gambetti, P., & Vrins, F. (2021). Forecasting recovery rates on non-performing loans with machine learning. *International Journal of Forecasting*, *37*(1), 428–444.

Chehrazi, N., Glynn, P. W., & Weber, T. A. (2019). Dynamic credit-collections optimization. *Management Science*, *65*(6), 2737–2769.

Chung, D. J., Steenburgh, T., & Sudhir, K. (2014). Do bonuses enhance sales productivity? a dynamic structural analysis of bonus-based compensation plans. *Marketing Science*, *33*(2), 165–187.

Coleman, J. & Hoffa, T. (1987). *Public and Private High Schools: The Impact of Communities*. Basic Books.

Crespo, I. & Govindarajan, A. (2018). The analytics-enabled collections model. *McKinsey & Company, April*.

Daljord, Ø., Misra, S., & Nair, H. S. (2016). Homogeneous contracts for heterogeneous agents: aligning sales force composition and compensation. *Journal of Marketing Research*, *53*(2), 161–182.

Devriendt, F., Berrevoets, J., & Verbeke, W. (2021a). Why you should stop predicting customer churn and start using uplift models. *Information Sciences*, *548*, 497–515.

Devriendt, F., Berrevoets, J., & Verbeke, W. (2021b). Why you should stop predicting customer churn and start using uplift models. *Information Sciences*, *548*, 497–515.

Devriendt, F., Moldovan, D., & Verbeke, W. (2018a). A literature survey and experimental evaluation of the state-of-the-art in uplift modeling: A stepping stone toward the development of prescriptive analytics. *Big Data*, *6*(1), 13–41. PMID: 29570415.

Devriendt, F., Moldovan, D., & Verbeke, W. (2018b). A literature survey and experimental evaluation of the state-of-the-art in uplift modeling: A stepping stone toward the development of prescriptive analytics. *Big data*, *6*(1), 13–41.

Fan, Q., Hsu, Y.-C., Lieli, R. P., & Zhang, Y. (2021). Estimation of conditional average treatment effects with high-dimensional data.

Fayyad, U. (1996). Data mining and knowledge discovery: Making sense out of data. *IEEE expert*, *11*(5), 20–25.

Friberg, R. & Sanctuary, M. (2017). The effect of retail distribution on sales of alcoholic beverages. *Marketing Science*, *36*(4), 626–641.

Ghoshal, A., Mookerjee, V. S., & Sarkar, S. (2021). Recommendations and cross-selling: Pricing strategies when personalizing firms cross-sell. *Journal of Management Information Systems*, *38*(2), 430–456.

Glymour, C., Zhang, K., & Spirtes, P. (2019). Review of causal discovery methods based on graphical models. *Frontiers in Genetics*, *Volume 10 - 2019*.

Gubela, R. M., Lessmann, S., & Jaroszewicz, S. (2020). Response transformation and profit decomposition for revenue uplift modeling. *European Journal of Operational Research*, *283*(2), 647–661.

Gupta, S. & Zeithaml, V. (2006). Customer metrics and their impact on financial performance. *Marketing Science*, *25*(6), 718–739.

Han, J., Kamber, M., & Pei, J. (2012). *Data Mining, Concepts and Techniques* (Third ed.). In The Morgan Kaufmann Series in Data Management Systems.

Hansotia, B.; Rukstales, B. (2002). Direct marketing for multichannel retailers: Issues, challenges and solutions. *J Database Mark Cust Strategy Manag*, *9*, 259–266.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media.

Hinz, O., Skiera, B., Barrot, C., & Becker, J. U. (2011). Seeding strategies for viral marketing: An empirical comparison. *Journal of Marketing*, *75*(6), 55–71.

Höppner, S., Stripling, E., Baesens, B., vanden Broucke, S., & Verdonck, T. (2017). Profit driven decision trees for churn prediction. *European Journal of Operational Research*, *284*.

Jiang, P., Liu, Z., Abedin, M. Z., Wang, J., Yang, W., & Dong, Q. (2024). Profit-driven weighted classifier with interpretable ability for customer churn prediction. *Omega*, *125*, 103034.

Kalkan, I. E. & Şahin, C. (2023). Evaluating cross-selling opportunities with recurrent neural networks on retail marketing. *Neural Computing and Applications*, *35*(8), 6247–6263.

Kim, J. & Kang, P. (2016). Late payment prediction models for fair allocation of customer contact lists to call center agents. *Decision Support Systems*, *85*, 84–101.

Künzel, S. R., Sekhon, J. S., Bickel, P. J., & Yu, B. (2019a). Metalearners for estimating heterogeneous treatment effects using machine learning. *Proceedings of the National Academy of Sciences*, *116*(10), 4156–4165.

Künzel, S. R., Sekhon, J. S., Bickel, P. J., & Yu, B. (2019b). Metalearners for estimating heterogeneous treatment effects using machine learning. *Proceedings of the National Academy of Sciences*, *116*(10), 4156–4165.

Lalonde, R. (1986). Evaluating the econometric evaluations of training programs with experiment data. *American Economic Review*, *76*, 604–20.

Latorre, P., Meza, A., Lopez-Ospina, H., Verbeke, W., & Pérez, J. (2025). A prescriptive analytics framework for jointly optimizing retention incentives and targeting. *Knowledge-Based Systems*, 114649.

Lemmens, A. & Gupta, S. (2020). Managing churn to maximize profits. *Marketing Science*, *39*(5), 956–973.

Lundberg, S., Erion, G., Chen, H., DeGrave, A., Prutkin, J., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S.-I. (2020). From local explanations to global understanding with explainable ai for trees. *Nature Machine Intelligence*, *2*.

Maldonado, S., Bravo, C., López, J., & Pérez, J. (2017). Integrated framework for profit-based feature selection and svm classification in credit scoring. *Decision Support Systems*, *104*, 113–121.

Maldonado, S., Domínguez, G., Olaya, D., & Verbeke, W. (2021). Profit-driven churn prediction for the mutual fund industry: A multisegment approach. *Omega*, *100*, 102380.

Maldonado, S., López, J., & Vairetti, C. (2020). Profit-based churn prediction based on minimax probability machines. *European Journal of Operational Research*, *284*(1), 273–284.

Maldonado, S., Álvaro Flores, Verbraken, T., Baesens, B., & Weber, R. (2015). Profit-based feature selection using support vector machines – general framework and an application for customer retention. *Applied Soft Computing*, *35*, 740–748.

Malinsky, D. & Danks, D. (2018). Causal discovery algorithms: A practical guide. *Philosophy Compass*, *13*(1), e12470. e12470 10.1111/phc3.12470.

Mantrala, M. K., Albers, S., Caldieraro, F., Jensen, O., Joseph, K., Krafft, M., Narasimhan, C., Gopalakrishna, S., Zoltners, A., Lal, R., et al. (2010). Sales force modeling: State of the field and research agenda. *Marketing Letters*, *21*(3), 255–272.

Misra, S. & Nair, H. S. (2011). A structural model of sales-force compensation dynamics: Estimation and field implementation. *Quantitative Marketing and Economics*, *9*(3), 211–257.

Mitchner, M. & Peterson, R. P. (1957). An operations-research study of the collection of defaulted loans. *Operations Research*, *5(4)*, 522–545.

Molnar, C. (2020). Interpretable machine learning: A guide for making black box models explainable. *URL https://christophm.github.io/interpretable-ml-book*.

Morgan, S. L. & Winship, C. (2014). *Counterfactuals and Causal Inference: Methods and Principles for Social Research* (2 ed.). Analytical Methods for Social Research. Cambridge University Press.

Ni, J., Neslin, S. A., & Sun, B. (2012). Database submission-the isms durable goods data sets. *Marketing Science*, *31*(6), 1008–1013.

Nyberg, O. & Klami, A. (2023). Exploring uplift modeling with high class imbalance. *Data Mining and Knowledge Discovery*, *37*(2), 736–766.

Olaya, D., Vásquez, J., Maldonado, S., Miranda, J., & Verbeke, W. (2020). Uplift modeling for preventing student dropout in higher education. *Decision Support Systems*, *134*, 113320.

Pearl, J. (2009). *Causality* (2 ed.). Cambridge University Press.

Peters, J., Janzing, D., & Scholkopf, B. (2017). *2017). Elements of causal inference: foundations and learning algorithms*. MIT press.

Petrides, G., Moldovan, D., Voets, L., Guns, T., & Verbeke, W. (2022). Cost-sensitive learning for profit-driven credit scoring. *Journal of the Operational Research Society*, *73*(2), 338–350.

Przybyłek, M., Przybyłek, A., Wojciechowski, K., & Shkroba, I. (2025). Towards a smart debt collection system: a design science research approach. *Journal of Big Data*, *12*(1), 193.

Radcliffe, N. J. & Surry, P. D. (2012). Real-world uplift modelling with significance-based uplift trees. In *Business, Computer Science*.

Reyna, J. (1968). Peter blau y otis dudley duncan. the american occupational structure. nueva york : John wiley and sons, 1967. 520 p. *Estudios Demográficos y Urbanos*, *2*, 117.

Seiler, S., Yao, S., & Wang, W. (2017). Does online word of mouth increase demand? (and how?) evidence from a natural experiment. *Marketing Science*, *forthcoming*.

Simsek, S., Dag, A., Tiahrt, T., & Oztekin, A. (2020). A bayesian belief network-based probabilistic mechanism to determine patient no-show risk categories. *Omega*, *100*, 102296.

Sivamayilvelan, K., Rajasekar, E., Vairavasundaram, S., Balachandran, S., & Suresh, V. (2024). Flexible recommendation for optimizing the debt collection process based on customer risk using deep reinforcement learning. *Expert Systems with Applications*, *256*, 124951.

Sivamayilvelan, K., Rajasekar, E., Vairavasundaram, S., Balachandran, S., & Suresh, V. (2025). Building explainable artificial intelligence for reinforcement learning based

debt collection recommender system using large language models. *Engineering Applications of Artificial Intelligence*, *159*, 111622.

Thomas, L., Crook, J., & Edelman, D. (2017). *Credit Scoring and its Applications* (Second ed.). Society for Industrial and Applied Mathematics.

Tran, C. M. & Tran, T. H. (2025). Optimizing recovery debt collection process by using machine learning and operation research. In *Asian Conference on Intelligent Information and Database Systems*, (pp. 199–215). Springer.

Vafeiadis, T., Diamantaras, K., Sarigiannidis, G., & Chatzisavvas, K. (2015). A comparison of machine learning techniques for customer churn prediction. *Simulation Modelling Practice and Theory*, *55*.

Vairetti, C., Gennaro, F., & Maldonado, S. (2024). Propensity score oversampling and matching for uplift modeling. *European Journal of Operational Research*, *316*(3), 1058–1069.

van den Bulte, C. & Wuyts, S. (2007). *Social Networks in Marketing*. MSI Relevant Knowledge Series. Cambridge MA: Marketing Science Institute.

van der Geer, R., Wang, Q., & Bhulai, S. (2018). Data-driven consumer debt collection via machine learning and approximate dynamic programming. *SSRN Electronic Journal*, 1–32.

Verbeke, W., Olaya, D., Berrevoets, J., Verboven, S., & Maldonado, S. (2020). The foundations of cost-sensitive causal classification. *arXiv preprint arXiv:2007.12582*.

Verbeke, W., Olaya, D., Guerry, M.-A., & Van Belle, J. (2023). To do or not to do? cost-sensitive causal classification with individual treatment effect estimates. *European Journal of Operational Research*, *305*(2), 838–852.

Vu, V.-H. (2024). Predict customer churn using combination deep learning networks model. *Neural Computing and Applications*, *36*(9), 4867–4883.

Wager, S. & Athey, S. (2017). Estimation and inference of heterogeneous treatment effects using random forests.

Weeks, M. F., Kulka, R. A., & Pierson, S. A. (1987). Optimal call scheduling for a telephone survey. *Public Opinion Quarterly*, *51*(4), 540–549.

Wolff, J., Erichsen, P., Kevrekidis, I., & Rotermund, H. (2000). Die Musterbildung bei der CO-Oxidation auf Pt (110) unter dem Einfluß von Laserlicht. Technical report, Fritz Haber Institute, Berlin.

Zhang, H., Fam, K.-S., Goh, T.-T., & Dai, X. (2018). When are influentials equally influenceable? the strength of strong ties in new product adoption. *Journal of Business Research*, *82*, 160–170.

Zhang, L., Priestley, J., DeMaio, J., Ni, S., & Tian, X. (2021). Measuring customer similarity and identifying cross-selling products by community detection. *Big data*, *9*(2), 132–143.

Zhang, L., Wang, J., & Liu, Z. (2023). What should lenders be more concerned about? developing a profit-driven loan default prediction model. *Expert Systems with Applications*, *213*, 118938.

Zhou, X., Mayer-Hamblett, N., Khan, U., & Kosorok, M. R. (2015). Residual weighted learning for estimating individualized treatment rules. *Journal of the American Statistical Association*, *112*, 169 – 187.

Zurada, J. & Lonial, S. (2005). Comparison of the performance of several data mining methods for bad debt recovery in the healthcare industry. *Journal of Applied Business Research (JABR)*, *21*(2).

Apéndice A

Descripción base de datos cobranza

En el estudio realizado en esta tesis, se construyó un conjunto de datos de cobranza donde cada fila representa una llamada a un deudor la cual tiene como objetivo cobrar la deuda al cliente para que la pague. En base a esto, se divide en dos tipos de variables para analizar cómo influye cada conjunto de datos en las tareas de predicción.
El primer conjunto de datos corresponde a la información financiera de los clientes (FD) y el segundo conjunto de datos corresponde a la información del *contact center* (CCD). A partir de esto, las variables se dividen en 5 grupos que se describen a continuación:

- *FD1.Deuda*: Variables que caracterizan la deuda pendiente. Estas variables son bien conocidas en tareas como la puntuación del comportamiento (Baesens, Roesch & Scheule, 2016; Thomas, Crook & Edelman, 2017).

- *FD2.Deudor*: Variables que describen al deudor. Estas variables se basan en estudios previos de cobro de deudas, como (Bahrami, Bozkaya & Balcisoy, 2020; van der Geer, Wang & Bhulai, 2018; Zurada & Lonial, 2005).

- *CCD1.BTTC*: Variables que permiten identificar el mejor día y hora para llamar a un cliente. Aunque la literatura científica sobre este tema es bastante escasa y obsoleta (Weeks, Kulka & Pierson, 1987), existen varios informes empresariales y otros recursos que indican variables que pueden construirse.

- *CCD2.Intentos*: Variables que describen los intentos recientes y pasados de contactar al cliente. Estas variables se basan en estudios previos de análisis de centros

de contacto, como (Bahrami, Bozkaya & Balcisoy, 2020; Bellotti, Brigo, Gambetti & Vrins, 2021).

- *CCD3.Contactos*: Podemos utilizar el historial de contactos anteriores para caracterizar al deudor en términos de su comportamiento con el centro de contacto. Se construyen variables como el número de contactos exitosos, el número de promesas de pago realizadas en el pasado y el número de promesas pagadas por el deudor.

Estos 5 grupos incluyen las siguientes variables:

FD1.Deuda:

- Unpaid-Debt: Cantidad de deuda pendiente.
- Delinquent-Debt: Cantidad de deuda morosa.
- Days-Late: Días de atraso en el pago de la deuda.
- Days-Debt-Assignment: Número de días de atraso en el pago de la deuda en el momento que se realizó la asignación que corresponde a la gestión de cobro por parte de la empresa..
- Daily-Debt-Amount: Cantidad de deuda morosa por día.

FD2.Deudor:

- Campaign: Categorización de los clientes de acuerdo a los días de atraso en el pago de su deuda, se divide en 3 categorías:
  - Campaign-Punishment, cuando un cliente tiene un atraso de más de 30 días en el pago de su deuda.
  - Campaign-Flow, cuando el cliente tiene un atraso de menos de 5 días en el pago de su deuda.
  - si no pertenece a ninguna de las categorías anteriores corresponde a que el cliente tiene un atraso entre 5 y 30 días en el pago de su deuda.

- Section: Categorización de los clientes de acuerdo a los días de atraso en el pago de su deuda. Hay 7 tramos del 0 al 6, donde 0 son los clientes con menos días de atraso en el pago y 6 los clientes con más días de atraso en el pago de la deuda.
- Risk-Mark: Marca de riesgo para el cliente basada en su historial de pago. Se divide en 3 categorías, alto, media y bajo.
- AVG-Payments: Promedio de pagos del deudor.
- Sd-Payments: Desviación estándar de los pagos del deudor.
- Amount-Payments: Cantidad de pagos que ha hecho el deudor.
- CellPhone: Si el número de teléfono del cliente está en la base de datos.
- Card-VISA: Entidad financiera a la que pertenece la tarjeta del cliente.
- Business-phone: Si el número de teléfono de trabajo del cliente esta en la base de datos.
- Alternative-phone: Si hay un número de teléfono alternativo del cliente en la base de datos.

CCD1.BTTC:

- Time-Call: Hora de la llamada que se está analizando.
- Morning-Call: La llamada que se está analizando se hizo en la mañana. Se considera mañana de las 8 hrs a las 13 hrs.
- Call-Monday-Thursday: La llamada que se está analizando se realizó entre lunes y jueves.
- Start-Month-Call: La llamada que se está analizando se hizo entre el día 1 y 10 del mes.
- Half-Month-Call: La llamada que se está analizando se realizó entre el día 11 y 20 del mes.
- Start-Month-Promises: La fecha en que el deudor promete pagar la deuda es entre el día 1 y 15 del mes.

CCD2.Attempts:

- Attempts: Número de intentos para contactar al deudor.
- MaxAttempts: Número máximo de intentos de contactar al deudor al momento de realizar la llamada.

CCD3.Contacts:

- Past-Calls: Número de llamadas que se le han realizado al deudor al momento de realizar la llamada.
- Past-Contacts: Número de contactos con el deudor antes de la llamada.
- Past-Promises: Número de promesas de pago que ha hecho el deudor antes de la llamada.
- Past-Paid-Promises: Número de promesas pagadas por el cliente en el pasado.
- Days-Call-PromisePay: Número de días entre la fecha en que se hace la promesa y la fecha en que se promete pagar.
- Promised-Amount-Yes: Al hacer una promesa de pago indica el monto a pagar.
- Promides-Amount-High: El monto que promete pagará es mayor a 100,000 pesos chilenos.

Apéndice B

Estadísticas base de datos venta cruzada

En la Tabla B.1 se presenta un resumen de las estadísticas de la base de datos de venta cruzada:

Cuadro B.1: Resumen estadísticas venta cruzada

Variable	Promedio (1)	D.E. (2)	Min. (3)	Max. (4)	Obs. (5)
Resultado y Tratamiento:					
Recomendación exitosa	0.1994	0.3995	0	1	61,363
Recomendación con campaña	0.0741	0.2619	0	1	61,363
Covariantes de los receptores:					
Masculino	0.2304	0.4211	0	1	30,574
Casado	0.7172	0.4503	0	1	30.574
Edad	37.54	7.48	23.14	70.21	30,574
Experiencia (Años)	3.22	3.37	0.00	27.17	30,574
Educacion					30,322
Escolar	0.0878				
Tecnico	0.3201				
Universitario	0.5414				
Postgrado	0.0506				
Cargo					30,574
Rep. de ventas	0.6526				
Supervisor	0.2687				
Manager	0.0787				
Covariantes de los emisores:					
Masculino	0.4897	0.4999	0	1	35,518
Casado	0.8311	0.3747	0	1	35.518
Edad	37.69	7.50	20.29	70.21	35,517
Experiencia (Años)	3.76	3.52	0.00	27.17	35,517
Educacion					35,432
Escolar	0.0350				
Tecnica	0.1402				
Universitaria	0.7943				
Postgrado	0.0306				
Cargo					35,519
Rep.de ventas	0.6995				
Supervisor	0.2251				
Manager	0.0753				

Notes: Las covariables para receptores y emisores corresponden a 605 y 816 personas diferentes a lo largo del tiempo, respectivamente.

Apéndice C

Construcción de variables para venta cruzada

Para el modelo predictivo, se creó una serie de variables que reflejan la relación entre el emisor y el receptor, así como factores adicionales que pueden ser relevantes según estudios de marketing, estas se describen a continuación:

- *Indegree*: Se crean variables de redes sociales, como la centralidad de indegree y outdegree, que pueden reflejar comportamientos del proceso de venta cruzada, como se ha sugerido en otras tareas de marketing, como el marketing viral y la segmentación de clientes (Hinz, Skiera, Barrot & Becker, 2011). Indegree, se define como el número de recomendaciones que un receptor gestiona en un período de tiempo, esto entrega información sobre su capacidad para convertir las referencias entrantes en nuevas ventas. Dado que este proceso puede requerir mucho tiempo, se considera que los vendedores con un alto indegree pueden verse abrumados por la carga de trabajo y, por lo tanto, son menos eficaces para que estas recomendaciones tengan éxito, en promedio.

- *outdegree and senders'tenure*: Esta variable se calcula como el número de conexiones que tiene un remitente. Según Hinz, Skiera, Barrot & Becker (2011), los clientes bien conectados tienen más probabilidades de hacer recomendaciones exitosas que los clientes seleccionados aleatoriamente en un contexto de marketing viral. Se considera que este patrón puede extenderse a los vendedores. En la misma dirección, se analiza la influencia de la antigüedad de los remitentes, asu-

miendo que los empleados con alta experiencia tienen más éxito en transformar las referencias en nuevas ventas.

- *Strengh of tie*: Otra variable de redes sociales que se analiza es la intensidad de la relación entre el remitente y receptor. Según Zhang, Fam, Goh & Dai (2018), la fuerza de los vínculos tiene una influencia importante en los clientes en el contexto de la adopción de nuevos productos. Se extiende este razonamiento a este análisis y se sugiere que los vínculos fuertes, definidos como el número de referencias del emisor al receptor, pueden tener un impacto positivo en la venta cruzada, ya que pueden acelerar el proceso gracias a la confianza mutua entre los vendedores involucrados.

- *Homophily*: Otro factor clave para el éxito en la venta cruzada puede ser la similitud entre el emisor y el receptor. Estudios previos han demostrado que la homofilia puede estar relacionada con las ventas, ya que genera confianza y reciprocidad (van den Bulte & Wuyts, 2007). Se crea una variable ficticia que refleja si el emisor y el receptor son similares, basándonos en las variables reportadas en Tabla B.1. Se creó un modelo de agrupamiento utilizando K-medias, empleando el índice de Davies-Bouldin para evaluar la compacidad del conglomerado. La homofilia se define entonces como si los dos vendedores pertenecen o no al mismo conglomerado. Si esto sucede y comparten atributos comunes, se cree que es más probable que produzcan recomendaciones exitosas.

- *Hierarchy and normative pressure*: En caso de que el emisor tenga una jerarquía superior a la del receptor en cuanto al puesto, se considera que este último se encuentra bajo presión normativa, en el sentido de que realizaría esfuerzos adicionales para lograr una recomendación exitosa. La presión normativa se ha estudiado como un factor de influencia social, ya que los empleados tienden a obedecer a otros en una jerarquía superior porque creen que les afectan directa o indirectamente. Se crea una variable ficticia que indica si el emisor tiene una jerarquía superior a la del receptor.

- *Natural relationships*: Finalmente, el éxito de una recomendación puede vincularse a si existe o no una secuencia natural y lógica de consumo del cliente, como se ha

sugerido en estudios previos (Gupta & Zeithaml, 2006). En el ámbito bancario, un patrón natural de venta cruzada sería derivar a un cliente de un producto más sencillo, como una cuenta corriente, a uno más complejo, como hipotecas o inversiones. Nuestra hipótesis es que las derivaciones entre unidades de negocio con una relación natural son más efectivas que las que no lo son. Creamos una variable ficticia que indica esta relación con base en la clasificación realizada por los gerentes del banco. Esta es la única variable que no se basa en las características de los vendedores, sino en la naturaleza de la derivación en sí.

Apéndice D

Descripción bases de datos para Bounded ITE

Se consideran los siguientes conjuntos de datos para analizar la evaluación del enfoque Bounded ITE:

- Venta cruzada (*Crosssell*): Mismo conjunto de datos utilizado para la aplicación de modelos causales basado en beneficio descrito en el capítulo 3.
- Hillstrom 'All'(*HillAll*): Este conjunto de datos proviene de una campaña de marketing enfocada en ropa para hombres y mujeres. Está dividido en tres: un tercio corresponde a correos promocionales de ropa masculina, otro tercio a correos de ropa femenina y el resto a clientes que no fueron objetivo de la campaña (grupo de control). Las etiquetas indican si la campaña resultó en una compra (Devriendt, Moldovan & Verbeke, 2018b). Se trata de un conjunto de datos público disponible en el blog MineThatData [1].
- Hillstrom 'Men' (*HillMen*): Es un subconjunto del conjunto de datos HillAll que incluye únicamente la campaña dirigida a ropa de hombre junto con el grupo de control.
- Hillstrom 'Women'(*HillWomen*): Es un subconjunto del conjunto de datos *HillAll* que incluye solo la campaña de ropa de mujer y el grupo de control.

[1] `https://blog.minethatdata.com/`

- ISMS Durables 2 (*Informs*): Este conjunto de datos se refiere a una campaña de correo promocional navideño realizada por un minorista estadounidense de productos electrónicos. Cerca del 50 % de los clientes fue elegido de manera aleatoria para recibir la promoción (Ni, Neslin & Sun, 2012). El conjunto de datos está disponible públicamente a través de la revista Marketing Science de Informs.

- Seguros (*Insurance*): Este también es un conjunto de datos de una campaña de marketing directo, en el que las etiquetas indican si la iniciativa de marketing genera una venta (Devriendt, Moldovan & Verbeke, 2018b). Este es un conjunto de datos público relacionado con el sector de los seguros y está disponible en el paquete R Ínformation'.

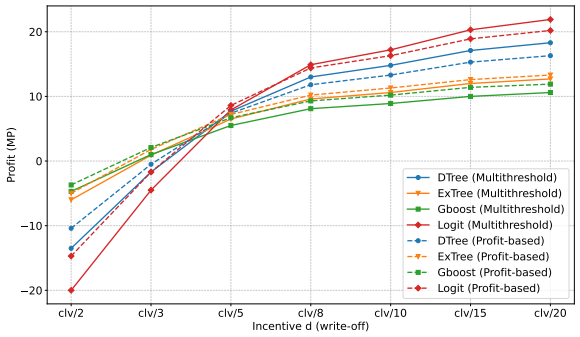
- Udemy Retail (*Udemy*): Este conjunto de datos consiste en una campaña de marketing para un minorista. El objetivo principal es predecir las respuestas de los clientes hasta dos meses después del contacto (Devriendt, Moldovan & Verbeke, 2018b). Este es un conjunto de datos públicos disponible en la página web de Udemy[2].

[2] `https://www.udemy.com/uplift-modeling/`

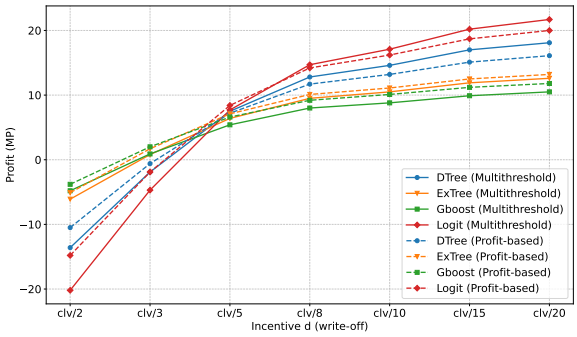
Apéndice E

Resultados profit para modelos causales en cobranza

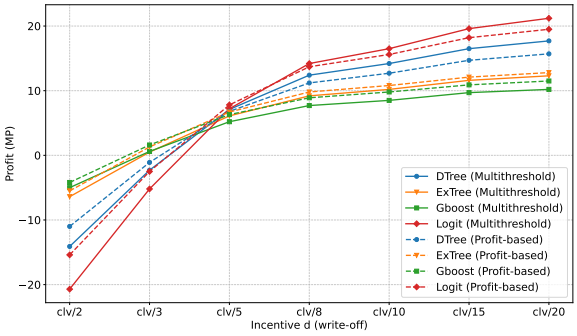
Este anexo presenta los resultados de profit para modelos causales estándar y la estrategia multisegmento para las 3 estrategias: MOA, S-learner y T-learner.



(a) Costo de contacto $f = 1$.

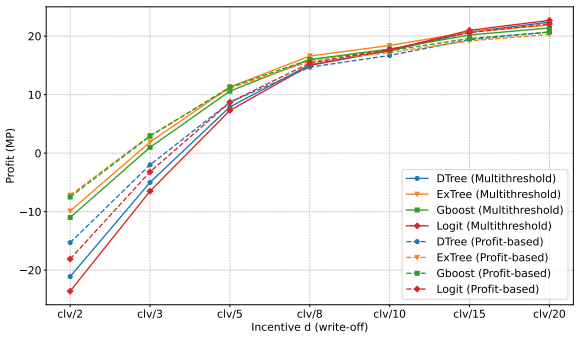


(b) Costo de contacto $f = 2$.

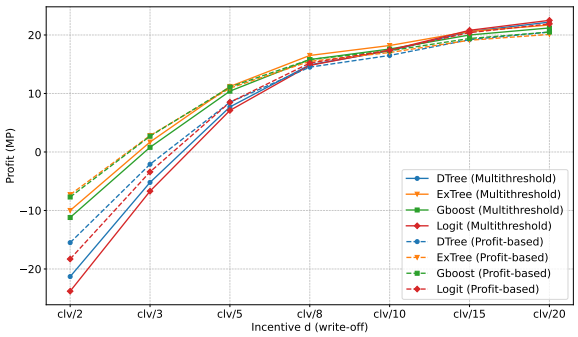


(c) Costo de contacto $f = 5$.

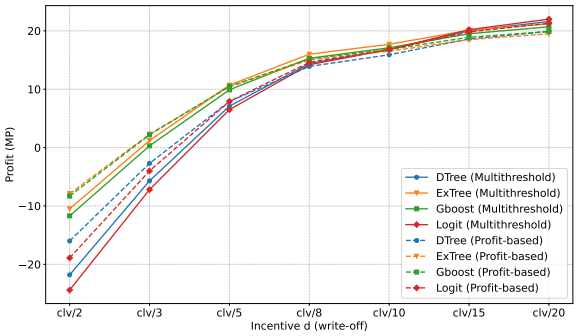
Figura E.1: Resultados profit para modelos causales estándar y la estrategia multisegmento utilizando el enfoque MOA.



(a) Costo de contacto $f = 1$.

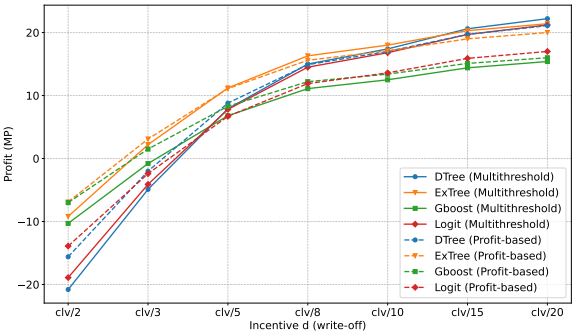


(b) Costo de contacto $f = 2$.

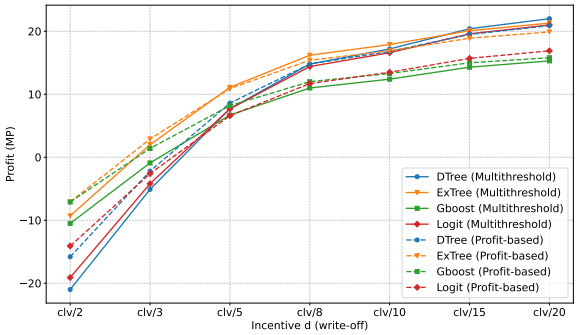


(c) Costo de contacto $f = 5$.

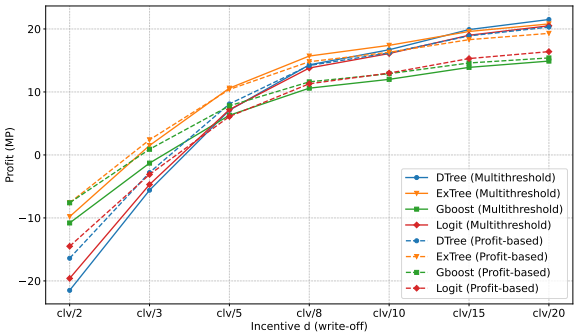
Figura E.2: Resultados profit para modelos causales estándar y la estrategia multisegmento utilizando el enfoque S-learner.



(a) Costo de contacto $f = 1$.



(b) Costo de contacto $f = 2$.



(c) Costo de contacto $f = 5$.

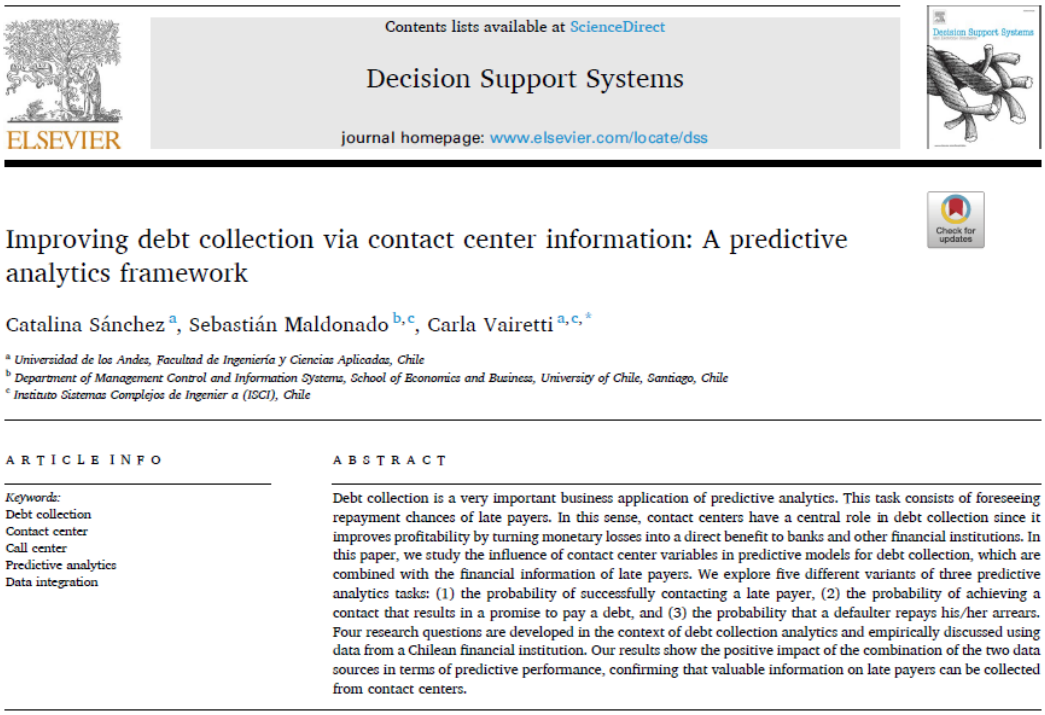
Figura E.3: Resultados profit para modelos causales estándar y la estrategia multisegmento utilizando el enfoque T-learner.

Apéndice F

Publicaciones

Como parte de esta investigación, tres trabajos fueron publicados, y un cuarto trabajo está en proceso de revisión.

- Sánchez, C., Maldonado, S. y Vairetti, C., Improving debt collection via contact center information: a predictive analytics framework. Decision Support Systems, Volume 159, 2022, 113812, ISSN 0167-9236.



Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

Improving debt collection via contact center information: A predictive analytics framework

Catalina Sánchez [a], Sebastián Maldonado [b,c], Carla Vairetti [a,c,*]

[a] Universidad de los Andes, Facultad de Ingeniería y Ciencias Aplicadas, Chile
[b] Department of Management Control and Information Systems, School of Economics and Business, University of Chile, Santiago, Chile
[c] Instituto Sistemas Complejos de Ingenier a (ISCI), Chile

ARTICLE INFO

Keywords:
Debt collection
Contact center
Call center
Predictive analytics
Data integration

ABSTRACT

Debt collection is a very important business application of predictive analytics. This task consists of foreseeing repayment chances of late payers. In this sense, contact centers have a central role in debt collection since it improves profitability by turning monetary losses into a direct benefit to banks and other financial institutions. In this paper, we study the influence of contact center variables in predictive models for debt collection, which are combined with the financial information of late payers. We explore five different variants of three predictive analytics tasks: (1) the probability of successfully contacting a late payer, (2) the probability of achieving a contact that results in a promise to pay a debt, and (3) the probability that a defaulter repays his/her arrears. Four research questions are developed in the context of debt collection analytics and empirically discussed using data from a Chilean financial institution. Our results show the positive impact of the combination of the two data sources in terms of predictive performance, confirming that valuable information on late payers can be collected from contact centers.

Figura F.1: Paper Improving debt collection via contact center information: a predictive analytics framework

- Vairetti, C., Vargas, R., Sánchez, C. et al. Improving incentive policies to salespeople cross-sells: a cost-sensitive uplift modeling approach. Neural Comput & Applic 36, 17541–17558 (2024).

Neural Computing and Applications (2024) 36:17541–17558
https://doi.org/10.1007/s00521-024-10051-2

ORIGINAL ARTICLE

Check for updates

Improving incentive policies to salespeople cross-sells: a cost-sensitive uplift modeling approach

Carla Vairetti[1,4] · Raimundo Vargas[1] · Catalina Sánchez[1] · Andrés García[1] · Guillermo Armelini[2] · Sebastián Maldonado[3,4]

Received: 28 March 2024 / Accepted: 19 June 2024 / Published online: 28 June 2024

Abstract
In this study, we present a novel cost-sensitive approach for uplift modeling in the context of cross-selling and workforce analytics. We leverage referrals from sales agents across business units to estimate the individual treatment effects of incentives on the cross-selling outcomes within a company. Uplift modeling is employed to predict relationships between salespeople that should be encouraged based on the probability of successful cross-selling - defined when a customer accepts the product suggested by sales agents. We conducted experiments on data from a Chilean financial group, evaluating both statistical and profit metrics. Exploring various machine learning classifiers for predictive purposes, we observed a significant improvement over the current approach, which exhibits an uplift below 0.01. Finally, we show that selecting the best classifier with profit metrics results in a 31.6% improvement in terms of average customer profit. This emphasizes the importance of defining an adequate compensation scheme and integrating it into the modeling process.

Keywords Uplift modeling · Cost-sensitive learning · Business analytics · Cross-selling · Workforce analytics

1 Introduction

Figura F.2: Paper Improving incentive policies to salespeople cross-sells: a cost-sensitive uplift modeling approach

- Sánchez, C., Vairetti, c., Coussement, k. y Maldonado, S., Bounded Individualized Treatment Effect: A Novel Approach for Uplift Modeling, in IEEE Access, vol. 13, pp. 145599-145610, 2025

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 31 July 2025, accepted 13 August 2025, date of publication 15 August 2025, date of current version 22 August 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3599482

RESEARCH ARTICLE

Bounded Individualized Treatment Effect: A Novel Approach for Uplift Modeling

CATALINA SÁNCHEZ[1], CARLA VAIRETTI[1,4], (Member, IEEE), KRISTOF COUSSEMENT[2], AND SEBASTIÁN MALDONADO[3,4], (Senior Member, IEEE)

[1]Facultad de Ingeniería y Ciencias Aplicadas, Universidad de Los Andes, Santiago 7620001, Chile
[2]CNRS, UMR 9221-Lille Economie Management (LEM), IÉSEG School of Management, University of Lille, 59000 Lille, France
[3]Department of Management Control and Information Systems, School of Economics and Business, University of Chile, Santiago 8330111, Chile
[4]Instituto Sistemas Complejos de Ingeniería (ISCI), Santiago 8370449, Chile

Corresponding author: Carla Vairetti (cvairetti@uandes.cl)

This work was supported in part by the Agencia Nacional de Investigación y Desarrollo (ANID) Programa de Investigación Asociativa (PIA)/PUENTE under Grant AFB230002 and in part by the Fondo Nacional de Desarrollo Científico y Tecnológico (FONDECYT)-Chile under Grant 12200007 and Grant 1250045. The work of Catalina Sánchez was supported by the Becas Doctorado Nacional under Grant 21230040.

ABSTRACT The advent of artificial intelligence has greatly facilitated the prediction of individualized treatment effects, expanding the field of causal inference beyond the traditional realm of econometric studies. This shift has enabled prescription of personalized treatments via predictive analytics, allowing businesses to deliver personalized and relevant interactions with customers across various touchpoints. In this paper, we propose a novel approach for uplift model estimation called Bounded Individualized Treatment Effect (BITE), which is designed for business analytics tasks. The goal is to prioritize customers who will be persuaded by the treatment, by filtering individuals based on their baseline probabilities. Our experiments show the virtues of the BITE approach on seven benchmark datasets and 21 uplift classifiers, achieving best average performance compared to traditional uplift modeling and standard predictive classification.

INDEX TERMS Uplift modeling, machine learning, business analytics, individualized treatment effects.

I. INTRODUCTION

Causal inference is concerned with learning about cause medical treatments [7], [8], and to target students to prevent dropout [9].

Figura F.3: Paper Bounded Individualized Treatment Effect: A Novel Approach for Uplift Modeling