



Proyecto pasantía
Gestión comercial BCIPlus+
Sección 13

Nombre: Nicolás Alejandro Zapata Chávez

Docente: Rafael Cereceda

Carrera: Ingeniería civil informática / Ingeniería civil industrial

Fecha: 26 de noviembre del 2023

Índice

1. Resumen Ejecutivo	5
2. Abstract	6
3. Introducción	7
3.1. La empresa	7
3.2. Contexto y problema	8
3.3. Metas BCIPPlus+ y conclusiones del problema	12
3.4. Importancia del proyecto	13
4. Objetivo del proyecto	13
4.1. Objetivo general	13
4.2. Objetivos específicos	13
5. Estado del arte	14
5.1. Machine learning	14
5.2. Plataformas de inteligencia empresarial	15
5.3. Maximización de beneficios asociados al análisis de campañas comerciales	16
5.3.1. Modelos predictivos de ventas	17
5.3.2. Modelos de optimización	17
6. Solución	20
6.1. Matriz de riesgos	21
6.2. Evaluación económica	22
6.3. Metodología	24
6.4. Plan de implementación	26
6.4.1. Etapa 1: Nivel bruto (agosto)	27
6.4.2. Etapa 2: Análisis de sensibilidad (septiembre)	27
6.4.3. Etapa 3: Coeficientes de impacto y modelos de optimización (octubre-mediados de noviembre)	27
6.4.4. Etapa 4 Pruebas finales y cierre de proyecto (mediados de noviembre y diciembre)	27
6.5. Medidas de desempeño	28
6.5.1. Tiempo	28
6.5.2. Eficacia	28
6.6. Desarrollo del proyecto	29
6.6.1. Etapa 1: Nivel bruto	29
6.6.2. Etapa 2: Análisis de sensibilidad	31
6.6.3. Etapa 3: Coeficientes de impacto y modelos de optimización	35
7. Conclusiones	36
7.1. Resultados cualitativos y cuantitativos	36
7.2. Etapa 4: Pruebas finales y cierre del proyecto	36
7.3. Comentarios finales	38
8. Bibliografía	40
9. Anexo	41

Índice de tablas

Tabla 1: Métricas calculadas por el banco.	12
Tabla 2: Extracto de la matriz de riesgos.	21
Tabla 3: Evaluación económica considerando los principales costos asociados al proyecto en pesos chilenos.	22
Tabla 4: Evaluación económica del proyecto en pesos chilenos.	23
Tabla 5: Comparación del rendimiento entre los experimentos.	33
Tabla 6: Comparación de rendimiento entre las acciones de campañas antiguas y el rendimiento esperado por el robot.	37
Tabla 7 (Anexo): Matriz de riesgo del proyecto	43
Tabla 8 (Anexo): Rendimientos de modelos de predicción.	45
Tabla 9 (Anexo): Decisiones de campañas en BCIPPlus+ para el mes de octubre.	52
Tabla 10 (Anexo): Decisiones de campañas del robot para el mes de octubre.	52

Índice de ilustraciones

Ilustración 1: Proceso de campañas de BCIPPlus+ en Bizagi.	11
Ilustración 2: Flujo de trabajo informático.	24
Ilustración 3: Flujo de trabajo industrial.	25
Ilustración 4: Carta Gantt del proyecto BCIPPlus+.	28
Ilustración 5: Panel del robot BCIPPlus+ etapa 1.	30
Ilustración 6: Visualización del impacto en las sucursales de la marca Adidas.	31
Ilustración 7 (Anexo): Diseño del robot en su segunda etapa de desarrollo.	48
Ilustración 8 (Anexo): Vista de simulador y ventas en el diseño del robot en la tercera etapa de desarrollo.	50
Ilustración 9 (Anexo): Vista de Benchmark en el diseño del robot en la tercera etapa de desarrollo.	51
Ilustración 10 (Anexo): Vista de Campaign Manager en el diseño del robot en la tercera etapa de desarrollo.	51

Índice de gráficos

Gráfico 1: Actividad de clientes MACH en las variables de actividad de la plataforma desde Agosto 2022 hasta Agosto 2023.	9
Gráfico 2: Actividad total de la plataforma MACH desde Enero 2023 hasta Julio 2023.	10
Gráfico 3: Promedio del impacto de ventas con cashback esperadas en una marca mediana.	34
Gráfico 4 (Anexo): Promedio del impacto de ventas con cashback esperadas en una marca grande.	47
Gráfico 5 (Anexo): Promedio del impacto de ventas con cashback esperadas en una marca pequeña.	47

1. Resumen Ejecutivo

La asignación efectiva de recursos es esencial para el éxito y el buen rendimiento de los diferentes procesos que existen en una determinada organización, en especial si los recursos existentes son limitados y se busca responder con mayor velocidad y precisión a los requerimientos de la sociedad.

Las buenas prácticas para la asignación efectiva de recursos se pueden encontrar en diferentes áreas, como por ejemplo, la manufactura, tecnología, salud y los negocios. Esto es visto en la práctica, por ejemplo, en el desarrollo de productos, uso de tierras agrícolas, gestión de proyectos de *software*, asignación de personal e incluso en campañas comerciales. Una buena asignación de recursos permite tomar decisiones inteligentes, eficientar procesos, mitigar riesgos, e incluso, generar una mayor competitividad en el mercado.

El proyecto del practicante se elaboró en vías de mejorar la eficiencia y efectividad de las alianzas comerciales en su programa de beneficios BCIPlus+. A través de la combinación de la lógica ingenieril, tanto del área industrial como del área informática, se aborda el desafío de automatizar y mejorar la toma de decisiones relacionadas con las alianzas comerciales y el *cashback* en el programa.

Este proyecto busca orientar las decisiones sobre las campañas comerciales hacia la obtención de clientes activos para el negocio, disminuir la brecha de información sobre las capacidades de los comercios potenciales para el banco, y, generar un proceso fluido de información que permita mejorar la eficiencia en sus procesos de alianzas y campañas comerciales.

Para ello, se investigó el impacto del *cashback* en el programa a través de los datos, utilizando el desarrollo tecnológico sobre modelos de aprendizaje automático para estimar su comportamiento, y posteriormente, gestionar las campañas de beneficio.

Combinando las capacidades de ambos mundos ingenieriles, con el uso apropiado de los datos y la gestión inteligente, es que se conformó el proyecto de gestión comercial BCIPlus+.

2. Abstract

Effective resource allocation is crucial for the success and optimal performance of various processes within an organization, especially when dealing with limited resources and the need to respond swiftly and accurately to social requirements.

Best practices for resource allocation can be found in different areas such as manufacturing, technology, healthcare, and business. This can be seen in practice, for example, in product development, agricultural land use, software project management, personnel allocation, and even commercial campaigns. A good resource allocation strategy enables intelligent decision-making, process efficiency, risk mitigation, and fosters increased competitiveness in the market.

The internship project was developed to enhance the efficiency and effectiveness of commercial partnerships within the BCIPPlus+ benefits program. By combining engineering logic from both industrial and IT domains, the challenge of automating and improving decision-making related to commercial partnerships and cashback in the program is addressed.

This project aims to guide decisions regarding commercial campaigns towards acquiring active customers for the business, reduce information gaps about the potential capabilities of partner businesses for the bank, and establish a streamlined information process to enhance efficiency in alliance and commercial campaign processes.

To achieve these objectives, the project investigates the impact of cashback in the program through data analysis, leveraging technological development on machine learning models to estimate behavior, and subsequently managing benefit campaigns.

With both engineering realms, and employing data and smart management effectively, the BCIPPlus+ commercial management project was conceptualized.

3. Introducción

3.1. La empresa

El Banco de Crédito e Inversiones (desde ahora en adelante BCI), es uno de los bancos más importantes y antiguos de Chile. Fue fundado en 1937 por Juan Yarur Lolas, con el objetivo de apoyar a las empresas pequeñas y medianas del país. Durante los años, el banco se ha caracterizado por la innovación y calidad de sus servicios, buscando satisfacer y resolver las necesidades financieras de tanto las personas como las empresas.¹

Algunos de los productos que ofrece esta entidad financiera son: cuentas corrientes, créditos a corto plazo para el capital de trabajo, financiamiento, inversiones, seguros, entre otros. A lo largo del tiempo, el banco ha sido reconocido dentro de la industria, ganando diversos premios, como por ejemplo, el de sostenibilidad en sus operaciones, experiencia de cliente, innovación, y liderazgo en retener talento joven.¹

A grandes rasgos, el banco BCI se compone de 4 grandes áreas: Banca personas, Banca empresas, Inversiones y trading, y Ecosistema digital. Estas áreas se subdividen en subáreas, por ejemplo, dentro del Ecosistema digital encontramos: *New ventures*, *Data & analytics*, MACH, y, *Loyalty*.

El área de inserción del practicante es en la subárea *New ventures* del Ecosistema digital, específicamente en el equipo de BCILabs, desempeñando el rol de estrategia de inteligencia de negocios (desde ahora en adelante denominado BIS). Este rol se dedica a las investigaciones de negocios, proponiendo soluciones e investigación a través de los datos y apoyar, con ello, los diferentes proyectos del equipo.

El equipo de BCILabs se encarga de la investigación e innovación del banco y tiene principalmente 4 objetivos:

- A. **Innovación interna:** Se refiere al diseño, seguimiento y ejecución de procesos de innovación junto a áreas internas del banco para desarrollar nuevos productos y servicios.
- B. **Innovación abierta:** Se refiere a traccionar capacidades junto a Pymes o FINTECH del ecosistema para complementar propuesta de valor del banco BCI.
- C. **Exploración de tendencias:** Se refiere a investigar y recopilar tendencias de negocios, metodologías de innovación y nuevas tecnologías como fuente de oportunidades que diferencien al banco del resto.

- D. **Relaciones públicas:** Se refiere a generar una red de conocimiento compartido entre el interior de la corporación y el ecosistema de negocio.

3.2. Contexto y problema

El equipo de BCILabs cuenta con diferentes proyectos enfocados en los objetivos mencionados en el punto 3.1, los cuales están directamente relacionados con las jugadas estratégicas del banco. El proyecto del practicante estará enfocado dentro del programa de BCIPlus+.

BCIPlus+ es un programa de beneficios abierto y masivo que entrega *cashback*^A a los clientes del banco mediante la cuenta MACH o BCI. Este programa, como modelo de negocio, busca generar beneficios tanto para los clientes como para los comercios adheridos.

Desde el punto de vista de las personas, se obtiene un beneficio monetario flexible, ya que no posee restricciones de medio de pago (puede ser débito o crédito), y, además, tiene múltiples opciones para acumular el saldo y usarlo en los comercios.

Por otra parte, respecto al punto de vista de los comercios. Estos obtienen beneficios a través del programa debido a la masiva cantidad de clientes pertenecientes al banco que acceden a sus servicios y/o productos, permitiendo así no solo generar más ventas, sino que también mejorar su visibilidad. Como adicional, el banco les brinda apoyo con datos de las compras que les permite tomar mejores decisiones.

Por su parte, el banco también recibe beneficios de este programa. Este obtiene un ingreso mediante una comisión por venta derivada^B, la cual se define entre el banco y el comercio adherido al inicio de la conformación de la campaña de beneficio.

Actualmente, el programa busca extenderse al medio físico a través de la aplicación MACH, y posicionarse así como el principal centro de beneficios del banco debido a la gran cantidad de clientes que la plataforma posee. Se ha adentrado a realizar un plan piloto dentro de esta, sin embargo, en el tiempo que llevan de progreso, han notado que existe una baja actividad de parte de los clientes con el uso de la cuenta, y se espera que el programa ayude a incrementarla.

Por lo que se identifica un problema que está ocurriendo en el área de ecosistema digital, específicamente en el establecimiento del programa de beneficios BCIPlus+, pues existe una gran cantidad de personas que tienen el producto MACH, medio por el cual el programa es impulsado, pero no lo utilizan frecuentemente.

Para poder validar este problema y verificar esta baja actividad, se obtuvo una serie de datos de variables de actividad y el total de usuarios inscritos en la plataforma. A este total de usuarios en MACH lo llamaremos como “Stock MACH”, el cual corresponde a 3,8 millones de clientes actualmente. Mientras que las variables de actividad se definen de la siguiente forma:

- **Usuarios P2M:** Clientes que han generado pagos hacia comercios.
- **Usuarios CASH IN:** Clientes que han generado ingreso de dinero a sus cuentas de MACH.
- **Usuarios PLUS+:** Clientes que han generado pagos que han implicado un *cashback*.
- **Usuarios P2P:** Clientes que han generado pagos hacia otras personas naturales.
- **Nuevo ingreso:** Cantidad de clientes nuevos inscritos en la plataforma MACH en un determinado periodo de tiempo.

Con la información de la empresa, se observará la actividad de los clientes en el año 2023, contraponiendo el stock de clientes MACH frente a las demás variables de actividad. Por otro lado, se analizará también el comportamiento del flujo de clientes nuevos de MACH en cada periodo del mes del año 2023 frente a las variables de actividad. Los resultados se pueden observar en el gráfico 1 y en el gráfico 2 a continuación.

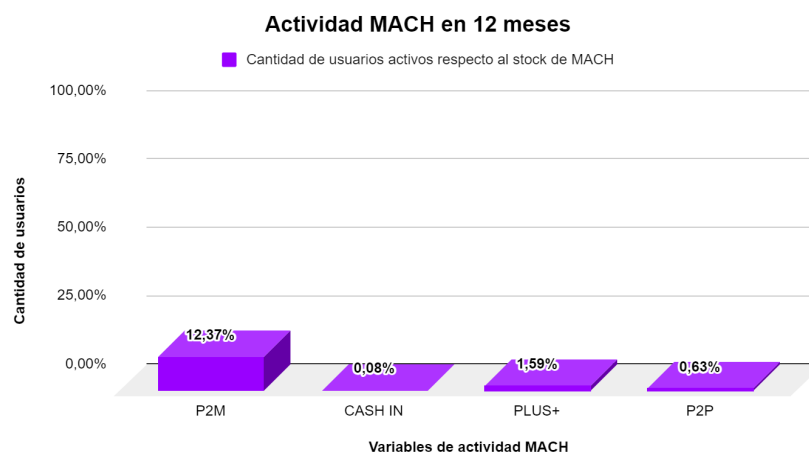


Gráfico 1: Actividad de clientes MACH en las variables de actividad de la plataforma desde agosto 2022 hasta agosto 2023.

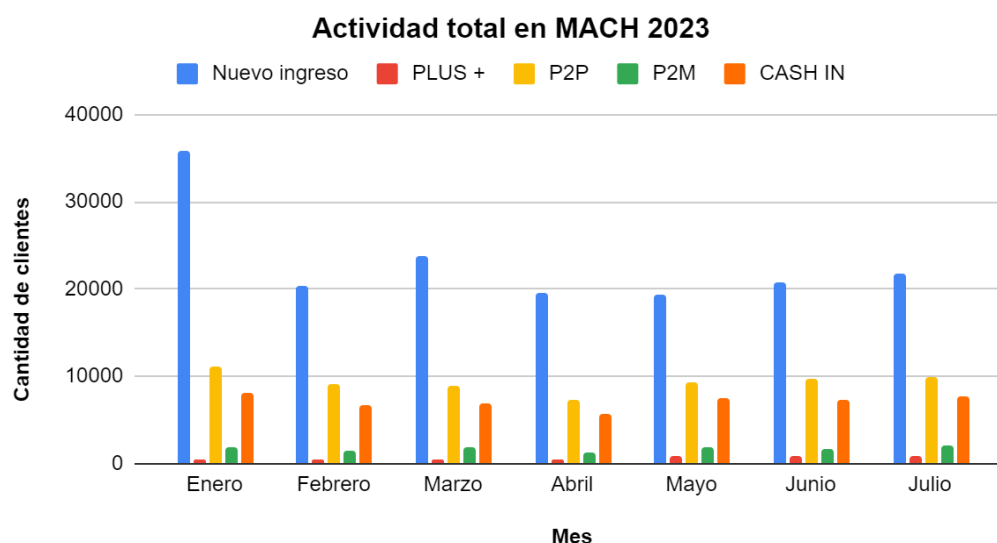


Gráfico 2: Actividad total de la plataforma MACH desde enero de 2023 hasta julio 2023.

Analizando los gráficos, podemos determinar, en primer lugar, que como máximo, sólo un 12,37% del stock de MACH utiliza la plataforma para P2M, además, menos de la mitad de los nuevos ingresos realiza alguna actividad relevante dentro de la plataforma. Por lo que, en general, el uso que se le está dando actualmente no llega a superar el 13% de su capacidad total, y considerando que será la principal fuente de despliegue del programa de BCIPlus+ ese número necesita aumentar.

Por otra parte, junto a BCIPlus+ y el equipo de BCILabs, se pudieron detectar los siguientes dolores:

1. BCIPlus+ está recién incorporándose al ecosistema digital y hay certeza de que las campañas con los comercios generan actividad en los usuarios, por lo que se están buscando formas de que mejoren las tomas de decisiones de estas mismas. Sin embargo, no hay certeza de qué empresas entregan valor al banco.
2. No se han gestionado métricas que ayuden en la toma de decisión.
3. No existe un presupuesto organizado respecto a las campañas.

Dado que se conoce que las campañas generan actividad en los usuarios, cabe mencionar cómo estas se realizan, pues en la actualidad no existe formalidad durante el proceso de estas mismas, en otras palabras, no se conoce cómo están funcionando, ni qué procesos las componen, y mucho menos, con qué negocios conviene realizarlas. Por lo que

para su mayor entendimiento, el practicante realizó un diagrama en Bizagi que se puede observar en la ilustración 1, describiendo sus fases y roles.

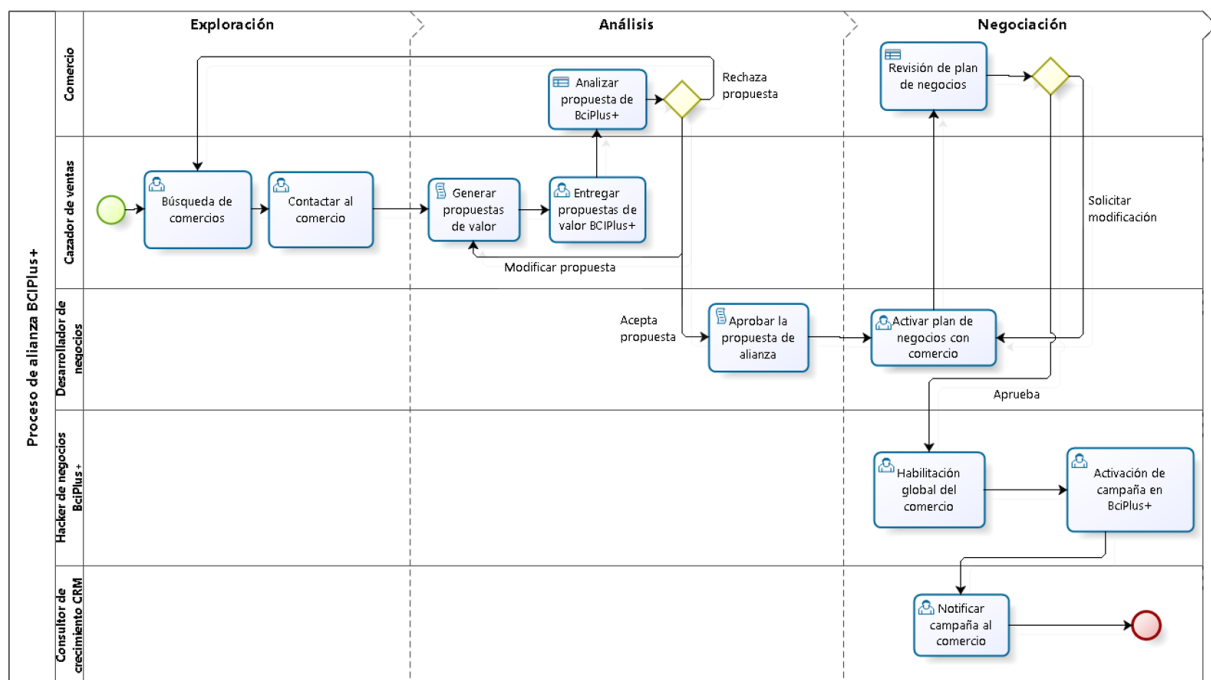


Ilustración 1: Proceso de campañas de BCIPlus+ en Bizagi.

Entonces, con el apoyo del diagrama, se debe considerar que durante el proceso de campaña existen 3 principales fases: exploración, análisis, y negociación.

La fase de exploración es donde se genera un primer acercamiento sobre los comercios existentes y se contacta con ellos para promover y dar un conocimiento inicial sobre los beneficios que BCI posee para apoyarlos. Tiene una duración aproximada de una semana, y busca limitar el potencial de las empresas disponibles en el mercado basándose en el rubro, tramo de ventas, ubicación, etc.

Luego se pasa a la fase de análisis, que también tiene una duración de una semana. En esta fase se busca principalmente analizar el potencial de ventas, clientes activos por periodo y comisiones potenciales que podría generar, recopilando la información para negociar con el comercio (transacciones de clientes BCI, potencial de crecimiento, competencia, etc.).

En esta fase se empieza a trabajar con una propuesta más desarrollada para una posible alianza temporal junto a BCI con el objetivo de generar campañas. Se analizan las capacidades del comercio y se construyen las principales propuestas de valor para el banco,

que, posteriormente, revisará el comercio. Cabe mencionar que en este punto, el comercio puede tanto aceptar, rechazar o querer modificar estas propuestas, por lo que el flujo de actividades puede volver tanto a la fase de exploración con otro comercio o volver a generar otras propuestas.

En el caso de que se acepte la propuesta, ingresamos a la fase de negociación. Esta fase tiene una duración de dos semanas y en ella, el desarrollador de negocios ^c aprueba la propuesta formalmente a nivel interno y activa un plan de negocios enfocado principalmente al uso de *cashback* en el comercio. El plan de negocios vuelve a ser presentado, pudiendo ser aceptado o modificado.

Una vez que el plan de negocio esté aprobado, y se hayan acordado las comisiones respectivas entre el banco y el comercio, el hacker de negocios ^d de BCIPlus+ habilita el comercio en los sistemas digitales y verifica el correcto funcionamiento interno de la campaña.

Finalmente, con todo puesto en marcha, el consultor de crecimiento CRM^e, notifica las campañas al comercio y se encarga de instruir al comercio sobre los usos y funcionamiento de las campañas en terreno.

3.3. Metas BCIPlus+ y conclusiones del problema

Las estrategias realizadas en el área de BCIPlus+ son evaluadas con métricas de desempeño que se enfocan en el rendimiento de las conformaciones de alianza y las campañas comerciales. En la tabla 1 se puede observar los valores que tenían previamente calculados en el banco.

Métrica	Nivel de cumplimiento
MAUS: Clientes activos por mes (Plus+) ^f	12%
GMV: volumen de ventas (Plus+) ^g	11%

Tabla 1: Métricas calculadas por el banco.

Por lo que podemos determinar que, a nivel estratégico, el programa de BCIPlus+ posee un bajo rendimiento en su desempeño, y, por lo tanto, lo analizado por el practicante y lo que el banco ya tenía estimado, indica que el problema existe y la inactividad de parte de los usuarios presentes en la aplicación de MACH no cumple con las expectativas del banco.

3.4. Importancia del proyecto

La estrategia actual del banco es que BCIPlus+ sea el medio principal por el cual se le entregue beneficios al cliente, no solo a través del *cashback* que permite tener mayor recurrencia de compra, sino que también con descuentos comunes.

La idea de la centralización de beneficios es que, a largo plazo, BCIPlus+ pase a ser un medio robusto en el cual los clientes quieran adquirir otros productos del banco, generar saldos en sus cuentas, realizar mayores transacciones, tener un mayor volumen de ventas, metas sobre comisiones, entre otros.

Por parte de los comercios, estar presentes en esta plataforma con accesos masivos de clientes también tiene un valor agregado al programa que BCI les está ofreciendo. Por otra parte, de cara al público y de forma estratégica, han podido validar que el *cashback* promueve una mayor liquidez en la mente de las personas, mayor funcionalidad y flexibilidad, dado que, este saldo que se devuelve no está ligado a algún catálogo de ofertas o restricción. El cliente ve lo que gana mediante el uso.

4. Objetivo del proyecto

4.1. Objetivo general

El objetivo general del proyecto es aumentar la captura de MAUS de las campañas en un 30% de aquí a finales del año 2023.

4.2. Objetivos específicos

Los objetivos específicos planteados para el proyecto son:

- Automatizar el análisis de empresas con las cuales BCIPlus+ podría generar una alianza.
- Generar estimaciones de ventas sobre los comercios y/o variables relevantes.
- Entender el impacto del *cashback* en el contexto del programa BCIPlus+ y los comercios adheridos.
- Resolver las mejores decisiones de campaña para cada una de las empresas que estén dentro o fuera del programa BCIPlus+, con tal de maximizar la cantidad de MAUS a obtener.

5. Estado del arte

A continuación, se expondrá una visión general del conocimiento existente para el tema de investigación. Esto brindará la oportunidad de analizar algunas metodologías o investigaciones relevantes para resolver problemas similares, considerando los objetivos principales del proyecto y la problemática descubierta.

5.1. Machine learning

Actualmente, la inteligencia artificial y el aprendizaje automático se utilizan ampliamente por las empresas para la toma de decisiones basadas en datos. Esto les brinda una ventaja competitiva en el mercado y les permite entender cómo se mueven los flujos de información con otros comercios y clientes. Una de las formas de lograr esto es a través del machine learning⁵.

Pero, ¿qué es machine learning? Machine learning es una disciplina que se basa en algoritmos, en la cual se dota a un determinado activo analítico la capacidad de identificar patrones en grandes volúmenes de datos, generando respuestas inteligentes dependiendo de los comportamientos descubiertos. Su base lógica es mediante el uso de herramientas matemáticas y la capacidad tecnológica para llevar a cabo estas tareas. Algunos sectores clave de aplicación del machine learning son la fabricación de productos, la investigación biológica, servicios financieros, y, entretenimiento y multimedia^H.

Los algoritmos que se utilizan en la herramienta pueden ser categorizados de la siguiente manera⁸:

- **Aprendizaje supervisado:** Algoritmos con un aprendizaje previo (entrenamiento) mediante el empleo de etiquetas asociados a los datos para generar decisiones y/o predicciones, como por ejemplo, algoritmos de regresión, bayesianos y redes neuronales.
- **Aprendizaje no supervisado:** Algoritmos sin aprendizaje previo. Se enfrentan a tareas de agrupamiento conocidas como clustering o segmentación, en la cual buscan identificar grupos similares a raíz del conjunto de datos general y su debido comportamiento, como por ejemplo, agrupamiento jerárquico y K-means.

Por otra parte, el uso de machine learning tiene tanto ventajas como desventajas. Por parte de las ventajas se tiene⁹:

- o Mayor grado de especificidad. Se logran identificar patrones complejos en los datos que el humano puede pasar por alto.
- o Mayor precisión. Al ir aprendiendo de los datos, el sistema se vuelve más preciso con el tiempo.
- o Pueden trabajar sin intervención humana una vez que están contruidos, por lo que pueden entregar constantemente información.
- o Alta capacidad tecnológica en cuanto al grado de procesamiento de datos.

Respecto a las desventajas, se puede mencionar que:

- o Se necesita una base de datos robusta para aplicar los modelos.
- o Dependiendo de la configuración de los softwares requeridos, en algunos casos, puede haber una alta inversión inicial.
- o Es necesario saber cómo aplicar e interpretar las herramientas. Grado de dificultad en el aspecto técnico.

Con el fin de automatizar el análisis con el cual la empresa podría generar una alianza, se verifica la necesidad de entender los patrones existentes en la información histórica a través de este tipo de herramientas. Es importante saber con qué información será necesario trabajar y aplicar el modelo indicado para resolver el problema.

5.2. Plataformas de inteligencia empresarial

En la misma orientación del punto 5.1, es decir, de un mundo gobernado por datos y cambiante dentro de los negocios, la ventaja de automatizar análisis y generar un control sobre la toma de decisiones en las empresas es fundamental.

Las plataformas BI o de inteligencia de negocio son herramientas que permiten agilizar estos trabajos. Estas permiten importar, limpiar y analizar datos, e incluso, hacer pronósticos con ellos⁶. En otras palabras, permite mantener y gestionar de mejor manera el flujo de algún proceso en particular de trabajo.

Su uso más común radica en, por ejemplo, análisis de ventas y marketing, análisis de clientes y segmentación, gestión de inventario, control de experimentos tecnológicos, visualización de estados de procesos, y gestión de proyectos.

Conceptualmente, los programas de inteligencia de negocios permiten englobar procedimientos que transforman la información entrante, en términos de datos, en conocimientos para la organización.

Dentro de las ventajas del uso de las plataformas de inteligencia empresarial se tiene:

- o Alto grado de sencillez en los análisis. Facilidad en la recopilación y procesamiento de datos, a la vez de que es ágil en términos del flujo de información.
- o Permite automatizar algunas tareas analíticas, disminuyendo la complejidad de procesos comerciales.
- o Al generar procesos analíticos más automatizados, eleva la productividad y el uso eficiente de los recursos disponibles y el tiempo.

Por otro lado, las desventajas de su uso radican en:

- o Pueden existir costos altos para la implementación de software. Sobre todo, cuando se generan integraciones específicas para alguna empresa. Poco asequible para empresas de menor tamaño.
- o En algunos casos, la implementación de estos sistemas puede demorar años según el grado de complejidad y uso que se requiera.

Con el fin de poder recopilar los resultados obtenidos del trabajo a realizar, en el ámbito del análisis de datos y la gestión de recursos, se observa una gran oportunidad en utilizar este tipo de herramientas para complementar una automatización en el proceso requerido. Sin embargo, será necesario identificar con qué tipo de plataforma de inteligencia empresarial se cuenta, y seleccionar la que se adapte mejor al problema, junto con aprender sobre su utilización y sus limitaciones particulares.

5.3. Maximización de beneficios asociados al análisis de campañas comerciales

Para poder aumentar el rendimiento de las campañas se han generado diferentes alternativas o metodologías como los modelos predictivos de ventas y los modelos de optimización. A continuación se analizarán ambas metodologías mencionadas.

5.3.1. Modelos predictivos de ventas

Los modelos de pronósticos de demanda o venta son modelos basados en datos históricos para estimar ventas futuras. Por ejemplo, existen los pronósticos en gestión de operaciones, en donde tenemos métodos regresivos y series temporales, alisado exponencial y medias móviles. Estos métodos ocupan como medida de desempeño principalmente el error¹.

Este error puede ser medido de diferentes formas, como por ejemplo, el error cuadrático medio (MSE), error absoluto porcentual medio (MAPE) y el error medio absoluto (MAE). Estas medidas se diferencian en la capacidad de penalizar los errores de comparación entre el valor real y pronosticado, siendo el MSE el método que refleja un mayor cambio cuando se está en presencia de alguna magnitud de error grande, dado que sus cambios son cuadráticos, como se puede apreciar en la fórmula adjunta en el anexo¹.

Cabe destacar que los modelos de regresión de machine learning también trabajan bajo algunas de estas medidas de desempeño, entrenando los datos en favor de la menor distancia entre el valor real y el pronosticado.

Los modelos predictivos se comportan como una herramienta complementaria para poder maximizar el beneficio de las campañas comerciales, proyectando lo que puede pasar en el futuro.

5.3.2. Modelos de optimización

La programación matemática u optimización es una herramienta que nos permite, a grandes rasgos, analizar cómo podemos conseguir los mejores resultados en un conjunto de múltiples soluciones para un problema. Como menciona el libro Modelos deterministas: Optimización lineal: “Sirve para encontrar la respuesta que proporciona el mejor resultado, la que logra mayores ganancias, mayor producción o felicidad o la que logra el menor costo, desperdicio o malestar”².

Las características generales que constituyen un modelo de optimización son las siguientes:

- **Función objetivo:** Medida o fórmula que se desea maximizar o minimizar. Algunos ejemplos comunes tenemos las utilidades, ganancias, costos, clientes, etc. Cualquier característica relevante en nuestro estudio a mejorar.
- **Variables de decisión:** Son las variables que generan las diferentes soluciones para el modelo (espacio de búsqueda) y que posteriormente se ajustan para encontrar la combinación específica que maximiza o minimiza la función objetivo.
- **Restricciones:** Son las limitaciones del modelo, las cuales permiten asegurar que las respuestas sean lógicas en cuanto al contexto del problema y se brinde espacio a soluciones factibles. Ejemplo de ello tenemos los presupuestos, limitaciones de recursos, etc.

Con el uso de estas variables, el proceso de optimización busca encontrar la solución que satisfaga tanto la relación de maximización o minimización como las restricciones del problema. Esta herramienta tiene una amplia variedad de campos de aplicación como la economía, logística e ingeniería.

Además, los modelos de optimización se pueden categorizar de la siguiente forma:

- **Programación lineal:** Enfoque en la cual se optimiza una función lineal sujeta a un conjunto de restricciones lineales. Este tipo de problemas involucran variables continuas y se resuelven en la búsqueda de valores óptimos de las variables objetivo a maximizar o minimizar.
- **Programación entera:** Extensión de la programación lineal en la que se requiere que algunas o todas las variables sean números enteros. Mayor grado de complejidad que un problema de programación lineal.
- **Programación no lineal:** Problemas de optimización en donde la función objetivo o las restricciones no son lineales. También suelen ser más complejos que la programación lineal y requieren de algoritmos específicos.

Ahora, dentro de las ventajas del uso de estos modelos de programación, se encuentra que:

- Estos modelos ayudan a encontrar la mejor solución en términos de costos, recursos o rendimiento, lo que se traduce en mayor eficiencia en la toma de decisiones.
- Pueden automatizar la toma de decisiones en situaciones que se repiten a lo largo del tiempo, lo que ahorra tiempo y recursos.
- Se basan en datos, lo cual proporciona decisiones con mayor precisión en vez de la suposición o plena intuición de los datos.

Por otra parte, las desventajas presentes en el uso de estos modelos son que:

- La complejidad puede ser la principal barrera a la hora de utilizar esta herramienta, por lo que se requiere los conocimientos adecuados sobre su desarrollo y la propia resolución. Además, es necesario no solo modelar y utilizarlos, sino también interpretar los resultados y aplicarlos en el contexto del problema.
- Algunos problemas pueden llegar a ser costosos en términos de requerir alto rendimiento computacional y tiempo para poder resolverse.

A través de esta investigación, se corroboró la existencia de diferentes herramientas y metodologías con el cual podemos resolver los objetivos expuestos en el proyecto y que actualmente se ocupan en la industria. Esto ha sido un repaso general de las diferentes vías existentes, sin embargo, también es necesario entender con qué herramientas contamos para la elaboración del proyecto, la capacidad técnica para aplicar estas teorías y la búsqueda de una configuración apropiada que permita resolver el problema en los plazos establecidos.

6. Solución

Para la solución se decide optar por una opción híbrida entre las herramientas y metodologías mencionadas en el punto 5, dado que cada una de ellas aportará con lo requerido para completar los objetivos del proyecto.

La solución, a la cual llamaremos “Robot BCIPlus+”, corresponde a un sistema de información del tipo “apoyo a las decisiones”. Como su nombre lo menciona, estos sistemas poseen como característica primordial la toma de decisiones a un nivel estratégico medio y alto, tomando altos volúmenes de datos brutos para su posterior procesamiento en conocimiento para la empresa.

Esta herramienta se ubicará de cara a los usuarios del equipo de BCIPlus+, principalmente para la etapa de análisis y parte de la exploración del proceso de generación de una alianza, con el objetivo de tomar decisiones con mayor conocimiento, con eficacia para responder ante los requerimientos de los objetivos, y, además, con procesos automatizados.

Esta herramienta, contará con el uso de modelos de *machine learning* y modelos de optimización para gestionar las campañas comerciales, en función de maximizar la obtención clientes MAUS para el negocio, teniendo en cuenta el impacto que genera el *cashback* en las ventas y el presupuesto determinado para dichas campañas comerciales.

El resultado esperado es obtener las recomendaciones que permitirán establecer la cartera de inversión con marcas de comercios que mejor se ajusten a un presupuesto establecido.

6.1. Matriz de riesgos

Para identificar y mitigar futuros problemas en el proyecto, se construyó una matriz de riesgos que se puede apreciar en la tabla 2 (la tabla completa junto con su desarrollo está adjunta en el punto K del anexo). En ella se incluyó las principales fuentes de riesgo del proyecto, su categoría en términos del impacto y las probabilidades de ocurrencia, junto a su respuesta de mitigación.

ID	Tipo de riesgo	Riesgo		Respuesta
		Fuente	Consecuencia	
1	Técnico	Problemas para el acceso a las principales plataformas de trabajo: Databricks, PowerBI y AWS (Caídas del sistema, problemas con las credenciales o bases de datos particulares).	No permite el avance del proyecto ni acceso a las principales fuentes de datos.	Comunicar a Scrum master y BIS (analistas de datos) del equipo de BCILabs.
2	Técnico	Intercepción de datos cruciales del banco y ciberseguridad.	Al manejar un equipo de la empresa y tener acceso a los principales datos, se es vulnerable a potenciales amenazas que puedan interferir. Podría no realizarse el proyecto a tiempo.	Se tiene que mantener una buena cultura en cuanto a la seguridad informática, uso de los equipos, herramientas de trabajo y comunicación. En caso sospechoso de actividad en los sistemas del equipo, comunicar al equipo de soporte TI BCI.
3	Calidad	El control de calidad de los datos es deficiente.	Demora en extracción de datos y tiempos de ejecución del proyecto.	Verificar en detalle la integridad de los datos antes de su procesamiento.
4	Técnico	Deterioro del equipo informático.	No se podría entrar a las principales herramientas de trabajo y comunicación con los equipos.	Se tiene que mantener el debido cuidado e integridad al equipo entregado por BCI. En caso de algún desperfecto, avisar de forma interna al equipo de BCILabs y dirigirse al soporte TI BCI para tramitar un cambio o revisión del equipo.
5	Gestión	Cambio de objetivos proyecto BCIPlus+	El proyecto se tendría que volver a construir en base a la nueva orientación del proyecto, y por ende, su retraso.	Se tiene que idear nuevas estrategias para abordar los problemas requeridos.
6	Experiencia	El equipo BCIPlus+ no trabaja al ritmo necesario.	El proyecto podría no realizarse a tiempo.	Se presionará al equipo para acelerar el proceso, tratando de emitir con tiempo los requerimientos asociados y avisando al supervisor el estado de ello.
7	Gestión	Equipo de TI BCI no realiza las aprobaciones correspondientes para la obtención de licencias de programas necesarios.	El proyecto no podría obtener las principales herramientas de trabajo para su construcción. Retraso o cambio en la metodología del proyecto.	Se tiene que idear nuevas formas de realizar el trabajo, verificar otras opciones de aplicaciones y uso de equipos externos.

Tabla 2: Extracto de la matriz de riesgos.

De forma general, los principales problemas al cual se enfrenta el proyecto tiene que ver con posibles desperfectos técnicos, uso de los sistemas y softwares de la empresa, y el flujo de intercambio de información clave para el desarrollo del proyecto y su investigación.

6.2. Evaluación económica

Para el proyecto, se realizó una evaluación económica sobre los principales costos asociados y cómo impactará en contraposición de lo que actualmente se utiliza. Por lo tanto, consideraremos una serie de información (actualizada hasta la fecha de este informe) sobre los principales requerimientos y costos del proyecto, y el costo del proceso actual que se está cuantificando.

Además, gracias a estos costos, generamos una evaluación del proyecto, considerando que el robot pueda generar mayores transacciones gracias a una mejor gestión de las campañas de BCIPPlus+. Ambas evaluaciones se observarán en las tablas 3 y 4 a continuación.

Gastos principales por mes	
Licencia Azure Databricks	\$ 10.200
Licencia PowerBi	\$ 9.500
Costos trabajador BCI promedio a nivel de rango y equipo BCIPPlus+ y BCILabs	\$ 3.900.000
Equipo (PC)	\$ 550.000
Costo promedio de practicante	\$ 500.000
Requerimientos del proceso de alianza	
4 Colaboradores BCIPPlus+ (4 roles del proceso)	\$15.600.000
4 Equipos (1 por colaborador)	\$ 2.200.000
1 Licencia Azure Databricks (Hacker de negocios BCIPPlus+)	\$ 10.200
Costo total mensual	\$17.810.200
Requerimientos del proyecto	
2 Colaboradores BCILabs en rol de Analista y Dueño de producto (Uso de un 25% de su tiempo para el proyecto)	\$ 1.950.000
1 Practicante	\$ 500.000
1 Licencia Azure Databricks (uso común)	\$ 10.200
2 Licencias PowerBi (Analista y practicante)	\$ 19.000
3 Equipos (1 por colaborador y 1 practicante, incluye tiempo de uso)	\$ 825.000
Costo total mensual	\$ 3.304.200
Ahorro esperado mensual	
Costo por etapa de análisis (1 semana)	\$ 4.155.713
Costo por etapa de análisis nueva (1 día)	\$ 593.673
Ahorro del proceso con la solución	\$ 3.562.040
Porcentaje de ahorro sobre el proceso completo	20%

Tabla 3: Evaluación económica considerando los principales costos asociados al proyecto en pesos chilenos.

Evaluación del proyecto BCIPPlus+			
Ingresos principales	Mes 1	Mes 2	Mes 3
Cientes MAUS BCIPPlus+	18768	21468	24473
Cientes MAUS en campañas de costo	3739	2969	4895
Cientes MAUS esperado por proyecto	27607	23258	33822
Cientes MAUS adicionales	23868	20289	28927
Transacciones en BCIPPlus+	84204	86772	113198
Transacciones en campañas de costo	21208	16494	32804
Transacciones ingreso	62996	70278	80394
Cantidad de transacciones promedio por cliente	3,356564365	3,273616546	3,285007968
Transacciones adicionales por proyecto	80114,47826	66418,40609	95025,42549
Valor promedio de ingreso por transacción	\$ 480	\$ 480	\$ 480
Flujo de ingreso del proyecto	\$ 38.454.950	\$ 31.880.835	\$ 45.612.204
Gastos principales	Mes 1	Mes 2	Mes 3
2 Colaboradores BCILabs (Uso de un 25% de su tiempo para el proyecto)	\$ 1.950.000	\$ 1.950.000	\$ 1.950.000
1 Practicante	\$ 3.900.000	\$ 3.900.000	\$ 3.900.000
1 Licencia Azure Databricks (uso común)	\$ 10.200	\$ 10.200	\$ 10.200
2 Licencias PowerBi (Analista y practicante)	\$ 19.000	\$ 19.000	\$ 19.000
3 Equipos (1 por colaborador y 1 practicante, incluye tiempo de uso)	\$ 825.000	\$ 825.000	\$ 825.000
Costo mensual del proyecto	\$ 6.704.200	\$ 6.704.200	\$ 6.704.200
Flujo Neto del proyecto	\$ 31.750.750	\$ 25.176.635	\$ 38.908.004
Inversión inicial (2 meses de desarrollo para investigación)	\$ 13.408.400		
Tasa de descuento	20%		
VAN	\$ 53.050.538		
TIR	223%		

Tabla 4: Evaluación económica del proyecto en pesos chilenos.

En la tabla 3 se puede observar que los principales gastos asociados al proyecto son las licencias para los softwares de trabajo y el costo de los colaboradores de BCIPPlus+ y BCILabs. En el caso de BCI, el costo asociado a un colaborador no incluye el costo del equipo (PC), esto fue conversado con las personas de la empresa, entendiendo que, al existir diferentes rangos, en términos jerárquicos de la empresa, los costos por colaborador pueden variar, pero el equipo posee un costo medio aproximado.

Además, dado que se espera reducir el tiempo del proceso de análisis de 1 semana a 1 día, verificamos que el ahorro que se genera gracias al robot es de aproximadamente 3,5 millones de pesos, y, respecto al proceso completo, del 20% del gasto total.

Finalmente, se genera una evaluación del proyecto mediante los gastos antes considerados, junto con el flujo de ingreso que generaría el robot, la inversión inicial, la cual corresponde a 2 meses de investigación, y la tasa de descuento declarada por la empresa

para la generación de sus proyectos. Con ello obtuvimos el VAN^L y la TIR^M asociada al proyecto, observando una gran viabilidad para el proyecto. Esto se debe, principalmente, a que se espera que el robot mejore las decisiones sobre las campañas, generando una mayor cantidad de transacciones de las que actualmente se logran, sobre todo con las campañas de costos, las cuales implican una inversión por parte del banco. Esta característica de las campañas podremos verlas más adelante en el punto 6.6.3 del informe sobre el desarrollo del proyecto o en el punto X del anexo, donde se explica la capacidad de pago en cuanto a comisiones de los comercios (*fee*).

6.3. Metodología

El trabajo se va a dividir en diferentes capas de desarrollo. Primero, contemplando la conexión a los datos y la plataforma *PowerBi*, luego, se aplicará minería de datos para entender los patrones de comportamiento del *cashback* en las marcas comerciales, y finalmente, con la información estudiada y los coeficientes de impacto asociados, se construirá un modelo de optimización para la captura de clientes MAUS en función de los costos y un presupuesto establecido.

Para mayor entendimiento, se construyeron algunos flujos de referencia sobre la conexión del trabajo tanto a nivel informático como industrial que pueden apreciar en las ilustraciones 2 y 3.

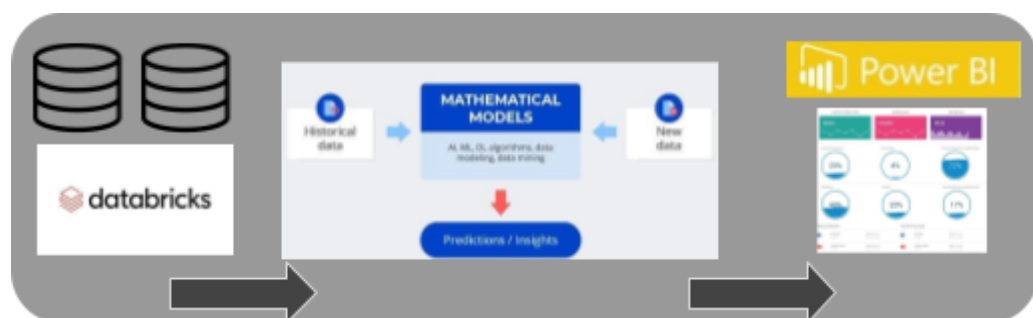


Ilustración 2: Flujo de trabajo informático.



Ilustración 3: Flujo de trabajo industrial.

Entonces, inicialmente se construirá un prototipo básico del sistema que queremos implementar, el cual, con el tiempo, recibirá las mejoras correspondientes tanto a nivel funcional como estético para el usuario final.

Para la conexión de los datos con la plataforma y la arquitectura básica del sistema, conectaremos los datos de las principales fuentes de información del banco, alojadas en el sistema de *Azure Databricks*. Esta plataforma alojada en la nube del banco BCI, nos permite navegar por las principales fuentes de datos, generar *scripts* para la automatización analítica y conectar dichos datos o procesos para su visualización en *PowerBi*.

Posteriormente, aplicando la metodología clásica de minería de datos *CRISP-DM* (*Cross-Industry Standard Process for Data Mining*)^N se investigará el comportamiento del *cashback*, utilizando para ello, la información de las marcas que han tenido ventas con *cashback* o están inscritas en el programa, verificando su rendimiento y generalizar su comportamiento según el tamaño que estas poseen.

El tamaño de estas marcas comerciales, dentro del contexto de la empresa, están divididas según su cantidad de transacciones promedio mensuales, por lo que el objetivo será identificar cuánto crecen las ventas de las marcas comerciales con el *cashback*, visualizando con ello su impacto (utilizando tanto las empresas con buen y mal rendimiento para no sesgar la información).

También será necesario realizar un proceso de limpieza de datos, identificando posibles datos anómalos, valores nulos, incoherencias, entre otros, para consecutivamente, procesar los datos en un modelo de *machine learning* y generar los principales experimentos de impacto que brindarán las respuestas del comportamiento del *cashback*.

El modelo a utilizar para este proyecto será sobre modelos de regresión ^o, por lo que sería asociado hacia al tipo de aprendizaje supervisado. Esto nos ayudará a entender cómo impacta el *cashback* en las diferentes marcas de comercio observando sus características significativas. Esto se explica en mayor detalle en el punto 6.6.2.

Además, utilizaremos *Python* como principal lenguaje de programación, dado que es un lenguaje simple, flexible y está enfocado principalmente al área de análisis y ciencia de datos. Por otra parte, este lenguaje tiene la flexibilidad para la automatización de tareas que es lo que estamos buscando para este proyecto.

Posteriormente, entendiendo el impacto del *cashback* como variable independiente, hacia las ventas como variable dependiente, se analizará los comercios que mayores beneficios y oportunidades traen al banco. Será necesario determinar qué variables van a ser las determinantes para estos beneficios, generando una función de maximización que permita responder ante esta problemática.

A grandes rasgos, dentro de la metodología de trabajo, la parte informática será la herramienta que provee el manejo de los datos, mientras que la parte industrial será quien tomará esta información para convertirlos en las fuentes de oportunidades para gestionar de mejor forma las campañas comerciales.

Finalmente, el resultado esperado es obtener un producto el cual desplegará en detalle las características de los diferentes comercios, sus indicadores principales, gráficos relevantes para la medición, y lo más importante, las recomendaciones para la gestión de las campañas de las diferentes marcas comerciales.

Cabe mencionar que *PowerBi* es la herramienta principal que utiliza el banco para generar esta visualización, y con la cual se tiene permiso de uso dentro del banco. La idea es que todas estas herramientas comentadas estén conectadas entre sí, con tal de que se pueda actualizar los datos de forma constante y finalmente impactar en el proceso de alianzas para BCIPlus+.

6.4. Plan de implementación

Para la construcción de este proyecto tendremos en consideración las siguientes etapas de desarrollo, cada una con su marco de tiempo estipulado y la temática e investigación a tratar, que podrán visualizar en la carta gantt de la ilustración 4.

6.4.1. Etapa 1: Nivel bruto (agosto)

En el nivel bruto buscaremos generar un prototipo básico del Robot BCIPPlus+, el cual se puede visualizar la conexión principal a los datos, se pueda generar una búsqueda general de los comercios, generar algunas proyecciones básicas en las transacciones de los comercios, y determinar indicadores de rendimientos sobre alguna campaña en particular (clientes activos, comisiones, porcentaje de ganancia para el banco, etc.). Esto debe estar conectado a las fuentes de datos, códigos de programación y al *PowerBi* de cara al usuario de BCIPPlus+.

6.4.2. Etapa 2: Análisis de sensibilidad (septiembre)

Esta etapa se orienta a los principales avances sobre los análisis del impacto del *cashback* y su comportamiento en las ventas. Para ello, se tendrá que hacer procesos de selección de datos transaccionales, limpieza de datos y procesar la información en los modelos de *machine learning*. Con ello se tendrán que generar su debido ajuste al modelo para descartar variables no significativas y generar curvas de proyección según variables relativas al modelo.

6.4.3. Etapa 3: Coeficientes de impacto y modelos de optimización (octubre-mediados de noviembre)

En esta etapa, tomando los análisis de impacto del *cashback*, se extraerán los coeficientes de las marcas comerciales respectivas segmentadas por tamaño. Esto nos permitirá incluirlas en el modelo de optimización que se construirá en esta etapa. Con las restricciones propias del modelo y un presupuesto base por tamaño, se buscará generar la mejor recomendación para las marcas existentes y obtener una respuesta certera sobre cómo invertir en las campañas comerciales con tal de maximizar la tasa de captura de clientes MAUS, siendo este su beneficio.

6.4.4. Etapa 4 Pruebas finales y cierre de proyecto (mediados de noviembre y diciembre)

Finalmente, para cerrar el proyecto se irá midiendo el comportamiento de los clientes con las nuevas recomendaciones, automatizadas y mejoradas, para obtener el resultado del proyecto. Se generarán las pruebas finales de implementación del robot.

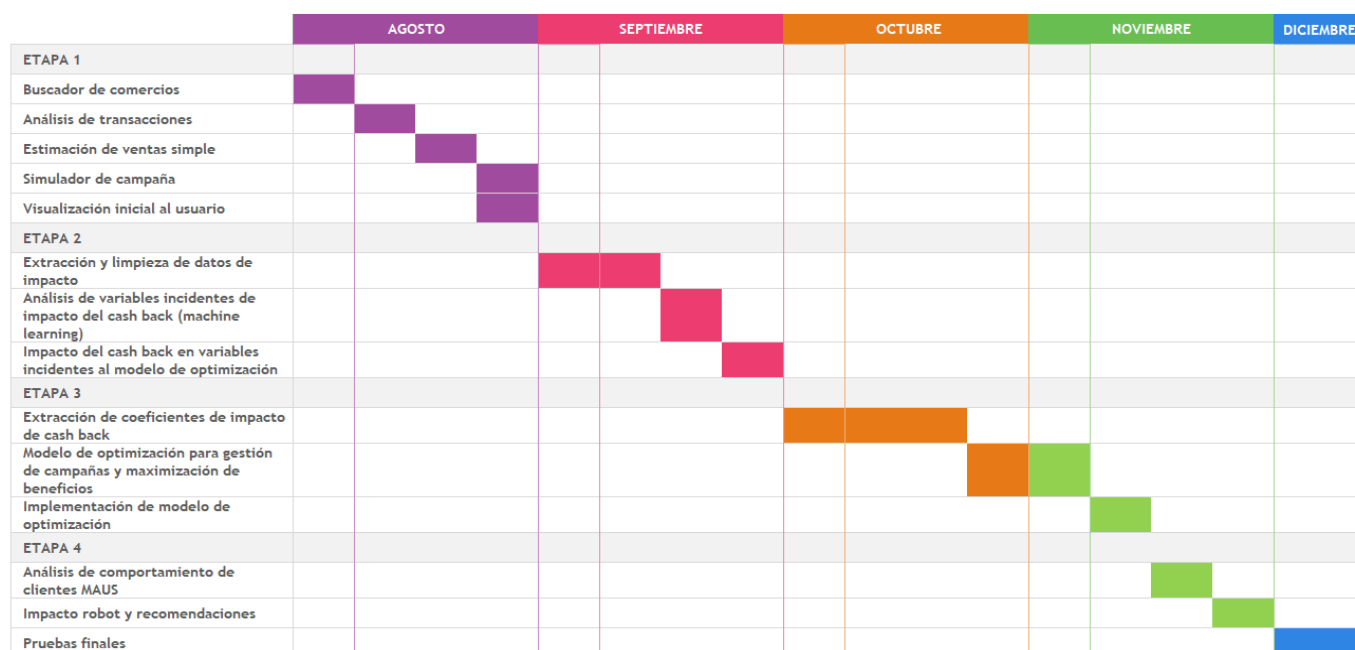


Ilustración 4: Carta Gantt del proyecto BCIPPlus+.

6.5. Medidas de desempeño

Para poder evaluar el desempeño del proyecto, de acuerdo con nuestros objetivos y metodologías, tendremos 2 principales indicadores: tiempo y eficacia.

6.5.1. Tiempo

El tiempo se refiere a cuánto se demora generar una recomendación sobre una alianza y decisiones de entrega de ofertas. Principalmente para la etapa de análisis y parte de la exploración para el proceso de las campañas BCIPPlus+.

Tenemos que considerar que el tiempo actual de esta etapa es de 1 semana, mientras que el esperado es de 1 día. La razón por la cual hacemos evaluación de esta métrica es para ahorrar tiempo y dinero que se puede utilizar para la gestión de campañas comerciales y búsqueda de oportunidades. Esta medida es propia de la automatización.

6.5.2. Eficacia

La eficacia es la medida porcentual de mejora respecto a la obtención de *MAUS* en las campañas realizadas. Se calculará comparando entre lo registrado en el sistema antiguo versus el robot. Se espera que esta medida de captura aumente en un 30%.

6.6. Desarrollo del proyecto

A continuación, se presentará el desarrollo de las diferentes etapas del proyecto, desde la etapa 1 hasta la etapa 3 de acuerdo al plan de implementación explicado en el punto 6.4. La etapa 4 se expondrá en las conclusiones finales.

6.6.1. Etapa 1: Nivel bruto

Para esta etapa se trabajó principalmente en generar los códigos de programación tanto en *SQL* como *Python* para la extracción de datos de los comercios y cálculos de las métricas de desempeño de las campañas. Además, se creó el primer diseño para el panel de *PowerBi*, en donde se desplegará la información histórica de los comercios a través de diferentes filtros.

Cabe destacar que para esta tarea se tuvo que aprender sobre cómo enlazar el panel al flujo de datos de la nube en *Databricks*, en donde, con los permisos necesarios propios del rol de BIS, se enlaza el sistema a la nube de innovación de BCILabs. Además, se tuvo que entender el funcionamiento propio del software de inteligencia empresarial *PowerBi*, las métricas asociadas a las campañas de cashback en BCIPlus+, trabajar activamente junto a los BIS del equipo de BCILabs, proponer las ideas del proyecto y enlazar los conocimientos propios de la universidad.

Los datos de BCI tienen sus principales repositorios en las nubes de Data & Analytics de la empresa. BCILabs trabaja con estos datos solo a nivel de visualización, por lo que no puede crear tablas de información ni generar cambios sobre estos. Sin embargo, BCILabs posee su propia nube de innovación, por lo que para este proyecto, fue necesario filtrar los datos asociados a los comercios en las diferentes bases de Data & analytics, descargar una copia de la información, y, posteriormente, generar una migración hacia la nube de innovación de BCILabs donde sí se poseen los permisos necesarios para la modificación.

Por otro lado, dentro del panel, se trabajó en una herramienta de simulación para las campañas de BCIPlus+. Esto permitirá al usuario del área evaluar una posible campaña con los rendimientos actuales de un comercio en particular. En esta simulación se poseen 3 controles esenciales: porcentaje de *cashback* al cliente, comisión de comercio y la duración de la campaña.

Con estos controles, el sistema calcula un impacto esperado a través de los MAUS, transacciones esperadas, GMV, monto promedio de venta y una comisión

resultante de la campaña. Estas métricas son reflejadas a través de las fórmulas adjuntas en el anexo en el punto P.

Luego, se empiezan las tareas para poder estimar el periodo futuro de los comercios en términos de los clientes distintos y la cantidad de transacciones, junto con el impacto del cashback que se realizará en la etapa 2. Los periodos sobre los comercios fueron hechos, por petición del equipo y la empresa, en tiempos mensuales, dado que el proceso de alianza gira en torno a este tiempo.

Con las metodologías mencionadas en el punto 5.3.1 efectuamos la estimación de los clientes distintos y la cantidad de transacciones en un periodo futuro (1 mes) para la base de comercios existentes, y, además, conectando esta información a *PowerBi*.

El periodo futuro lo estimamos con medias móviles, suavizado exponencial simple y regresión lineal. Observamos su rendimiento predictivo para las variables de clientes distintos y cantidad de transacciones, considerando cierto nivel de estacionalidad y el comportamiento de los datos. La razón por la que utilizamos estos modelos la pueden encontrar en el anexo en el punto Q.

Con las estimaciones ejecutadas sobre cada comercio, los filtros asociados y la sintaxis DAX^R propias del sistema de *PowerBi* es que se pudo construir las herramientas propias del panel que pueden visualizar en la ilustración 5.

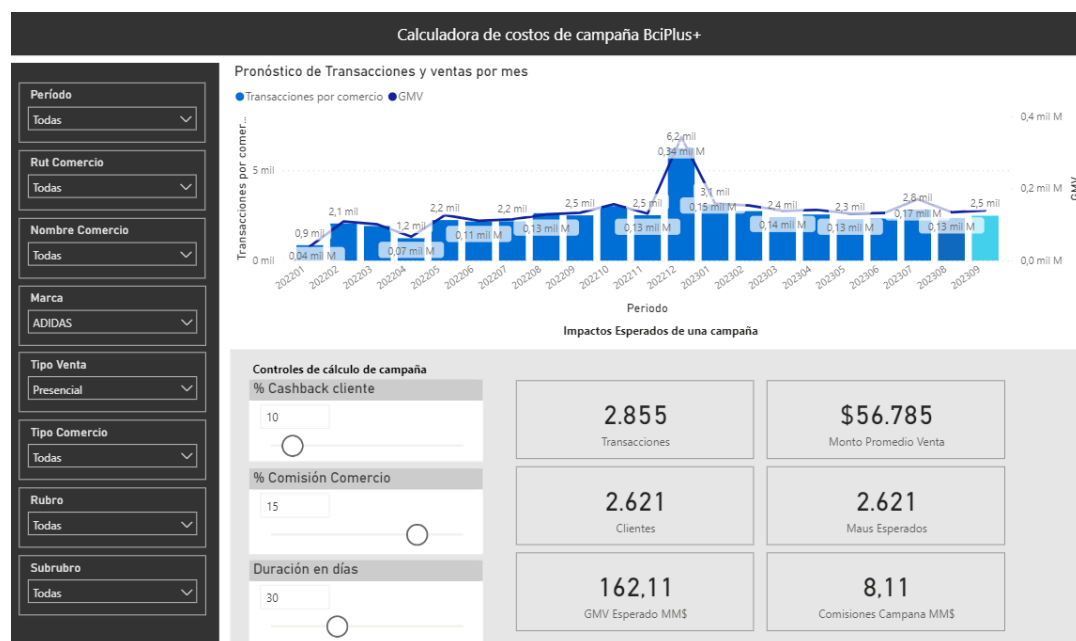


Ilustración 5: Panel del robot BciPlus+ etapa 1.

Finalmente, se construyó una segunda visualización tipo tabla, en donde se podrá ver el impacto sobre las sucursales de algún comercio en concreto. Esto se hizo para obtener una visión panorámica y enfocar los esfuerzos comerciales en los nichos de mayor rendimiento para que el usuario pueda comparar. Este ejemplo lo podemos ver en la ilustración 6.

Rut Comercio	Glosa Comercio	Transacciones	Clientes	Monto Promedio	GMV MM\$	Comisiones MM\$
99584660	ADIDAS SANTIAGO CL	9.398	8.179	\$72.448	680,86	34,043
78744360	ADIDAS MBFO COQUICOQUIMBO	4.276	3.813	\$44.698	191,14	9,557
78744360	ADIDAS NUNOA SANTIAGO	4.212	3.737	\$44.585	187,80	9,390
78744360	ADIDAS LA FABRICASANTIAGO	4.003	3.452	\$43.702	174,93	8,747
78744360	ADIDAS MAIPU SANTIAGO	3.708	3.381	\$46.116	170,98	8,549
78744360	ADIDAS ALTO LAS CSANTIAGO	3.680	3.440	\$60.393	222,24	11,112
78744360	ADIDAS MALL PLAZASANTIAGO	3.357	3.165	\$60.877	204,36	10,218
78744360	ADIDAS MARINA ARAVINA DEL MAR	3.125	2.913	\$59.383	185,55	9,277
78744360	ADIDAS PARQUE ARASANTIAGO	2.595	2.481	\$60.926	158,09	7,904
78744360	ADIDAS BCS PLAZA SANTIAGO	2.231	2.085	\$52.627	117,39	5,870
78744360	ADIDAS ORIGINALS SANTIAGO	2.029	1.957	\$59.847	121,41	6,070
78744360	ADIDAS OUTLET VIVTEMUCO	2.018	1.811	\$40.562	81,84	4,092
78744360	ADIDAS ARAUCO MAISANTIAGO	1.948	1.859	\$49.001	95,43	4,772
78744360	ADIDAS FLORIDA CESANTIAGO	1.862	1.800	\$48.552	90,39	4,519
78744360	ADIDAS OCS COSTANSANTIAGO	1.640	1.583	\$66.740	109,47	5,474
78744360	ADIDAS PORTAL NUNSANTIAGO	1.614	1.455	\$38.530	62,18	3,109
78744360	ADIDAS PLAZA OESTSANTIAGO	1.235	1.192	\$46.847	57,86	2,893
78744360	ADIDAS MARINA ARAUCO VINA DEL MAR CL	1.144	1.084	\$75.453	86,35	4,317
76832117	ADIDAS PLAZA EGANSANTIAGO	1.140	1.073	\$50.487	57,53	2,877
76832117	ADIDAS IQUIQUE IQUIQUE	1.118	1.062	\$58.325	65,19	3,260
78744360	ADIDAS SANTIAGO	976	895	\$44.525	43,47	2,173
76832117	ADIDAS TALCA	888	827	\$61.706	54,79	2,739
76832117	ADIDAS CALAMA BCSCALAMA	861	812	\$73.003	62,88	3,144
76735222	ADIDAS VALDIVIA	686	653	\$52.758	36,19	1,809
78744360	ADIDAS KIDS PLAZASANTIAGO	670	650	\$47.729	31,99	1,599
76735222	ADIDAS LOS ANGELES	647	612	\$51.722	33,47	1,674
76832117	ADIDAS COPIAPO COPIAPO	645	608	\$58.403	37,66	1,883
78744360	ADIDAS PLAZA EGANSANTIAGO	604	569	\$50.040	30,21	1,510
77581040	ADIDAS PUCON	500	465	\$54.363	27,16	1,358

Ilustración 6: Visualización del impacto en las sucursales de la marca Adidas.

A medida que se fue desarrollando esta primera fase del proyecto, se validó su información con los usuarios de BCIPPlus+, quienes nos brindaban retroalimentaciones, posibles cambios o mejoras. Además, con el apoyo de los BIS del área, se corroboró la integridad de la información y la lógica implementada.

6.6.2. Etapa 2: Análisis de sensibilidad

Durante esta etapa se generó la principal base de datos para las metodologías de *machine learning*, con el fin de obtener el parámetro de impacto de *cashback* de las fórmulas expuestas en la etapa 1 y entender cómo influye el *cashback* en el comportamiento de los diferentes comercios.

Para ello, se conformó una base con la mayor de cantidad de transacciones del banco en donde se incluyeron las marcas que hayan tenido ventas con y sin *cashback* que estuvieran dentro del programa de BCIPPlus+. Esta base fue generada a

través del sistema de *Databricks*, recopilando la información asociada a nuestro problema y realizando la limpieza de datos respectiva.

Posteriormente, las variables categóricas de la base fueron transformadas mediante la metodología *one hot encoding*⁵, dado que para el procesamiento de datos de machine learning era necesario tener variables numéricas.

El modelo que utilizaremos será el modelo regresivo debido a la naturaleza de los datos y el objetivo en el cual se requiere cuantificar el impacto específico de la variable *cashback* en las ventas. Los modelos regresivos y el porqué de su uso, para complementar, lo podemos encontrar en el anexo en el punto T.

La tabla de datos a investigar se conforma por las marcas comerciales junto con sus transacciones por mes dentro del año 2023, el porcentaje de *cashback* promedio, las columnas transformadas por *one hot encoding* equivalente a cada mes, el código de rubro correspondiente obtenido en las bases de datos del banco y el tamaño de la marca. Además, se incluyó una serie de variables categóricas de interacción entre el *cashback*, los meses, los rubros y los tamaños de las marcas comerciales, con tal de verificar si dichas variables se potencian una con la otra.

Con la tabla de datos conformada, se observó el panorama general de los datos, como primer experimento, calibrando cada coeficiente con mayor significancia y con diferentes grupos de variables. Cada uno de estos modelos, se observó su rendimiento en términos del error absoluto promedio y el error cuadrático promedio, junto con su coeficiente de determinación R^2 , el cual explica la variabilidad de los datos. Sin embargo, se observó que a nivel general, con esta agrupación de variables en los modelos, existe una tendencia a errores bastantes grandes que los esperados, por lo que tuvo que recurrirse a un segundo experimento que permitiera mitigar estos errores.

El segundo experimento consistió en generar un modelo regresivo específico por cada una de las marcas existentes dentro de la tabla para posteriormente segmentarlas según su tamaño. Este tamaño de marca del comercio es una medida entregada por el banco, la cual cataloga las marcas comerciales por la cantidad promedio de transacciones mensuales. En conversaciones con los BIS del equipo de BCILabs acordamos verificar el comportamiento del *cashback* mediante estos tamaños, dado los rendimientos del primer experimento.

Consecutivamente, se generaron los modelos icono de cada marca con sus coeficientes significativos, registrando con ello las mismas medidas de desempeño del experimento 1 en una tabla *Excel*, para poder comparar con el segundo experimento. Al generar los modelos respectivos por marca se comprueba el promedio de sus medidas de desempeño. La comparación entre los experimentos lo pueden ver en la tabla 5.

Rendimiento del modelo completo (antiguo)	
Coeficiente de determinación R2	0,7496
Error cuadrático medio (MSE)	220204083
Error absoluto medio (MAE)	6946
Rendimiento del modelo por marca (nuevo)	
Coeficiente de determinación R2	0,61746
Error cuadrático medio (MSE)	936904
Error absoluto medio (MAE)	283,45

Tabla 5: Comparación del rendimiento entre los experimentos.

Con ello, se verifica que el comportamiento a nivel específico por marca logró obtener menor grado de error en la calidad de los datos, sin embargo, disminuyó un 10% el coeficiente de determinación respecto al anterior, lo cual nos refleja que el modelo se ajusta en menor calidad a los valores reales.

Finalmente, entendiendo los resultados de estos procedimientos, se optó por los modelos por marca, gracias a que nuestro objetivo de disminuir los errores era más notable en este último, a pesar de la penalización en el coeficiente de determinación.

Con los modelos creados por marca, y sus coeficientes respectivos, se simulaban diferentes porcentajes de *cashback* para obtener la cantidad de ventas gracias a este beneficio, logrando diferenciarlo de las ventas normales. El *cashback* máximo entregado por el banco es del 40%, lo cual se respetó en este experimento.

Logrando determinar la cantidad de ventas con y sin *cashback* de las diferentes marcas de comercio, se procedió a determinar su impacto dividiendo la cantidad de ventas con *cashback* y ventas normales de cada porcentaje de simulación. Con ello se obtuvo el impacto de cada una de las marcas y se procedió a

segmentar por su tamaño, promediando los impactos respectivos. Gracias a ello, se logró obtener una curva tipo representativa para cada tamaño de marca de comercio. Se adjunta un ejemplo de marca mediana en el gráfico 3, las demás tablas se adjuntan en el anexo en el punto U.

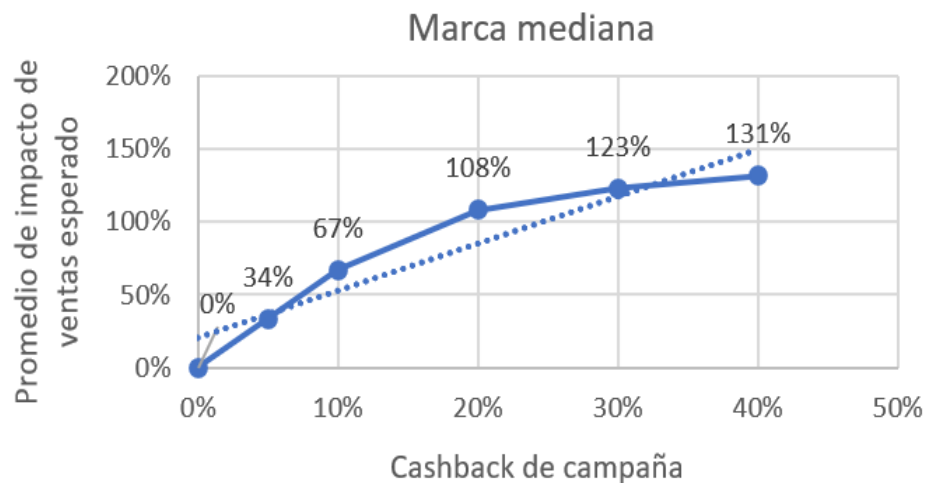


Gráfico 3: Promedio del impacto de ventas con cashback esperadas en una marca mediana.

En el gráfico se puede observar la cantidad de ventas que se logra aumentar con diferentes campañas de porcentaje de *cashback*, obteniendo así, el comportamiento de las marcas comerciales frente a diferentes porcentajes de *cashback*.

Gracias a estos experimentos y el estudio de los datos, se pudo entender que a medida que el tamaño de marca es mayor, dentro del contexto de desempeño del banco, el impacto del *cashback* sobre los comercios es menor. Estos resultados nos permiten hacer un leve pronóstico de rendimiento sobre las ventas y observar cómo se comportan los comercios frente al *cashback* según su tamaño.

Finalmente, se trabajó en mejorar el diseño del robot en *PowerBi* para obtener un evolutivo del producto y dar forma a los principales componentes de este producto a generar para BCIPPlus+. Esto es posible verlo en la imagen adjunta en el anexo en el punto V.

6.6.3. Etapa 3: Coeficientes de impacto y modelos de optimización

Para la última etapa de desarrollo se trabajó en la creación de los modelos de optimización para gestionar las campañas de *cashback*, utilizando para ello los coeficientes de impacto asociados a los diferentes tamaños de marcas de comercios.

Para el modelo de optimización se ocuparon las ecuaciones presentadas en la etapa 1 de desarrollo de proyecto. Nuestro objetivo es maximizar la captura de MAUS con la conformación de campañas a través de un presupuesto inicial establecido.

Lo primero que se tuvo que entender es que el costo de generar las campañas se genera cuando el comercio no da una comisión suficiente para abarcar el *cashback* entregado al usuario, financiando el banco lo que falta.

Para disminuir la complejidad del problema y enfocarnos en las marcas comerciales más importantes, se procede a capturar la información de las ventas de las 10 mejores marcas que mayor cantidad de usuarios activos poseían para el periodo futuro y con ellas generar una cartera de inversión óptima con el modelo. Estas decisiones se toman por periodos mensuales, de esta forma el usuario de BCIPPlus+, de acuerdo al transcurso de los meses, verifica el rendimiento de las marcas comerciales y puede proponer campañas de *cashback* en función de obtener la mayor cantidad de MAUS.

Se generó un modelo de optimización que pueden revisar en el punto W del anexo. Este modelo nos permite obtener cuales deben ser las campañas de *cashback* a entregar por marca, e incluso, el gasto a incurrir por cada una, dentro del grupo de los que mejor desempeño se espera que tengan en el mes futuro. El fee de comercio, en el contexto del modelo, es la variable que se ajusta a la disposición máxima que entregarán los comercios para generar las campañas esperadas.

Cabe destacar que este modelo se replica para cada tamaño de los comercios, haciendo que el usuario decida por cada tamaño cuánto desea invertir y cuánto esperan que los comercios entreguen para las campañas como variables de entrada.

Para procesar los resultados, se investigó cómo generar estos modelos mediante código *Python*, sin embargo, las librerías asociadas a estos modelos son poco flexibles en términos de las características del modelo, no logrando abarcar

nuestro problema específico para el banco. Por lo que se optó por procesar los resultados mediante el solucionador de *Excel*, ajustando las variables y restricciones respectivas. Esto posteriormente se automatizó a través de las macros.

Ahora, en el ámbito de la visualización en *PowerBi* se conversó, junto al equipo de BCIPlus+, cómo podrían distribuirse los presupuesto y el fee de comercio ^x, dado que en el panel del software *PowerBi* solo es posible visualizar los datos que uno le entregue. La solución a ello, fue generar distintos escenarios posibles de inversión, con diferentes tamaños y presupuestos como filtros dentro del panel.

Para finalizar esta etapa, teniendo en cuenta estos procedimientos realizados, también se mejoró el diseño de forma general, como parte de un evolutivo del robot. En esta etapa no solo era importante obtener los resultados, sino que completar la implementación completa del robot, dado que se expondría hacia el resto de las áreas del ecosistema digital BCI.

Por ello, se creó una cápsula informativa, en la que se demostrarían las funcionalidades completas del robot. La gestión de las campañas se ubicaría dentro de una pestaña *Campaign manager* del panel. Estos detalles del panel final y su descripción general es posible verlo en el punto Y del anexo.

7. Conclusiones

7.1. Resultados cualitativos y cuantitativos

A continuación exponaremos los resultados del proyecto a través de la etapa 4 considerada para el robot.

7.2. Etapa 4: Pruebas finales y cierre del proyecto

Una vez finalizado el robot, se pasa a medir el rendimiento de este. Sin embargo, a pesar de que la implementación fue concluida, no es posible ver de forma inmediata los resultados de la nueva gestión, ya que los tiempos, para los cuales se considera la realización del proyecto y el tiempo propio de demora para la conformación de campañas y alianzas comerciales, sobrepasa los límites estipulados. Entonces, para poder identificar los resultados del proyecto, nos basamos en contraponer los escenarios de decisión antiguos frente a como lo hubiera hecho el robot.

Para este trabajo, era necesario posicionar temporalmente al robot con las transacciones de los períodos antiguos y procesar el modelo de optimización, verificando la

captura de MAUS e incluyendo el impacto del *cashback* de la etapa 2 para los tamaños de las empresas. En este caso, consideramos un horizonte de tiempo de 3 meses (agosto, septiembre y octubre del año 2023), tiempo en el que se desarrolló el robot.

Gracias al equipo de BCIPlus+ se pudo obtener los presupuesto estimados utilizados en esos meses para las campañas con los diferentes tamaños de comercios. Además, consideramos las campañas de costos, que es el escenario que el robot optimiza explicado en la etapa 3.

Los resultados se observan a continuación:

Rendimiento general de BCIPlus+			
Mes	Cantidad de transacciones	MAUS	GMV
Agosto	84204	18768	\$1.475.422.488
Septiembre	86772	21468	\$1.775.441.892
Octubre	113198	24473	\$3.020.009.442
Rendimiento BCIPlus+ Campañas de costo			
Mes	Cantidad de transacciones	MAUS	GMV
Agosto	21208	3739	\$ 264.697.048
Septiembre	16494	2969	\$ 223.543.182
Octubre	32804	4895	\$ 442.394.744
Rendimiento Robot BCIPlus+ Campañas de costo			
Mes	Cantidad de transacciones	MAUS	GMV
Agosto	42439	27607	\$1.258.124.900
Septiembre	35235	23258	\$1.157.429.380
Octubre	49489	33822	\$1.556.821.430

Tabla 6: Comparación de rendimiento entre las acciones de campañas antiguas y el rendimiento esperado por el robot.

Para estos meses se puede verificar la cantidad de clientes activos (MAUS), la cantidad de transacciones y el GMV conseguido gracias a las campañas. Esto considerando los escenarios de campañas de costo, el escenario general y el escenario óptimo del robot, validando así las métricas de desempeño esperadas para este proyecto.

Respecto a los rendimientos generales, en contraposición del robot, tenemos un aumento promedio en esos meses de un 31,21% de clientes MAUS. Y comparando sólo los escenarios de costo, que son los que implican inversión del banco, se observa un aumento radical en la captura de clientes MAUS con una diferencia promedio de 24 mil clientes aproximadamente.

Además, en términos del tiempo a utilizar para el análisis del proceso de alianza y oportunidades de campaña, gracias a esta automatización que antes les demoraba 1 semana, probando junto al equipo de BCIPlus+, ahora es posible proponer opciones de campañas o alianzas en menos de un día (aproximadamente en menos de 3 horas), gracias a que la información de los comercios y su debida gestión están presentes en una sola herramienta.

Si bien estos resultados son alentadores, ¿a qué se debe este aumento radical en las transacciones, MAUS Y GVM? Observando las decisiones del robot, se pudo dar cuenta que al conformar la cartera de inversiones, los costos y las campañas que este propone invertir en cada marca con el modelo de optimización, el robot, diversifica el presupuesto en otros comercios potenciales para el banco, no concentrándose en sólo unos pocos como se estaba realizando hasta la fecha. Esto logra abarcar una mayor posibilidad de captura de clientes. A modo de ejemplo, en el apartado Z del anexo, es posible observar las decisiones que se tomaron en el mes de octubre para obtener los resultados expuestos anteriormente.

Por otra parte, el equipo de BCIPlus+ gracias a los resultados y las oportunidades que el robot les está entregando, ya se encuentran tramitando algunas campañas de alianza junto a París por su buen desempeño. El robot logra recomendar campañas que el usuario de BCIPlus+ no había considerado hasta la fecha, lo que es un buen indicador en términos de explorar nuevas posibles alianzas.

7.3. Comentarios finales

El proyecto desarrollado dentro del banco de BCI ha sido un camino de arduo aprendizaje, trabajo en equipo y con una comunicación excepcional. Trabajar dentro del contexto del análisis de datos requiere la capacidad de entender el negocio, buscando los enfoques necesarios para resolver cada tipo de problema a la cual se está expuesto.

Muchos de los procedimientos requeridos llevan múltiples verificaciones, teniendo incluso que volver a armar los modelos de datos necesarios desde cero para validar las propuestas de alguna solución. El desarrollo y los resultados del proyecto fueron validados junto a los equipos de BCIPlus+ y BCILabs, con tal de ir a la par con los requerimientos asociados y la comprobación correcta del uso de los datos.

A pesar de ser un banco con múltiples áreas y proyectos en constante desarrollo, la realidad es que son muy limitadas las herramientas con las que personalmente como practicante se posee, sobre todo si se habla desde el punto de vista de una automatización

importante como la presente durante este proyecto. *PowerBi* y otras herramientas presentadas para la creación de nuestro proyecto, fueron las herramientas disponibles durante la práctica, sin embargo, no eran necesariamente las mejores opciones que el mercado nos podría ofrecer. En ese sentido, se considera esto como sugerencias de mejoras que se podrían llevar a cabo en el futuro de este producto construido, y que fue debatido incluso con las personas de equipo.

Finalmente, el conocimiento y las aptitudes técnicas dentro del proyecto fueron importantes para llevar a cabo el desarrollo en cuestión. Poder evidenciar el manejo de los datos de la mano de profesionales del área, su estudio y su implicancia en contextos reales de negocios ayudaron a nutrir las capacidades del proyecto.

Ahora, desde el punto de vista de las áreas ingenieriles ocupadas para el proyecto, podemos verificar la importancia de las áreas de informática e industrial gracias a sus aportes respectivos dentro del trabajo. La gestión de los recursos, ya sea en una fábrica o en un banco, es necesaria para tomar decisiones estratégicas inteligentes frente a un mercado en constante cambio y en el cual es relevante monitorear. Así también las ventajas de la implementación de nuevas tecnologías de apoyo que nos permiten agilizar el trabajo, obtener una mayor precisión, y a complementar nuestras decisiones, esto ad-hoc a un contexto de alto volumen de información concebido dentro del mundo empresarial.

8. Bibliografía

1. Von Chrismar Carvajal, A. M. O. y. E. (27 de febrero de 2023). *MEMORIA INTEGRADA 2022*.
https://bci-cdn.azureedge.net/uploads/a6c0fcc7-b48f-43c2-838e-e89943afa410/original/MI_BCI_2022_compressed.pdf
2. *Modelos deterministas: Optimización lineal*. (1996, 25 febrero).
<https://home.ubalt.edu/ntsbarsh/business-stat/opre/spanishd.htm#:~:text=Los%20problemas%20de%20optimizaci%C3%B3n%20generalmente,con%20respecto%20a%20las%20variables.>
3. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (1st ed.) [PDF]. Springer.
4. Álvarez, P., & Alejandro, F. (2016). Desarrollo y evaluación de programa de incentivos para aumentar la fidelización de clientes en cuponera online.
<https://repositorio.uchile.cl/handle/2250/139196>
5. Corporativa, I. (s. f.). Descubre los principales beneficios del Machine Learning. Iberdrola. Recuperado 25 de noviembre de 2023, de
<https://www.iberdrola.com/innovacion/machine-learning-aprendizaje-automatico>
6. Herramientas de Business Intelligence (BI): cuál elegir. (s. f.). Tableau. Recuperado 25 de noviembre de 2023, de
<https://www.tableau.com/es-es/learn/articles/business-intelligence/choosing-bi-platforms>
7. Pérez, S. D. (s. f.). El impacto de la IA en la toma de decisiones empresariales. Intelequia. Recuperado 25 de noviembre de 2023, de
<https://intelequia.com/es/blog/post/el-impacto-de-la-ia-en-la-toma-de-decisiones-empresariales>
8. Rueda, J. F. V. (2019, agosto 4). Aprendizaje supervisado y no supervisado. healthdataminer.com.
<https://healthdataminer.com/data-mining/aprendizaje-supervisado-y-no-supervisado/>
9. (S. f.). Amazon.com. Recuperado 25 de noviembre de 2023, de
<https://aws.amazon.com/es/what-is/machine-learning/#:~:text=El%20machine%20learning%20es%20la,basarse%20en%20patrones%20e%20inferencias.>

9. Anexo

- A. **Cashback:** Consiste en la devolución de un porcentaje de las compras de los clientes al realizar una compra. Este monto se devuelve a su cuenta personal BCI con la libertad de utilizarla sin un catálogo de canje en específico.
- B. **Comisión por venta derivada:** Comisión de ventas provenientes de pagos realizados con productos BCI (Tarjetas de débito, crédito, cuenta MACH, etc.). Esta comisión permite el financiamiento del *cashback* para el cliente y un porcentaje de ganancia para el banco.
- C. **Desarrollador de negocios:** Desarrolla en mayor profundidad las propuestas comerciales sobre el comercio en cuestión y verifica la potenciación del negocio frente a la principal variable de beneficios de devolución de dinero (*cashback*).
- D. **Hacker de negocios BCIPlus+:** Encargado de la activación e implementación correcta de las campañas comerciales en los sistemas digitales de BCI.
- E. **Consultor de crecimiento CRM:** Da a conocer al comercio la propuesta comercial e implementación de la propuesta. Ayuda con la comunicación frente al comercio de los procesos que se llevan a cabo y de instruir sobre el funcionamiento de estas campañas comerciales.
- F. **MAUS (Month active users):** Clientes activos o diferentes del mes. Dentro del contexto de negocio del banco, esta medida permite observar la actividad de los clientes en los diferentes productos que ofrece el banco, por ejemplo, MAUS en transacciones específicas, proyectos, o en este caso, sobre las campañas de *cashback* en BCIPLUS+.
- G. **GMV (Gross merchandise value):** Valor total de bienes vendidos dentro de una plataforma. Dentro del contexto analítico del banco, esto corresponde a la multiplicación de las transacciones y el monto promedio de venta (También se le hace referencia como “volumen de venta”).
- H. **Sectores clave de aplicación del *machine learning* y ejemplos de uso:**
- **Fabricación de productos:** Mantenimiento de productos y derivados, ayuda en el control de calidad, logística, administración de inventarios, cadena de suministros y métodos inteligentes de producción.
 - **Investigación biológica:** Ayuda enfocada en la investigación, comportamiento sobre datos de múltiples áreas biológicas y la salud, investigación de enfermedades, diagnósticos y tratamientos.
 - **Servicios financieros:** Mejoran la regulación dentro de los activos financieros que entrega una determinada entidad, análisis de riesgo, análisis de comportamiento del mercado, calibración de carteras financieras y mitigaciones sobre indicios de fraude.

- **Entretenimiento y multimedia:** Principalmente, su uso se ha enfocado en entender el comportamiento de la audiencia, ofrecer contenidos de manera inteligente, certera y personalizada, e incluso, agilizar la producción de contenido.

I. Error: distancia entre el valor pronosticado por un modelo y el valor real en los datos.

J. Fórmulas de rendimiento de un modelo predictivo de ventas

- **Error absoluto porcentual medio:**

$$MAPE = \frac{1}{n} \times \sum \left| \frac{\text{Valor actual(real)} - \text{Valor pronosticado}}{\text{Valor actual(real)}} \right|$$

- **Error cuadrático medio:**

$$MSE = \frac{1}{n} \times \sum_{i=1}^n (\text{Valor real} - \text{Valor estimado})^2$$

- **Error medio absoluto (MAE)**

$$MAE = \frac{1}{n} \times \sum_{i=1}^n |\text{Valor real} - \text{Valor estimado}|$$

K. Considerando:

- I = Impacto (Alto/Medio/Bajo)
- P = Probabilidad (Alto/Medio/Bajo)
- Valor = Valor del riesgo del 1 al 9
- Nivel = Nivel del riesgo (Alto/Medio/Bajo)

ID	Tipo de riesgo	Riesgo		I	P	Evaluación		Respuesta	Responsable
		Fuente	Consecuencia			Valor	Nivel		
1	Técnico	Problemas para el acceso a las principales plataformas de trabajo: Databricks, PowerBI y AWS (Caídas del sistema, problemas con las credenciales o bases de datos particulares).	No permite el avance del proyecto ni acceso a las principales fuentes de datos.	M	B	2	B	Comunicar a Scrum master y BIS (analistas de datos) del equipo de BCILabs.	Soporte TI BCI / Equipo de gobierno de datos
2	Técnico	Intercepción de datos cruciales del banco y ciberseguridad.	Al manejar un equipo de la empresa y tener acceso a los principales datos, se es vulnerable a potenciales amenazas que puedan interferir. Podría no realizarse el proyecto a tiempo.	A	M	6	A	Se tiene que mantener una buena cultura en cuanto a la seguridad informática, uso de los equipos, herramientas de trabajo y comunicación. En caso sospechoso de actividad en los sistemas del equipo, comunicar al equipo de soporte TI BCI.	Nicolás
3	Calidad	El control de calidad de los datos es deficiente.	Demora en extracción de datos y tiempos de ejecución del proyecto.	M	A	6	A	Verificar en detalle la integridad de los datos antes de su procesamiento.	Nicolás / Leandro
4	Técnico	Deterioro del equipo informático.	No se permitiría entrar a las principales herramientas de trabajo y comunicación con los equipos.	A	B	3	M	Se tiene que mantener el debido cuidado e integridad al equipo entregado por BCI. En caso de algún desperfecto, avisar de forma interna al equipo de BCILabs y dirigirse al soporte TI BCI para tramitar un cambio o revisión del equipo.	Nicolás
5	Gestión	Cambio de objetivos proyecto BCIPPlus+	El proyecto se tendría que volver a construir en base a la nueva orientación del proyecto, y por ende, su retraso.	A	B	3	M	Se tiene que idear nuevas estrategias para abordar los problemas requeridos.	Equipo BCIPPlus+
6	Experiencia	El equipo BCIPPlus+ no trabaja al ritmo necesario.	El proyecto podría no realizarse a tiempo.	M	M	4	M	Se presionará al equipo para acelerar el proceso, tratando de emitir con tiempo los requerimientos asociados y avisando al supervisor el estado de ello.	Nicolás
7	Gestión	Equipo de TI BCI no realiza las aprobaciones correspondientes para la obtención de licencias de programas necesarios.	El proyecto no podría obtener las principales herramientas de trabajo para su construcción. Retraso o cambio en la metodología del proyecto.	M	M	4	M	Se tiene que idear nuevas formas de realizar el trabajo, verificar otras opciones de aplicaciones y uso de equipos externos.	Soporte TI BCI / Head BCILabs

Tabla 7 (Anexo): Matriz de riesgo del proyecto

- L. VAN (Valor actual neto):** Valor presente de los flujos de caja netos de algún determinado proyecto. Permite determinar la viabilidad económica de un proyecto, observando si la inversión generará más beneficios que costos en términos del valor presente.
- M. TIR (Tasa interna de retorno):** Porcentaje de ingresos que se obtiene de forma periódica debido a alguna inversión. Tasa a la cual los beneficios de una inversión igualan sus costos.
- N. Metodología CRISP-DM:**
- **Comprensión del negocio:** Objetivos y definición del problema de minería de datos.
 - **Comprensión de los datos:** Recopilación inicial, exploración y calidad en los datos.
 - **Preparación de los datos:** Limpieza de datos, selección de variables y conjuntos para el posterior modelado.
 - **Modelado:** Selección de técnicas de modelado, construcción del modelo y evaluación.
 - **Despliegue:** Implementación del modelo en el entorno de producción, junto al monitoreo y mantenimiento.
- O. Las ventajas de los modelos de regresión:**
- **Interpretación de resultados:** Proporciona coeficientes que pueden interpretarse directamente. Se puede verificar cómo cada variable de la base de datos a estudiar influyen en las ventas.
 - **Modelo ajustado a los datos:** Se ajustan específicamente para los datos disponibles, eliminando variables irrelevantes para el modelo y obtener resultados precisos.
 - Alta compatibilidad con softwares estadísticos, lo que facilita su implementación y uso.
 - Permite generar un modelo robusto y preciso gracias a las medidas de entrenamiento que usa el modelo, el cual minimiza los errores cuadráticos medios antes explicados en el estado del arte.
- P. Fórmulas de desempeño brindadas por BCI:**
- **MAUS:**

$$MAUS = \frac{\text{Clientes distintos del período futuro}}{30} \times \text{Duración (días)} \times \frac{\text{impacto del Cashback}}{100}$$
 - **Comisión de campaña:**

$$\text{Comisión de campaña} = \text{GMV campaña} \times \frac{(\text{Fee de comercio} - \text{Cashback de cliente})}{100}$$
 - **GMV de campaña:**

$$\text{GMV campaña} = \text{Transacciones campaña} \times \text{Monto promedio de venta}$$

- **Transacciones de campaña:**

$$\text{Transacciones campaña} = \frac{\text{Transacciones del periodo futuro}}{30} \times \text{Duración (días)} \times \frac{\text{Impacto del Cashback}}{100}$$

Q. Decisión sobre modelos de predicción

Los modelos a continuación fueron los resultantes de la experimentación en código dentro del ambiente de *Databricks*, los cuales fueron los que mejor respondieron en términos de la complejidad de implementación, velocidad de procesamiento, limitaciones sobre bibliotecas propias del ambiente de la nube y el nivel de escalabilidad al llevar esta tarea a múltiples comercios existentes dentro del ecosistema digital de BCI.

Consecutivamente, fue necesario medir el desempeño de estos modelos seleccionados para poder usar el mejor. Para ello, utilizamos las medidas del estado del arte en términos del cálculo del error, enfocándonos en 2 principales métricas por cada modelo, el error absoluto promedio (MAE) y el error cuadrático promedio (MSE). Los resultados, los podremos ver a continuación en esta tabla resumida, en la que medimos el desempeño de los 3 modelos seleccionados en sus mejores escenarios, para la estimación de las variables comentadas en la etapa 1 del desarrollo del proyecto:

Modelo		Modelo		Modelo	
Suavizado simple		Regresión lineal		Media móvil (n=3)	
Tiempo de ejecución	41 segundos	Tiempo de ejecución	15 min aprox	Tiempo de ejecución	30 segundos
MAE promedio	MSE promedio	MAE promedio	MSE promedio	MAE promedio	MSE promedio
6,898386063	962,9585776	8,757815315	880,3921734	9,813581476	3541,10915

Tabla 8 (Anexo): Rendimientos de modelos de predicción.

A raíz de los resultados de los errores calculados se observó el rendimiento de cada uno de los modelos, siendo el suavizado simple quien menos errores posee a nivel absoluto. Mientras que el modelo de regresión lineal posee un menor error a nivel cuadrático. Sin embargo, este último, posee un mucho mayor tiempo de procesamiento que los otros modelos, por lo que en estos términos, nos quedaremos con el método de suavizado simple para las estimaciones de los periodos futuros de las transacciones y clientes distintos. Cabe mencionar que esta fue una prueba de cálculo con las variables antes mencionadas, siendo los mejores escenarios de cada uno e implementando su ejecución en código *Python* para el flujo de ingreso de datos al panel de *PowerBi*.

- R. **Sintaxis DAX:** Expresiones de análisis de datos. Estas nos permiten obtener funciones, cálculos y valores, los cuales resultan en medidas en el *PowerBi* que pueden ser desplegadas en los paneles visuales.
- S. **One hot encoding:** lógica simple de transformación de variables categóricas, en la cual, se crean columnas extra por cada valor categórico y se coloca un 1 en la columna de la categoría que corresponda o 0 en caso contrario. Sirve para modelos de *machine learning* y cuantificar el impacto de las variables categóricas.
- T. **Uso de los modelos regresivos en el proyecto**

A grandes rasgos, los modelos regresivos investigan la relación entre una variable dependiente y una o más variables independientes, esto para entender la naturaleza de las relaciones entre las variables y/o para realizar predicciones.

Dentro de los modelos regresivos, las relaciones entre las variables se interpretan a través de coeficientes que verifican su impacto en la variable predictora, brindando un nivel de significancia por cada uno de estos coeficientes. Esto nos permite formar modelos robustos y de alta precisión, comparando su rendimiento en términos del error de la misma forma que se hizo en la etapa 1 de desarrollo.

La flexibilidad de interpretación, implementación y uso permite una mejor adaptación en el contexto de la empresa, en cuanto a las herramientas que se posee al alcance.

Los procedimientos ejecutados fueron apoyados gracias al libro *An introduction to statistical learning*³.

U. Curvas de impacto para las diferentes marcas comerciales (grande y pequeña)

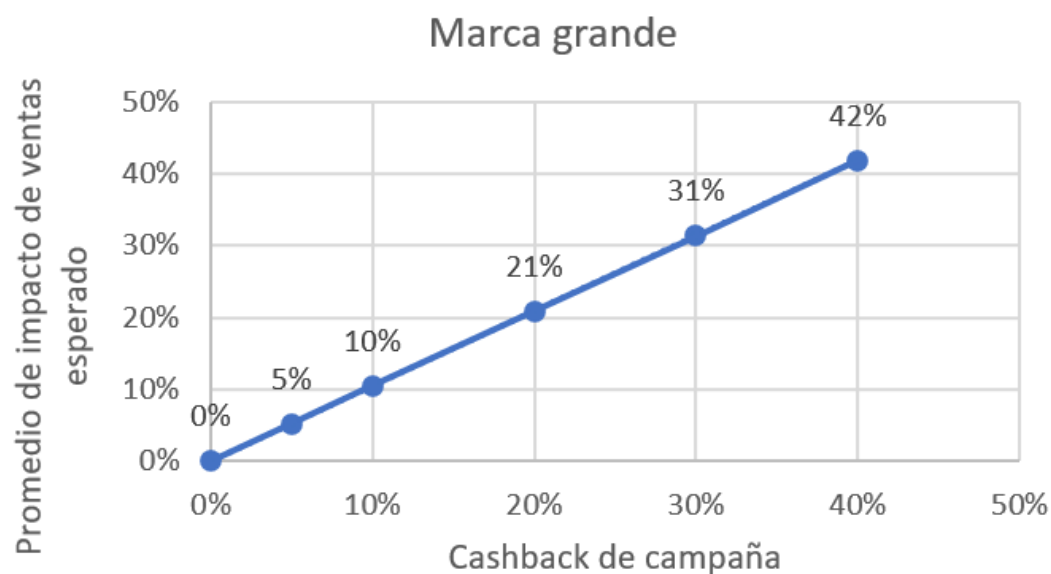


Gráfico 4 (Anexo): Promedio del impacto de ventas con cashback esperadas en una marca grande.

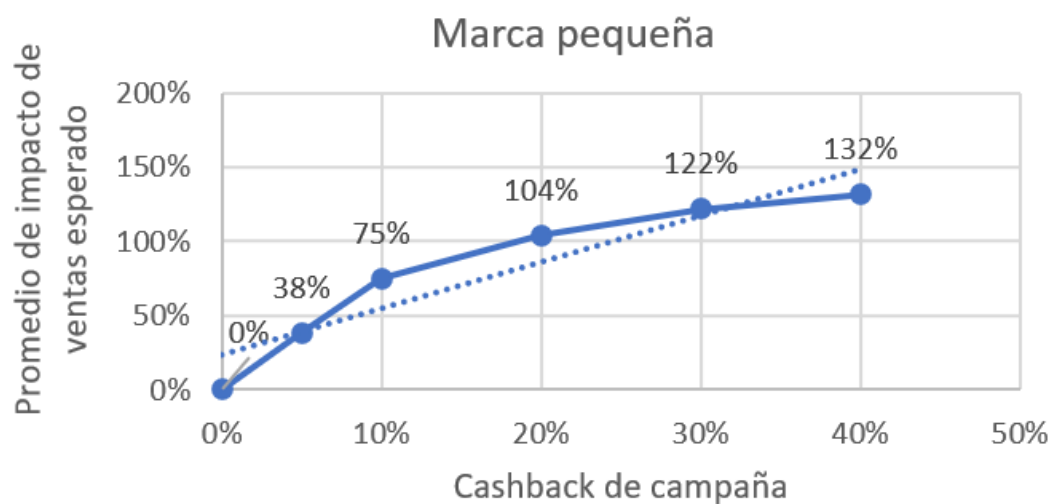


Gráfico 5 (Anexo): Promedio del impacto de ventas con cashback esperadas en una marca pequeña.

V. Diseño general etapa 2



Ilustración 7 (Anexo): Diseño del robot en su segunda etapa de desarrollo.

W. Modelo de optimización para campañas BCIPPlus+

Función objetivo

$$\text{MAX} \sum_t \text{MAUS}(t) \text{ (Clientes activos por mes)} \quad \forall t \text{ (para cada marca } t)$$

Variables de decisión

Fee de comercio(t): Porcentaje de comisión que pondrá el comercio para la campaña de la marca t .

Cashback campaña(t): Porcentaje de cashback a utilizar para la campaña de la marca t .

Restricciones

Capacidad total de gasto

$$\sum_t \text{Costo de campaña}(t) \leq \text{Presupuesto} \quad \forall t \text{ (marcas)}$$

Capacidad total de gasto por campaña

$$\text{Costo de campaña}(t) \leq \text{Gasto máximo individual} \quad \forall t \text{ (marcas)}$$

Comisión de comercio máxima a disposición de las campañas

$$\text{Fee de comercio}(t) \leq \text{Fee de comercio máximo} \quad \forall t \text{ (marcas)}$$

Cashback máximo a entregar en la campañas

$$\text{Cashback campaña}(t) \leq \text{Cashback máximo} \quad \forall t \text{ (marcas)}$$

Campañas de costo (Naturaleza del problema)

$$\text{Cashback campaña}(t) \geq \text{Fee de comercio}(t) \quad \forall t \text{ (marcas)}$$

Naturaleza de las variables

$$\text{Cashback campaña}(t), \text{Fee de comercio}(t) \geq 0 \quad \text{Enteros}, \forall t \text{ (marcas)}$$

Dentro del modelo, el conjunto de las marcas para el cual se invierte, son las mejores 10 marcas que consiguieron los mejores rendimientos en términos de usuarios activos. Esto se trabaja para cada tamaño de marca ad-hoc a su presupuesto de inversión respectivo. Cuando el modelo se procesa para un tamaño de marca en específico, se ocupan los coeficientes de impacto de ese tamaño a modo de estimación proveniente de la etapa 2 de desarrollo.

X. Fee de comercio y campañas de costo

El fee de comercio es la comisión que está dispuesto a pagar el comercio para las campañas de *cashback* dentro de un periodo de tiempo (las alianzas son temporales). En la situación de una campaña normal, el comercio dispone de la comisión necesaria para cubrir las campañas, como fue expuesto en la comisión por venta derivada en el anexo punto B. Sin embargo, en el contexto de BCIPlus+, el comercio puede no disponer de una comisión completa que abarque el *cashback* que se le entregará al cliente. En estos casos, el banco financia el porcentaje restante, siempre y cuando esta inversión valga la pena en términos de la captura de clientes MAUS. Por ello, fee de comercio es el término general en BCIPlus+ para entender la disposición de pago del comercio.

Y. Diseño final del robot BCIPlus+ y descripción general

- **Simulador y ventas:** Evolutivo de las ventas de los comercios y simulador de una posible campaña. La simulación nos entregará el potencial de MAUS, GMV, y comisiones vistas como la ganancia o el costo de la campaña simulada.

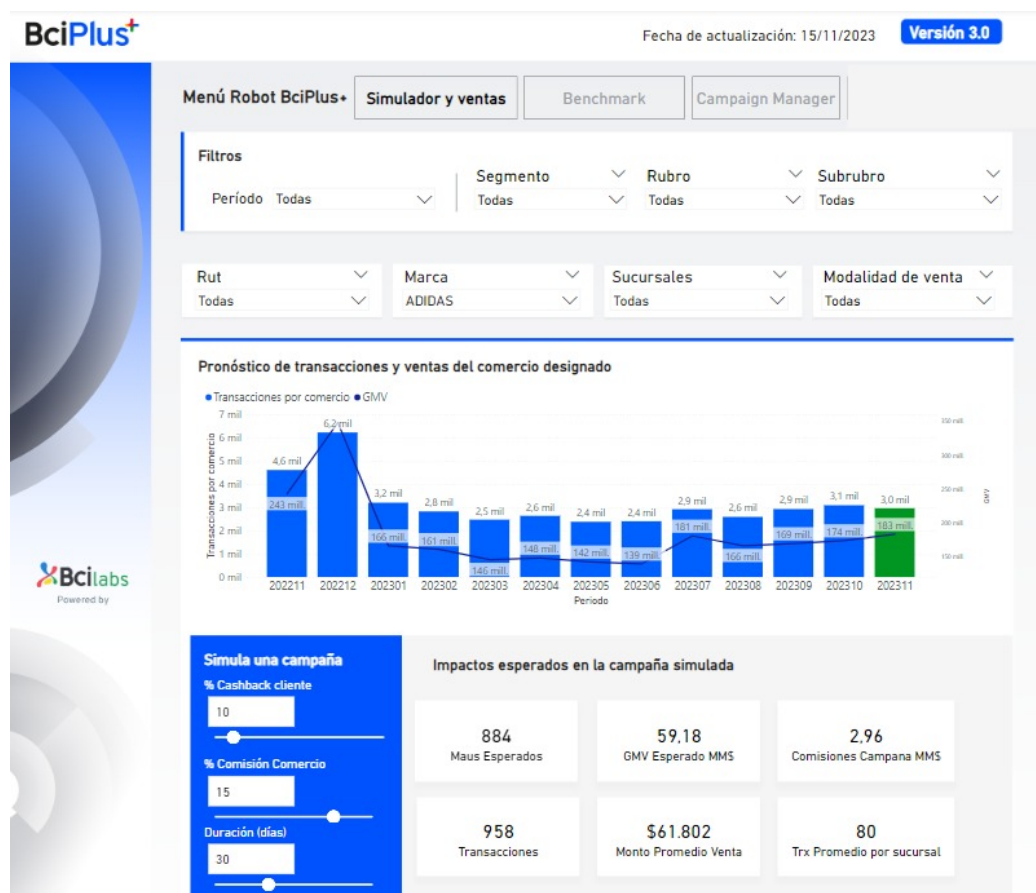


Ilustración 8 (Anexo): Vista de simulador y ventas en el diseño del robot en la tercera etapa de desarrollo.

- **Benchmark:** Acá podemos visualizar en detalle todas sucursales de los comercios seleccionados por los filtros correspondientes, permitiendo generar acciones comerciales sobre las sucursales de las marcas en la que se simula una campaña.

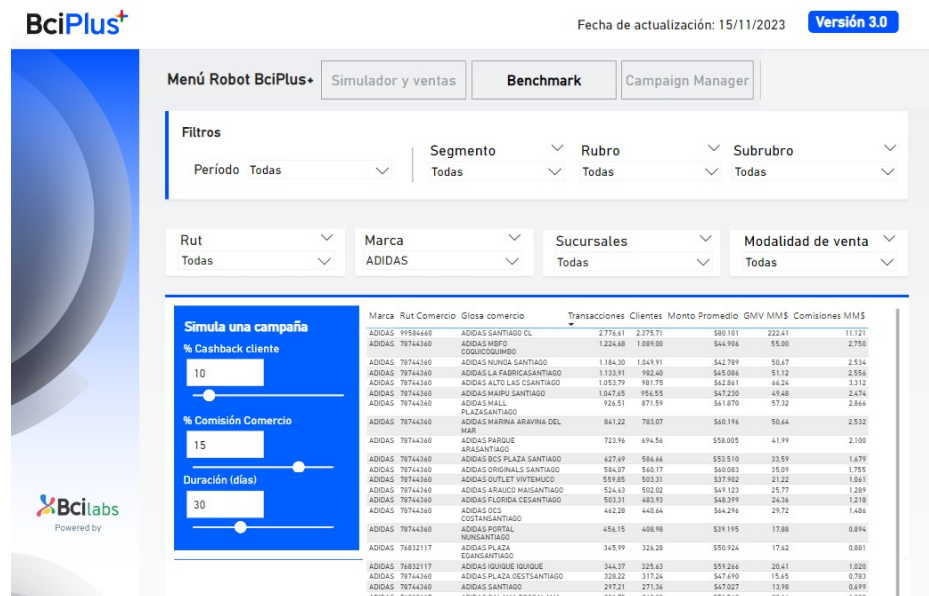


Ilustración 9 (Anexo): Vista de Benchmark en el diseño del robot en la tercera etapa de desarrollo.

- **Campaign Manager:** Función que nos permite crear estrategias de campañas mensuales con comercios, de tal forma que se maximice la ganancia de nuevos MAUS.

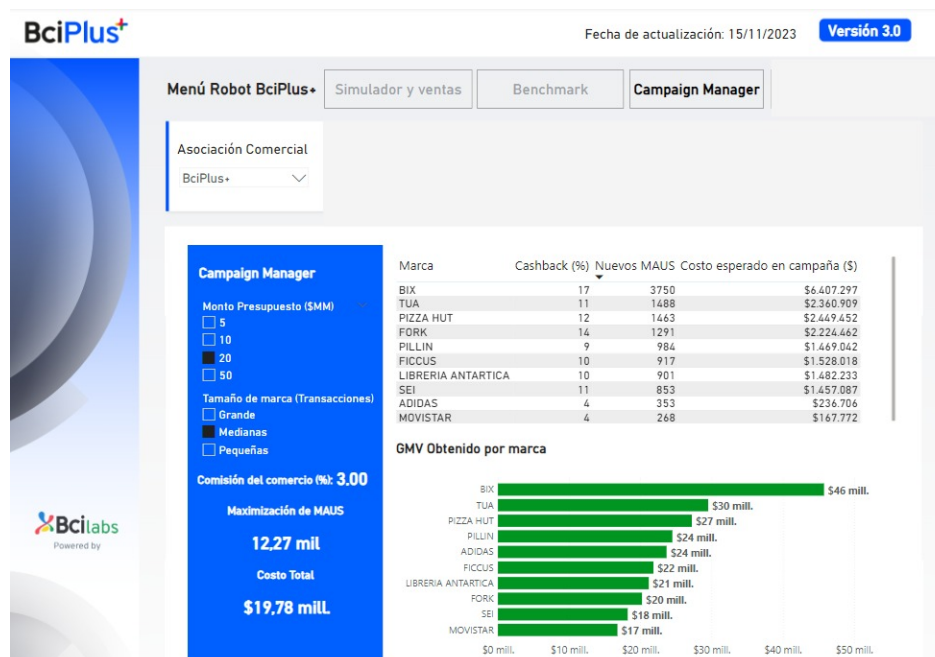


Ilustración 10 (Anexo): Vista de Campaign Manager en el diseño del robot en la tercera etapa de desarrollo.

Z. Ejemplo de decisiones de inversión el mes de octubre 2023

Los datos de las campañas de *cashback* específicas que se realizaron anteriormente no fue posible identificarlas, ya que no se pudo tener acceso a dicha información, sin embargo, sí era posible obtener el gasto que se utilizó por marca para los meses de exploración de resultados. De hecho, estos gastos mensuales en campañas fueron los presupuestos ingresados para el cálculo de optimización del robot con tal de seguir la lógica de las decisiones tomadas manualmente.

Decisiones BCIPPlus+		
Mes	Marca	Gasto incurrido
Octubre	SALCOBRAND	\$ 80.485.719
Octubre	BAR PORFIRIO	\$ 1.292.810
Octubre	BAR EXPENDIO	\$ 1.129.344
Octubre	CINEPOLIS / CINE HOYTS	\$ 1.141.208
Octubre	CORONA	\$ 563.231
Octubre	PALUMBO	\$ 442.579
Octubre	XIAOMI MI STORE CHILE	\$ 93.844
Octubre	UNCLE FLETCH	\$ 59.947
Octubre	TARINGUITA	\$ 50.000
Octubre	PILLIN	\$ 42.580
	TOTAL	\$ 85.301.262

Tabla 9 (Anexo): Decisiones de campañas en BCIPPlus+ para el mes de octubre.

Decisiones Robot BCIPPlus+		
Mes	Marca	Gasto incurrido
Octubre	SALCOBRAND	\$ 24.705.313
Octubre	PARIS	\$ 14.281.779
Octubre	PETROBRAS	\$ 7.679.101
Octubre	BCI SEGUROS	\$ 8.529.933
Octubre	EASY	\$ 5.759.145
Octubre	PALUMBO	\$ 5.519.066
Octubre	WOM	\$ 7.950.185
Octubre	PEDIDOSYA	\$ 2.879.863
Octubre	LA POLAR	\$ 1.951.833
Octubre	FASHIONS PARK	\$ 2.620.839
Octubre	CINE HOYTS	\$ 1.959.209
Octubre	CAT	\$ 204.982
Octubre	PALMERS	\$ 253.568
Octubre	EVERLAST	\$ 171.046
Octubre	PEPPERMINT	\$ 263.801
Octubre	COLUMBIA	\$ 54.961
Octubre	CUPONATIC	\$ 168.914
Octubre	VANS	\$ 78.512
Octubre	POTROS	\$ 118.784
Octubre	DOLLY	\$ 85.839
	TOTAL	\$ 85.236.673

Tabla 10 (Anexo): Decisiones de campañas del robot para el mes de octubre.