



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FISICAS Y MATEMATICAS
DEPARTAMENTO DE INGENIERIA INDUSTRIAL

EVALUACIÓN DE TÉCNICAS MEDIA MIX MODELLING PARA UNA AGENCIA DE
MARKETING DIGITAL EN CHILE

**MEMORIA PARA OPTAR AL TITULO DE INGENIERO CIVIL
INDUSTRIAL**

GERARDO ANTONIO ÁLVAREZ MEDINA

PROFESOR GUÍA:
PABLO ANDRÉS MARÍN VICUÑA

MIEMBROS DE LA COMISIÓN:
CAROLINA ANDREA SEGOVIA RIQUELME
EMILIO ERNESTO POLIT CORVALAN

SANTIAGO DE CHILE
2021

RESUMEN DE LA MEMORIA PARA OPTAR AL TÍTULO
DE: INGENIERO CIVIL INDUSTRIAL
POR: GERARDO ANTONIO ALVAREZ MEDINA
FECHA: 15/07/2021
PROFESOR GUIA: PABLO ANDRÉS MARÍN VICUÑA

EVALUACIÓN DE TÉCNICAS MEDIA MIX MODELLING PARA UNA AGENCIA DE MARKETING DIGITAL EN CHILE

El objetivo general del presente trabajo de título es evaluar la capacidad de un conjunto seleccionado de modelos de aprendizaje automático para predecir el valor que tendrá el costo por clic (CPC) de los anuncios publicados a nivel semanal por una agencia de marketing digital en Chile.

El trabajo se desarrolla en el contexto del servicio entregado por la agencia de la publicación de anuncios en línea para sus clientes. En la entrega de este servicio, si se toma como referencia un CPC igual a 0.08 [USD] y de 0.21 [USD] correspondientes a estándares a nivel de industria para las plataformas de Facebook ADS y Google ADS respectivamente, un 38.49% de las publicaciones semanales históricas han tenido un CPC mayor a este valor para Facebook ADS. Este último valor corresponde a 51.71% para plataformas de Google ADS. Estos resultados se deben en parte por la alta incertidumbre en la cual se ven envueltos los tomadores de decisiones al momento de publicar anuncios digitales, de esta manera, existe una ventana de oportunidad para reducir la incertidumbre por medio del uso de modelos predictivos que utilicen la información de desempeño histórica de la compañía.

Se evalúan siete algoritmos predictivos diferentes, donde cuatro de estos corresponden a algoritmos de clasificación, y los tres restantes, a algoritmos de predicción de variable continua. Los resultados muestran que, a pesar de faltar variables relevantes relacionadas al contenido de los anuncios y del contexto de la competencia al momento de publicación, los modelos de clasificación alcanzan niveles de acuraccy que van desde el 64.64% hasta 93.4% dependiendo de la plataforma de publicación. Por otro lado, los modelos de predicción del CPC como variable continua alcanzan errores que van desde 29.96% hasta 62.25% dependiendo de la plataforma, donde este valor significa el porcentaje que representa el error promedio absoluto de la predicción sobre la media histórica de CPC para cada plataforma.

La primera conclusión relevante de este trabajo se relaciona con que en general es el algoritmo XGBoost (Classifier y Regressor) el que mejor precisión entrega en las predicciones. Por otro lado, se concluye además que es mejor dividir el proceso de entrenamiento de modelos para cada una de las plataformas, esto debido a la heterogeneidad existente en el comportamiento de estas. Por último, se concluye que las variables más importantes a la hora de determinar el CPC alcanzado son el costo a nivel semanal, el cliente asociado, la cantidad de semanas que lleva publicado el anuncio y los días de la semana en los cuales el anuncio está en línea.

AGRADECIMIENTOS

Me gustaría agradecer en primer lugar a mi familia, quienes me han acompañado y apoyado desde el comienzo de mi formación hasta este momento.

Me gustaría agradecer también a mis amigos y amigas quienes me han dado mucha alegría a lo largo de toda mi carrera, lo cual ha sido muy importante para poder sacar adelante todo esto.

Finalmente, me gustaría agradecer a todos mis profesores y profesoras que he tenido tanto en el colegio como en la universidad, ya que sin sus enseñanzas y compromiso con la educación nada de esto habría sido posible.

TABLA DE CONTENIDO

1. Introducción	1
1.1. Características de la empresa	1
1.2. Descripción del mercado en el que participa la empresa	2
1.3. Descripción del servicio en el que se enfoca el trabajo	4
2. Problema detectado y justificación del problema.....	6
2.1. Problema existente en el servicio entregado	6
2.2. Elección del CPC como métrica de desempeño	8
3. Objetivos del trabajo de memoria	12
3.1. Objetivo general.....	12
3.2. Resultados esperados	13
3.3. Alcances del trabajo de título	14
4. Marco conceptual	15
4.1. Descripción de decisiones, métricas de desempeño y plataformas relevantes para el proceso de publicación de anuncios digitales	15
4.2. Consideraciones generales del proceso de publicación de anuncios en plataformas digitales	19
4.3. Importancia de los modelos de atribución para medir efectividad de una publicación digital	24
4.4. Marco teórico asociado a algoritmos de aprendizaje automático	29
4.4.1. Decision Tree Classifier	29
4.4.2. Redes Neuronales.....	31
4.4.3. Random Forest Classifier	33
4.4.4. XGBoost Classifier	33
4.4.5. Decision Tree Regressor	37
4.4.6. Random Forest Regressor.....	37

4.4.7. XGBoost Regressor	38
4.5. Métricas de evaluación de desempeño de modelos predictivos	38
4.5.1. Accuracy	38
4.5.2. Curva ROC (AUC)	39
4.5.3. MAE (Mean Squared Error)	41
4.5.4. RMSE (Root Mean Squared Error)	42
5. Metodología	42
5.1. Metodología general utilizada: CRISP-DM	42
5.2. Metodología en detalle utilizada para el desarrollo de modelos de aprendizaje automático.....	44
5.2.1. Metodologías generales para todos los algoritmos	45
5.2.2. Metodología particular para el desarrollo de modelos de redes neuronales	46
5.2.3. Herramientas y tecnologías utilizadas para el desarrollo de los modelos	48
6. Entendimiento de los datos.....	49
6.1. Descripción de los datos disponibles	49
6.2. Análisis descriptivo de datos	53
7. Desarrollo y evaluación de modelos predictivos	57
7.1. Modelos de clasificación	57
7.1.1. Primer modelo: clasificación de CPC por medio de un umbral único para todas las plataformas	58
7.1.2. Segundo modelo: clasificación del CPC por medio de un umbral diferenciado para cada una de las plataformas	63
7.1.3. Análisis de importancia de features	68
7.1.4. Conclusiones generales de los modelos de clasificación desarrollados	69
7.2. Modelos de predicción del CPC como variable continua.....	71

7.2.1. Predicción continua del CPC considerando todas las plataformas al mismo tiempo	71
7.2.2. Predicción continua del CPC diferenciando por plataforma	72
7.2.3. Conclusiones de los modelos predictivos del CPC como variable continua	74
8. Contraste de efectividad de modelos contra decisiones tomadas a criterio experto	74
9. Recomendaciones de implementación y propuestas de mejora para trabajos futuros.....	76
9.1. Recomendaciones de uso de modelos para realizar predicciones.....	77
9.2. Recomendaciones relacionadas a la disponibilidad de datos	78
9.2.1. Implementación de proceso de registro de datos y consideración de nuevas variables de control	78
9.2.2. Consideración de un mayor universo de datos	81
9.3. Propuestas para trabajos futuros.....	82
10. Bibliografía.....	84
11. Anexos	86
11.1. Resultados predictivos de modelos de clasificación que consideran un umbral de predicción diferente por plataforma	86
11.2. Resultados en detalle de modelos de predicción del CPC como variable continua	88
11.3. Hiperparametros de cada algoritmo optimizados a través de búsqueda de grilla (sin considerar redes neuronales).....	91
11.4. Procedimiento de optimización de hiperparametros para el algoritmo de redes neuronales	93
11.5. Situación de ejemplo expuesta en cuestionario para expertos.....	94
11.6. Ejemplos de publicaciones para cada una de las plataformas.....	95
11.7. Resultado de Shapely Values para modelos XGBoost Classifier (umbral diferenciado).....	99

INDICE DE TABLAS

Tabla 1: Evolución de la suma de la facturación en el tiempo de las agencias publicitarias consideradas el estudio	3
Tabla 2: Venta o facturación anual de las agencias al 31 diciembre del 2017.....	3
Tabla 3: Dispersión histórica de resultados medidos en CPC semanal	7
Tabla 4: Datos relacionados a la simultaneidad de uso de plataformas para un mismo cliente.....	52
Tabla 5: Umbrales de predicción establecidos tomando como referencia la mediana histórica de CPC para cada plataforma.....	57
Tabla 6: Resultados generales de predicción para el primer modelo evaluado	58
Tabla 7: Umbrales de predicción establecidos tomando como referencia la mediana histórica de CPC para cada plataforma.....	64
Tabla 8: Resultados generales de predicción para cada una de las plataformas	64
Tabla 9: Resultados de predicción de CPC continuo considerando todas las plataformas.....	71
Tabla 10: Resultados predictivos de modelo de predicción continua que diferencia por cada una de las plataformas	73
Tabla 11: Resultados generales de predicción para la plataforma Facebook.....	86
Tabla 12: Resultados generales de predicción para la plataforma Instagram	86
Tabla 13: Resultados generales de predicción para la plataforma Audience Network	86
Tabla 14: Resultados generales de predicción para la plataforma Messenger.....	87
Tabla 15: Resultados generales de predicción para la plataforma Google Search....	87
Tabla 16: Resultados generales de predicción para la plataforma Shopping	87
Tabla 17: Resultados generales de predicción para la plataforma Display	88
Tabla 18: Resultados de predicción de CPC continuo considerando solo Facebook..	88
Tabla 19: Resultados de predicción de CPC continuo considerando solo Instagram	88
Tabla 20: Resultados de predicción de CPC continuo considerando solo Audience Network.....	89

Tabla 21: Resultados de predicción de CPC continuo considerando solo Messenger	89
Tabla 22: Resultados de predicción de CPC continuo considerando solo Google Search	89
Tabla 23: Resultados de predicción de CPC continuo considerando solo Shopping	90
Tabla 24: Resultados de predicción de CPC continuo considerando solo Display	90
Tabla 25: Conjunto de hiperparametros optimizados por búsqueda de grilla para algoritmo Decision Tree Classifier [22]	91
Tabla 26: Conjunto de hiperparametros optimizados por búsqueda de grilla para algoritmo Random Forest Classifier [22]	91
Tabla 27: Conjunto de hiperparametros optimizados por búsqueda de grilla para algoritmo XGBoost Classifier [21]	91
Tabla 28: Conjunto de hiperparametros optimizados por búsqueda de grilla para algoritmo Decision Tree Regressor [22]	92
Tabla 29: Conjunto de hiperparametros optimizados por búsqueda de grilla para algoritmo XGBoost Regressor [21]	92
Tabla 30: Conjunto de hiperparametros optimizados por búsqueda de grilla para algoritmo Random Forest Regressor [22]	92
Tabla 31: Conjunto de hiperparametros considerados y sus valores testeados	93
Tabla 32: Orden de optimización de hiperparametros por medio de procedimiento de búsqueda de grilla	93

INDICE DE FIGURAS

Figura 1: Diagrama simplificado de proceso de publicación digital de anuncios	4
Figura 2: Porcentaje de anuncios con resultado insuficiente a nivel semanal	6
Figura 3: Funnel de conversión de compra	8
Figura 4: Relación ROAS y CPC para las plataformas de Facebook ADS	10
Figura 5: Relación ROAS y CPC para las plataformas de Google ADS	11
Figura 6: Efecto de respuesta inmediata (línea punteada corresponde a línea base de ventas)	20

Figura 7: Efecto de desgaste temporal (línea punteada corresponde a línea base de ventas).....	21
Figura 8: Efecto de forma (el eje horizontal representa el gasto en publicidad y el eje vertical la métrica de desempeño)	22
Figura 9: Efectos estacionales.....	23
Figura 10: Esquema del funnel de compra (4A).....	25
Figura 11: Esquema de funnel de compra (5A).....	26
Figura 12: Modelo de atribución de ultima interacción	27
Figura 13: Modelo de atribución de primera interacción	28
Figura 14: Modelo de deterioro en el tiempo.....	28
Figura 15: Esquema ejemplificativo del funcionamiento del algoritmo Decision Tree Classifier.....	30
Figura 16: Ejemplo de estructura red neuronal profunda.....	31
Figura 17: Esquema general algoritmo Xtreme Boost gradient.....	34
Figura 18: Curva ROC	40
Figura 19: Metodología CRISP-DM	43
Figura 20: Estructura inicial de la red neuronal utilizada para clasificar (los guiones representan los hiperparametros que deben se optimizados por medio del uso de búsqueda de grilla)	47
Figura 21: Estructura final de la red neuronal utilizada para clasificar (los hiperparametros a ser ajustados se muestran a través de un "-")	47
Figura 22: Magnitud histórica del CPC agregado a nivel semanal.....	53
Figura 23: Distribución del CPC para cada una de las plataformas de publicación...54	54
Figura 24: Distribución del CPC para cada una de los clientes considerados.....55	55
Figura 25: Cantidad de registros disponibles para cada una de las plataformas56	56
Figura 26: Accuracy detallado para cada una de las clases y cada una de las plataformas.....59	59
Figura 27: Cantidad de registros con clase positiva y negativa de cada plataforma en la base de testeo	60

Figura 28: Porcentaje de registros con clase positiva y negativa de cada plataforma en la base de testeo	61
Figura 29: Distribución del CPC para cada una de las plataformas de publicación en contraste con el umbral de predicción de 0.14 [USD].....	62
Figura 30: Porcentaje de registros con clase positiva y negativa de cada plataforma en la base de testeo	65
Figura 31: Cantidad de registros por clase en la base de testeo dividiendo por plataforma	66
Figura 32: Capacidad de predicción (accuracy) diferenciando por plataforma y por clase.....	67
Figura 33: Ejemplo de situación expuesta en cuestionario para ser resuelta por algún(a) experto(a).....	94
Figura 34: Anuncio de zapatillas en Google Search	95
Figura 35: Anuncio de educación impreso en el Feed de Facebook en un navegador web.....	95
Figura 36: Anuncio impreso en el feed de la aplicación móvil Instagram	96
Figura 37: Anuncio de zapatillas en Shopping.....	96
Figura 38: Anuncio de guitarras desplegado en alguna parte de un sitio web.....	97
Figura 39: Anuncio impreso en medio del uso de una aplicación que admite publicidad.....	97
Figura 40: Anuncio impreso en el feed de la red social Facebook, en cual desencadena una conversación al hacer clic	98
Figura 41: Shapely Values para la plataforma Facebook	99
Figura 42: Shapely Values para la plataforma Instagram	99
Figura 43: Shapely Values para la plataforma Google Search.....	100

1. Introducción

1.1. Características de la empresa

El trabajo de memoria se realiza en una empresa correspondiente a una agencia de marketing digital con operación a nivel nacional e internacional. Los servicios que esta organización ofrece a sus clientes pueden dividirse principalmente en 3 categorías:

1. **Servicios generales de marketing digital:** aquí se considera la creación de conceptos y campañas digitales, la gestión de redes digitales, servicios de email marketing, embajadores de marca, acciones de marketing "Bellow the Line" (BTL) y, por último, la generación de contenido digital, contenidos para blogs y mantenimiento web
2. **Servicios de diseño digital:** la agencia ofrece desarrollo de sitios web, sitios de e-commerce y de landing pages. Por otro lado, también se ofrece el diseño de newsletters, banners, gráficas, templates y material corporativo
3. **Servicios de publicidad online:** entre los servicios que la empresa ofrece se encuentran servicios de compra programática, servicios de publicación de anuncios en motores de búsqueda y redes sociales, generación de leads y de análisis en tiempo real

Los clientes de la agencia son principalmente empresas (negocio B2B) que deseen acceder a alguno de los servicios anteriormente mencionados.

La empresa comunica no poseer ninguna ventaja competitiva en particular, sin embargo, actualmente logran diferenciarse de las grandes agencias de marketing digital gracias al trato especializado que logran entregar a cada uno de sus clientes, el cual se ajusta a las necesidades de estos. Esto último se puede explicar debido a que las agencias de marketing digital de gran tamaño tienen mayores dificultades para entregar un servicio diferenciado a cada uno de sus clientes.

1.2. Descripción del mercado en el que participa la empresa

A continuación, se describirá brevemente el mercado en el cual se encuentra inmersa la compañía, para ello se entregarán algunas definiciones básicas relacionadas al marketing digital, se explicará de manera general el negocio de las agencias de marketing digital y se mencionará cuál es la situación actual de estas agencias a nivel nacional.

El Marketing Digital se define como: "Aplicación de tecnologías digitales para contribuir a las actividades de Marketing dirigidas a lograr la adquisición de rentabilidad y retención de clientes, a través del reconocimiento de la importancia estratégica de las tecnologías digitales y del desarrollo de un enfoque planificado para mejorar el conocimiento del cliente, la entrega de comunicación integrada específica, y los servicios en línea que coincidan con sus particulares necesidades". De esta manera, con el auge de las tecnologías de la información y su fácil acceso a la población, comienza también el auge de las agencias de Marketing Digital que entregan este servicio. [1]

Las agencias de marketing digital, por lo tanto, a nivel general, podrían describirse de la siguiente manera: "Una agencia digital es la encargada de crear estrategias, administrar contenidos, difundir los mensajes, escuchar y responder a los usuarios para mejorar la difusión y percepción de las marcas, pública contenidos digitales dinámicos y entretenidos que logren la participación del usuario. También es la encargada de alinear los objetivos comerciales, creando contenido en plataformas digitales que generen interacción y conversación, creando así, un estímulo para fomentar un vínculo emocional entre marca y usuarios". [2]

Para tener una idea del tamaño de la industria de las agencias de marketing digital en Chile, la Asociación Chilena de Publicidad (ACHAP) realizó el año 2018 un estudio titulado "Radiografía de las Agencias Publicitarias". [3] Este estudio considera datos de hasta fines del año 2017 correspondientes a 38 agencias de publicidad a nivel nacional, y muestra entre sus principales hallazgos que la dotación de estas agencias es de 66 personas en promedio y contabiliza que la facturación total realizada por todas las agencias consideradas en el estudio fue de aproximadamente 105 mil millones de USD para el año 2017.

El crecimiento en el tiempo de esta cifra puede apreciarse en la siguiente tabla:

Tabla 1: Evolución de la suma de la facturación en el tiempo de las agencias publicitarias consideradas el estudio

Año	Suma de facturación de las agencias consideradas en el estudio (\$Mill)
2014	53.361
2016	85.518
2017	105.744

(Fuente: "Radiografía de las agencias 2018" – Asociación Chilena de Publicidad)

Por otro lado, la siguiente tabla muestra los niveles de facturación de la industria para el año 2017 pero expuesta a través de diferentes estadísticos de interés que reflejan más bien el desempeño individual de cada una de las agencias:

Tabla 2: Venta o facturación anual de las agencias al 31 diciembre del 2017

Indicador	Facturación al 31 diciembre 2017 (\$Mill/año)
Percentil 25	1,300
Promedio	2,782
Mediana	2,221
Percentil 75	3,386

(Fuente: "Radiografía de las agencias 2018" – Asociación Chilena de Publicidad)

Considerando que la facturación anual de la agencia en la que se realiza el trabajo es menor, aunque cercana a 1,3 (\$Mill/año), se puede entonces considerar a la agencia como una empresa pequeña en relación con la competencia, ubicada cercana al percentil 25 en cuanto a sus niveles de facturación anual. Por otro lado, al tener 12 personas contratadas para su

operación, su cantidad de personal representa cerca del 18% del tamaño promedio de las plantillas que tienen las agencias de publicidad a nivel nacional.

1.3. Descripción del servicio en el que se enfoca el trabajo

El trabajo de título se enfoca en el servicio de la publicación de anuncios en línea a través del uso de las plataformas pertenecientes al grupo de Facebook ADS y Google ADS. Si bien la agencia trabaja también con otras plataformas, solo estos dos grupos son analizados por motivos que serán expuestos con mayor detalle en la sección 3.3 correspondiente los alcances del trabajo de título. Un diagrama simplificado que esquematiza el proceso de entrega de este servicio se muestra a continuación:

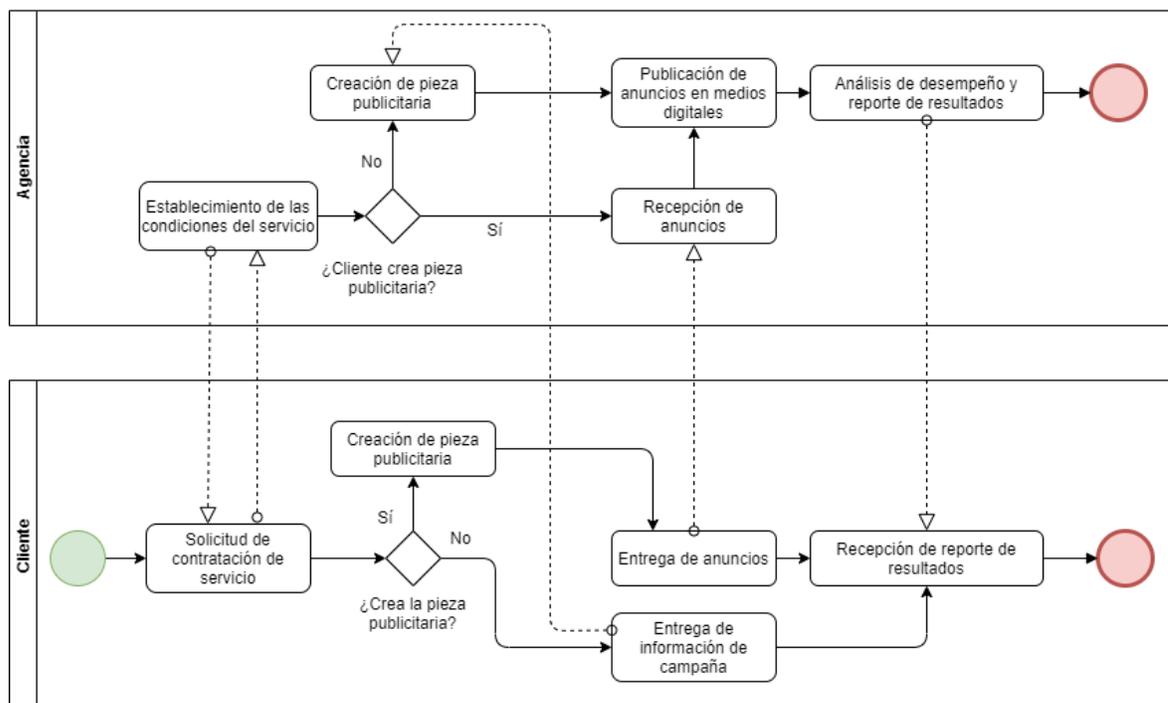


Figura 1: Diagrama simplificado de proceso de publicación digital de anuncios

(Fuente: elaboración Propia)

El proceso comienza cuando una empresa cliente contacta a la agencia para solicitar el servicio de la publicación de anuncios en línea, este servicio puede variar de cliente en cliente, sin embargo, el diagrama muestra la estructura general que se sigue para todos. Estas diferencias de cliente en cliente son determinadas en la etapa de establecimiento de los términos y condiciones de servicio, en la cual se definen, en conjunto con el cliente, las características del servicio a entregar. Estas variaciones en la entrega del servicio pueden ser de distinta índole, sin embargo, en su mayoría tienen que ver con la incidencia que tiene la agencia dentro de la campaña de marketing del cliente, y, por otro lado, el método y montos de cobro que se acuerdan con el cliente, a continuación, se habla más en detalle de estos dos puntos:

1. **Incidencia de la agencia en la campaña de marketing del cliente:** como se mencionaba anteriormente, la agencia destaca en su operación la posibilidad de entregar un servicio diferenciado para cada uno de sus clientes. De esta manera, existe un grupo de clientes que crea sus propias piezas publicitarias y se las entrega a la agencia para su posterior publicación. Por otro lado, existe también un grupo de clientes que deja en manos de la agencia tanto la creación como la publicación de estas piezas publicitarias. Por último, cualquier otro tipo de restricción particular que solicite el cliente, debe ser comunicada en la etapa de solicitud de contratación del servicio, para que posteriormente estas sean incorporadas como parte de las características del servicio que entregara la agencia
2. **Estrategia de pricing:** actualmente la manera en la que se les cobra a la mayoría los clientes es a través de una tarifa fija por cantidad de anuncios a publicar, es decir, independiente de los resultados que se tengan de las campañas publicitarias, los clientes deben pagar un monto fijo a la agencia para que esta cumpla con la tarea de la publicación de los anuncios. La excepción a esto último ocurre solo para un cliente, al cual se le tiene un sistema de pricing diferente, el cual consiste en la paga de un monto fijo al comienzo, y luego el cliente debe pagar montos adicionales por el cumplimiento de metas. Estas metas son medidas a través de la cantidad de conversiones generadas.

A pesar de estas diferenciaciones que pueden existir para cada uno de los clientes, las cuales se establecen como parte del contrato con este, siempre se sigue el proceso mostrado en la figura 1, por lo que será dentro de este proceso en donde se enmarcará el trabajo de título.

2. Problema detectado y justificación del problema

2.1. Problema existente en el servicio entregado

El problema existente y sobre el cual se basa el trabajo realizado corresponde a la alta incertidumbre en el proceso de toma de decisiones para la publicación en medios digitales. Esto último quiere decir que, al momento de tomar las decisiones de inversión para publicar un anuncio de forma digital, no se tiene un nivel de seguridad suficiente respecto a cuáles serán los resultados obtenidos de esa publicación.

Esto último se puede reflejar en los datos de desempeño históricos de la compañía, donde existen una gran cantidad de anuncios con métricas de desempeño insuficientes. A continuación, se muestra una gráfica que evidencia de manera visual este problema para la agencia:

Porcentaje de anuncios con resultados insuficientes

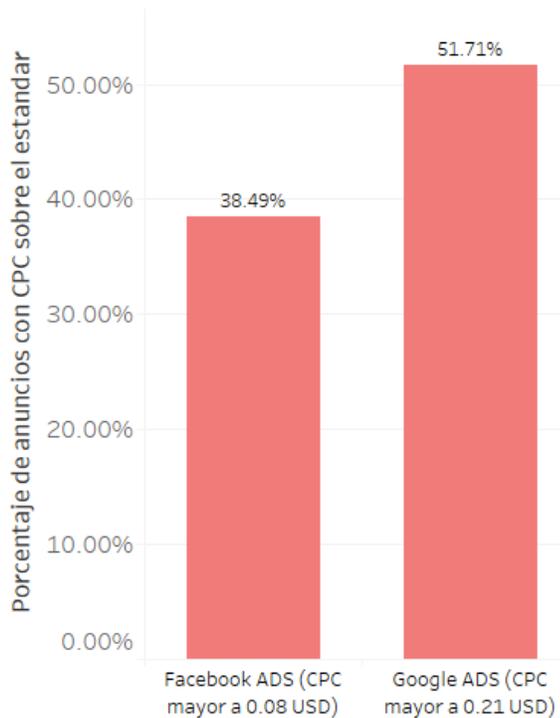


Figura 2: Porcentaje de anuncios con resultado insuficiente a nivel semanal
(Fuente: elaboración propia)

Aquí se muestra el porcentaje del total de campañas que a nivel semanal terminan teniendo un desempeño insuficiente, donde para este análisis se considera insuficiente un CPC mayor o igual al promedio estándar de la industria, el cual corresponde a 0.08 [USD] de CPC para plataformas entregadas por Facebook ADS y en 0.21 [USD] de CPC para las plataformas de Google ADS. [5][6][7]

Por otro lado, la dispersión histórica de los resultados obtenidos es también considerada alta, lo que evidencia también la incertidumbre en la cual se ven envueltos los tomadores de decisiones al momento de publicar contenido digital. La siguiente tabla, muestra la dispersión histórica a partir de los datos de cada una de las plataformas:

Tabla 3: Dispersión histórica de resultados medidos en CPC semanal

Plataforma	Porcentaje de representación de la desviación estándar sobre el promedio histórico de CPC
Facebook	144.37%
Instagram	145.97%
Audience Network	287.77%
Messenger	84.03%
Google Search	72.36%
Display	69.65%
Shopping	84.44%

(Fuente: elaboración propia)

Debido a estas informaciones es que se decide enfocar el trabajo a la búsqueda de herramientas que le permitan a la agencia anticiparse a los resultados de inversión, por medio de la entrega de estimaciones que le permitan tomar decisiones más informadas y que posibiliten la mejora de los resultados obtenidos hasta el momento.

2.2. Elección del CPC como métrica de desempeño

Un punto muy importante dentro de este problema tiene que ver con la explicación del porqué se decide elegir al CPC como métrica de medición de desempeño de una campaña publicitaria. Actualmente, para medir el desempeño de una publicación, la agencia divide el proceso publicitario en tres etapas claves, las cuales se muestran en la siguiente gráfica:



Figura 3: Funnel de conversión de compra

(Fuente: elaboración propia)

Esta gráfica muestra como varía la cantidad de usuarios que participan del proceso publicitario en cada una de las distintas fases. La primera etapa corresponde a la etapa de impresiones, en la cual a un subconjunto del segmento objetivo se le imprime la publicación en algún dispositivo para que el usuario pueda ser expuesto a la publicidad. La segunda etapa, corresponde a la etapa en la cual los usuarios hacen clic sobre el anuncio impreso, en esta, un subconjunto de los usuarios a los cuales se les imprimió el anuncio termina haciendo clic sobre este. Por último, está la última parte, la cual corresponde a conversiones, donde en esta se generan finalmente las compras o las contrataciones de los servicios, en esta, un subconjunto de los usuarios que hicieron clic sobre el anuncio terminó generando una conversión.

En cada una de estas etapas existen métricas que miden el desempeño del anuncio en esa respectiva fase, donde las más importantes corresponden al

CPM en la etapa de impresiones, el CPC y CTR en la etapa de clics y por último el ROAS en la etapa de conversión.

Intuitivamente, la métrica más útil para medir el desempeño de un anuncio debería ser el ROAS, ya que entrega de manera concreta cual fue la ganancia monetaria de haber publicado ese anuncio, sin embargo, para este trabajo de título esta métrica no es la más adecuada por los siguientes motivos:

1. No es posible determinar de manera exacta las relaciones causales entre la publicación del anuncio y la generación de conversiones, lo que termina provocando ruido sobre esta métrica de desempeño
2. Cada plataforma, para poder entregar esta métrica, estima su valor a través de un **modelo de atribución**, el cual suele variar de plataforma en plataforma, lo que impide por lo tanto la comparación de desempeño de las diferentes plataformas utilizando esta métrica de desempeño

Debido a esto, es que se decide utilizar una métrica que esté un paso atrás dentro del proceso de conversión mostrado en la figura 3, es decir, utilizar una métrica de desempeño correspondiente a la etapa de generación de clics sobre un anuncio. De esta forma, se determina que la métrica de desempeño más útil en la etapa de generación de clics corresponde al CPC, debido a los siguientes motivos:

El primero, tiene que ver con la relación que tiene el CPC con la eficiencia de un anuncio, entendiendo como eficiencia a la capacidad de un anuncio de entregar resultados de inversión positivos para la agencia. En los siguientes gráficos se puede ver la relación que tiene el CPC con el ROAS obtenido al finalizar una publicación semanal:

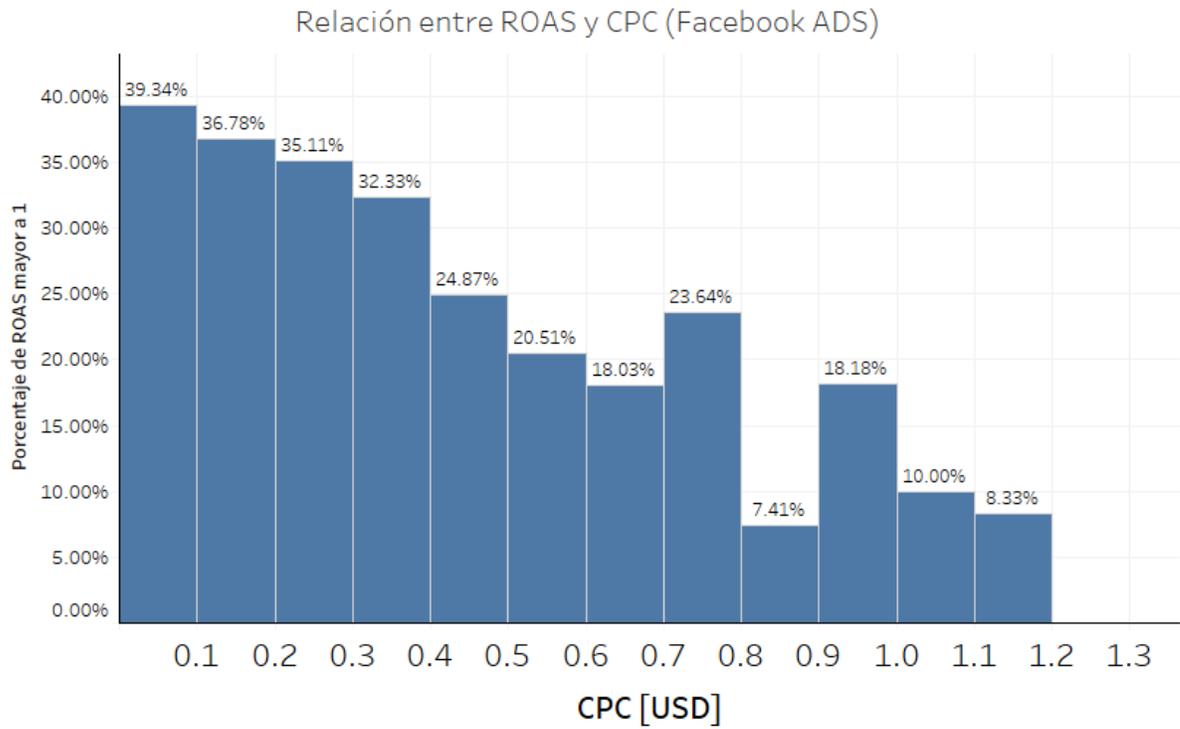


Figura 4: Relación ROAS y CPC para las plataformas de Facebook ADS
(Fuente: elaboración propia)

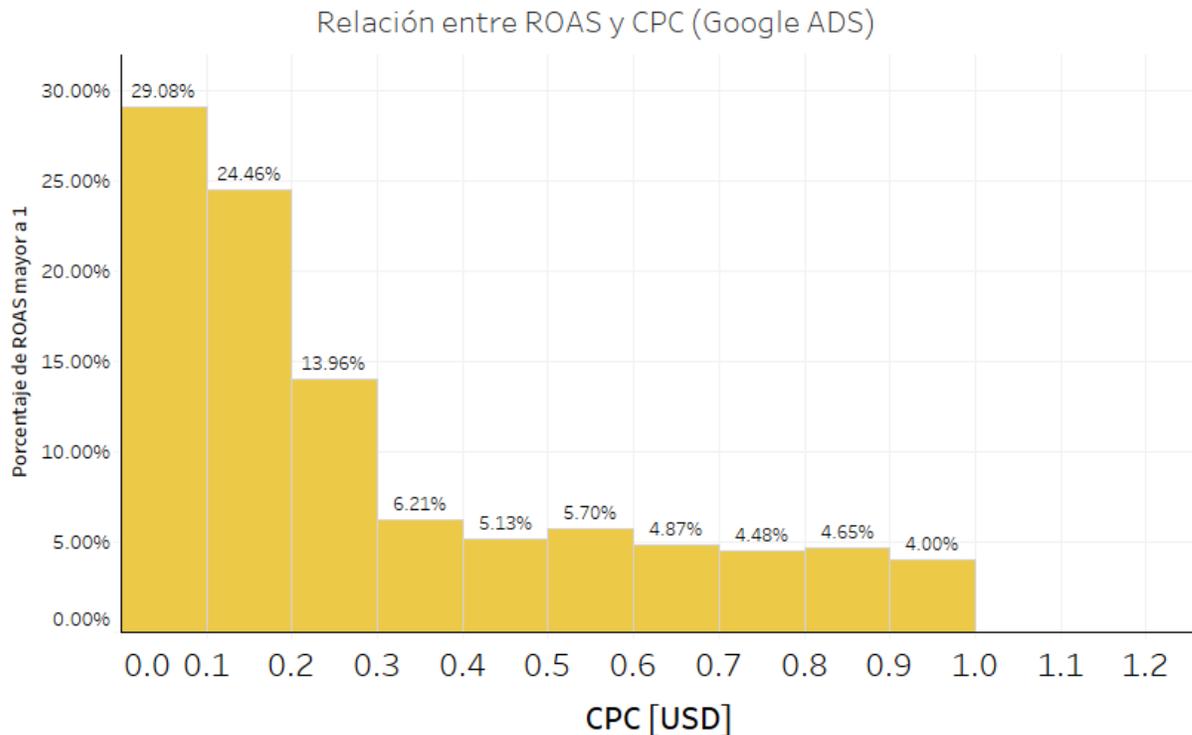


Figura 5: Relación ROAS y CPC para las plataformas de Google ADS

(Fuente: elaboración propia)

En estas gráficas, se muestra en porcentaje la cantidad de publicaciones a nivel semanal que tuvieron un ROAS mayor a uno para cada uno de los niveles de CPC, los cuales están representados en intervalos de 0.1 [USD] cada uno. Estos resultados evidencian que el CPC tiene una visible relación con los resultados monetarios de la agencia, ya que en general a un mayor CPC (peor), el porcentaje de anuncios que terminan generando ganancias monetarias disminuye, es decir, a medida que empeora el CPC, también empeoran los resultados de inversión para la agencia. Para fortalecer este punto, se puede también apreciar el hecho de que esta relación se cumple tanto para las plataformas de Google ADS como para las de Facebook ADS, a pesar del hecho de que estos dos conjuntos de plataformas presentan modelos de atribución que pueden ser completamente diferentes.

Un segundo motivo para elegir el CPC como métrica de desempeño, corresponde a que, en contraste a otras métricas que existen en la etapa de generación de clics, el CPC considera el costo que tuvo la publicación del anuncio, lo cual es beneficioso ya que, de no considerarse el costo, se estará

sobreestimando el desempeño de las plataformas que tienen un buen resultado para atraer a el usuario, pero que su alto costo no termina justificando su buen nivel de desempeño.

Un último motivo tiene que ver con que es importante el hecho de que no existe ambigüedad respecto a su forma de cálculo para cada una de las plataformas, ya que es una métrica muy concreta y fácil de medir, lo que permite por lo tanto comparar la efectividad de estas distintas plataformas por medio del CPC, lo cual no era posible cuando se consideraba el ROAS como métrica de comparación.

Para finalizar, un punto importante a mencionar en esta sección, es que no es correcto concluir a partir de los gráficos 4 y 5 que el conjunto de plataformas de Facebook ADS es más eficiente debido a que en magnitud tienen mayor cantidad de anuncios con ROAS mayor a 1 para los distintos niveles de CPC. Esto último debido a que no necesariamente tiene que ser así, ya que estos porcentajes están directamente relacionados con la forma particular de estimar conversiones que tiene cada plataforma. En este caso en particular, Facebook ADS es reconocido por sobreestimar la cantidad de conversiones que logra, ya que su modelo de atribución por ejemplo considera ventanas temporales de conversión más grandes que las que suele utilizar Google ADS. También, otro motivo, es que Facebook ADS suele atribuir una conversión al último anuncio sobre el cual el usuario hizo clic, pasando por alto totalmente los anuncios a los que pudo haber sido expuesto el usuario previamente. Es por esto que la finalidad de los gráficos mostrados en esta sección tiene que ver con explicar la relación entre CPC y ROAS más que mostrar la magnitud de resultados positivos de cada grupo de plataformas.

3. Objetivos del trabajo de memoria

3.1. Objetivo general

El objetivo general del trabajo de memoria es evaluar la capacidad de un conjunto seleccionado de modelos de aprendizaje automático para predecir el valor que tendrá la métrica de CPC de un anuncio publicado a nivel semanal, permitiéndole a la agencia poder anticiparse a los resultados de inversión.

De forma más concreta, los objetivos específicos de este trabajo de memoria son los siguientes:

1. Identificar las variables independientes que tengan una mayor relación con el desempeño de las publicaciones digitales
2. Desarrollar un conjunto seleccionado de modelos de aprendizaje automático que logren predecir el CPC alcanzado por un anuncio publicado a nivel semanal
3. Analizar y documentar ventajas, desventajas y limitantes de cada uno de los modelos desarrollados
4. Entrega de recomendaciones a la empresa basadas en los resultados y aprendizajes recolectados a lo largo del trabajo

3.2. Resultados esperados

Se espera a nivel general que este trabajo de memoria sea la primera aproximación de la agencia a comenzar a predecir métricas de desempeño por medio del uso de modelos que incorporen la información de resultados históricos de la empresa.

En una primera etapa, se espera encontrar la relevancia que tienen las diferentes variables disponibles por la empresa para modelar el CPC de los anuncios publicados a nivel semanal. En este sentido, se espera también hacer un análisis de cuáles podrían ser las eventuales variables con las que la empresa no cuenta, ya que esto será de utilidad para proponer recomendaciones y trabajos futuros.

Se espera también obtener información acerca del desempeño de diferentes algoritmos de aprendizaje automático para predecir el CPC de un anuncio publicado a nivel semanal. Un buen desempeño en esta parte se entiende como un modelo que logra tener una buena precisión en sus predicciones, que factible de implementar por la agencia y, por último, que las predicciones logren entregar conocimiento que genere valor para la compañía.

Por otro lado, se espera también entender cuáles son las características de los diferentes algoritmos que los hacen más o menos adecuados para resolver el problema en el cual se trabajará. Esto quiere decir, que se obtendrá en detalle cuáles son las características que definen la predicción de cada uno de los algoritmos y como estos se comportan ante los datos disponibles por la agencia.

Por último, se espera que, a partir de los aprendizajes logrados a lo largo de todo el trabajo, se pueda entregar a la empresa una serie de recomendaciones de como trabajar con los modelos desarrollados para resolver el problema de la alta incertidumbre en el proceso de inversión para la publicación de anuncios digitales, además, recomendaciones que le permitan a la empresa tomar acción para mejorar el desempeño de estos modelos en un futuro.

3.3. Alcances del trabajo de título

Como se mencionaba anteriormente, el objetivo principal de este trabajo de título corresponde al desarrollo y evaluación de un conjunto de modelos para predecir la métrica de desempeño CPC asociada a los anuncios publicados de manera digital, de esta manera, los alcances establecidos para este trabajo son los siguientes:

Las predicciones realizadas consideran la información perteneciente al conjunto de las 7 plataformas más utilizadas e importantes para la agencia, las cuales son Facebook, Instagram, Messenger, Audience Network, Google Search, Shopping y Display, cada una de estas fueron descritas con mayor detalle en la sección 1.3 de este informe. De esta manera, los modelos desarrollados entregan información a la agencia relacionada con que plataforma podrá ser más o menos efectiva, pero limitándose solo al conjunto de plataformas que se acaba de mencionar.

Por otro lado, en cuanto al desempeño de los modelos predictivos, se considera solo hacer una evaluación de estos, es decir, no se tiene como meta el superar un nivel de precisión de predicción en particular. Esto último es relevante, ya que el principal valor a aportar por medio de este trabajo de tesis corresponde al conocer las posibilidades de predicción que tiene la agencia con su estructura de datos actual. Por otro lado, otra razón que motiva el establecimiento de este alcance tiene que ver con el hecho de que la agencia es consciente de que no cuenta actualmente con un registro de toda la información relevante que en la teoría es clave para determinar el desempeño de las campañas publicadas de manera digital, como lo es por ejemplo la información relacionada a acciones de la competencia o también información relacionada a el contenido que posee cada uno de los anuncios (esto se ve más en profundidad en la sección 4.1 de este informe).

Por último, el entregable que se acordó con la compañía, corresponderá a la entrega de la estructura de cada uno de los modelos, además de los estudios

y recomendaciones que se desarrollen. Queda fuera del alcance de este trabajo de título el proceso de implementación de estos modelos predictivos dentro de las operaciones de la agencia.

4. Marco conceptual

4.1. Descripción de decisiones, métricas de desempeño y plataformas relevantes para el proceso de publicación de anuncios digitales

Dentro del proceso de publicación de anuncios digitales, una de las etapas más importantes es la de publicación de anuncios en los medios digitales, en esta se deben tomar una serie de decisiones de inversión que impactarán en el desempeño de los anuncios, a continuación, se detallan cada una de estas decisiones:

1. **Elección de la plataforma y formato:** la agencia tiene un conjunto de plataformas que puede elegir para publicar un anuncio de manera digital. Entre las más utilizadas por la agencia se encuentran aquellas ofrecidas por Facebook ADS y Google ADS, las cuales concretamente son:

- Facebook ADS
 - Facebook
 - Instagram
 - Messenger
 - Audience Network

- Google ADS
 - Google Search
 - Shopping
 - Display

Al final de esta sección se describen en detalle cada uno de estas plataformas

2. **Monto invertido:** en cada publicación, es necesario tomar la decisión del monto que se quiere invertir en esa publicación. Si bien el monto invertido depende del tiempo de publicación, también existe la posibilidad de pagar más para aumentar la intensidad de publicación, por ejemplo, aumentando el alcance de la publicación
3. **Elección del día de publicación:** para cada publicación, la agencia puede decidir qué día las publicaciones permanecerán en línea. Esta decisión es relevante ya que factores como la estacionalidad y eventos especiales pueden influir en el desempeño de la publicación
4. **Dispositivo de impresión del anuncio:** algunas plataformas, como por ejemplo las derivadas de Facebook ADS o de Google ADS, dan la opción de limitar la impresión del anuncio a algunos tipos de dispositivos de impresión, los cuales pueden ser por ejemplo smartphones, tablets, Smart TV, entre otros dispositivos
5. **Variables de targeting:** algunas plataformas permiten elegir o forzar la estrategia de targeting del anuncio, por ejemplo, dándole la posibilidad de elegir al publicista el rango etario objetivo del anuncio, intereses, ubicación geográfica, entre otras variables. Sin embargo, es necesario mencionar que las plataformas más utilizadas por la agencia (las derivadas de Facebook ADS y Google ADS) tienen su propio sistema de optimización de targeting a través del uso de variables de remarketing [12]

Un punto importante, es que estas decisiones son tomadas actualmente a criterio experto, lo que provoca que muchas veces, y como se verá más adelante en mayor detalle, exista una alta incertidumbre en relación a los resultados que se obtendrán luego de tomar las decisiones.

De esta manera, luego de pasar por el proceso de toma de decisiones para cada anuncio, y de haberse obtenido resultados, se crea un reporte que resume la información de la efectividad del anuncio, o de un conjunto de anuncios publicados en un periodo de tiempo establecido. Estos reportes son de utilidad tanto para la agencia como para el cliente y muestran las principales métricas de efectividad de los anuncios, las cuales se detallan a continuación:

1. **Costo:** corresponde a la suma del gasto que se debió realizar para mantener la publicación activa en alguna plataforma en particular
2. **Cantidad de impresiones:** corresponde a la cantidad de veces que el anuncio fue impreso (mostrado) en los dispositivos del público objetivo
3. **Cantidad de clics:** corresponde a la cantidad de clics que el público objetivo hizo sobre los anuncios
4. **Cantidad de conversiones:** corresponde a la cantidad de compras del producto o servicio que fueron generadas por efecto de la publicación del anuncio. Es importante mencionar que cada plataforma tiene su propia manera particular de determinar cuándo una compra fue causada por la exposición del usuario objetivo al anuncio. Esto último se analiza más en profundidad en la sección 4.2 de este informe
5. **CPC (Costo por clic):** corresponde al total del gasto dividido por la cantidad de clics hechos sobre el anuncio
6. **CTR (Click-through rate):** corresponde al total de clics obtenidos divididos por la cantidad de impresiones del anuncio
7. **ROAS (Return on Ad Spend):** corresponde a la suma de los montos de las ventas logradas por efecto de la publicación del anuncio dividida por el gasto total que se debió realizar para mantener el anuncio en línea. De esta manera si el ROAS tiene un valor mayor a 1, quiere decir que se recuperó la inversión en publicidad, en el caso contrario, las ganancias no lograron justificar el gasto en publicidad, es decir, se perdió dinero por efecto de la publicación. Es importante mencionar que las ganancias por conversión son calculadas de cara al cliente, y dependiendo del acuerdo de cobro que se tenga con este, se distribuyen las ganancias entre el cliente y la agencia

Por último, se describirán las 7 plataformas que son utilizadas por la agencia y que formarán parte de este trabajo de título (en el anexo 11.6 se puede

encontrar un ejemplo gráfico de un anuncio perteneciente a cada una de estas):

1. **Google Search:** son aquellos anuncios que son publicados a través del motor de búsqueda de Google. Corresponde a aquellos que simulan ser un resultado de búsqueda, pero que, sin embargo, corresponden realmente a publicidad que direcciona a la página web del anunciante. Estos anuncios son aquellos que aparecen en las primeras posiciones de los resultados de búsqueda y pueden diferenciarse por tener escrita la palabra "Ad" en una de sus esquinas
2. **Facebook:** corresponde a aquellos anuncios que aparecen entre las publicaciones en el feed o en las historias de la red social Facebook. Al realizarse un clic en el anuncio, el usuario es redirigido a la página web del anunciante
3. **Instagram:** corresponde a aquellos anuncios que aparecen entre las publicaciones en el feed de la red social Instagram o también aquellos que aparecen en formato de historias dentro de esta misma red social. Al realizarse un clic en el anuncio, el usuario es redirigido a la página web del anunciante
4. **Shopping:** es otro tipo de anuncio relacionado al motor de búsqueda de Google, corresponde a una pequeña sección que aparece por lo general antes o al costado de los resultados de búsqueda, y que muestra un pequeño Marketplace donde se imprimen directamente los productos ofrecidos por los anunciantes. En este caso, un clic sobre el anuncio, redirecciona a la página web del anunciante, pero específicamente a la vista en la cual aparece el producto que fue anunciado
5. **Display:** corresponden a aquellos anuncios que aparecen dentro de diversas páginas web, aparecen en algún lugar de la pantalla del usuario y su posición dentro de la pantalla va a depender de la estructura que tengan cada uno de los sitios web

6. **Audience Network:** corresponden a aquellos anuncios que aparecen dentro de la navegación de alguna aplicación para celular o tablets, aparecen principalmente en aquellas aplicaciones que obtienen ingresos a través de la entrega de espacios en publicitarios su aplicación para la colocación de estas campañas
7. **Messenger:** corresponde a aquellos anuncios que son colocados en el feed de la red social Facebook, pero que sin embargo la interacción provoca el comienzo de una conversación por chat con la empresa que ofrece el producto o servicio

4.2. Consideraciones generales del proceso de publicación de anuncios en plataformas digitales

La literatura relacionada al problema que se intenta abordar es diversa, sin embargo, la mayoría de las publicaciones relevantes estudiadas han sido publicadas a lo largo de las última dos décadas, ya que es en este periodo de tiempo donde se comienza a masificar el uso del internet, y también, por lo tanto, la inclusión de este como nuevo canal para realizar campañas publicitarias.

De manera general, los artículos relacionados con este tema coinciden en los siguientes puntos que son claves para el problema que se aborda en este trabajo [8][9][10]:

1. Para obtener buenos resultados de predicción, es necesario tener una gran cantidad de información de desempeño histórico. Por un lado, una alta cantidad de registros, y por otro, una gran cantidad de variables de control, ya que, como se verá más adelante, las variables que pueden alterar el desempeño de un anuncio son múltiples
2. Modelar el desempeño de los anuncios utilizando distintas técnicas en una situación en la cual los datos son insuficientes, repercutirá en que las diferentes técnicas entreguen resultados muy diferentes entre sí, lo que genera desconfianza en cuanto a los resultados obtenidos
3. Los experimentos controlados no son un buen método para encontrar efectos causales relacionados a la inversión en publicidad en las

diferentes plataformas, esto debido a la dificultad de controlar variables exógenas (una vez que el anuncio es publicado dentro de una plataforma, se pierden muchas variables de control)

4. Es difícil registrar las variables de control relacionadas a una publicación, sin embargo, existen oportunidades de mejora en esta línea

Por otro lado, la bibliografía relacionada menciona también una serie de efectos que hay que tener en cuenta en el proceso de publicación de anuncios, a continuación, se muestran cada uno de estos en detalle (por simplicidad, en el análisis se utilizará la cantidad de ventas como métrica de respuesta a la exposición a la publicidad, sin embargo, en la práctica estos efectos están presentes para cualquier métrica de desempeño elegida) [11]:

- 1. Efecto de respuesta inmediata:** este efecto corresponde a la magnitud de incremento en las ventas en el mismo periodo en el cual la publicidad comienza su funcionamiento. Este efecto por lo general es bajo en magnitud si se contrasta por ejemplo con como el precio de un artículo puede hacer variar la cantidad de ventas de este. El siguiente gráfico esquematiza este efecto:

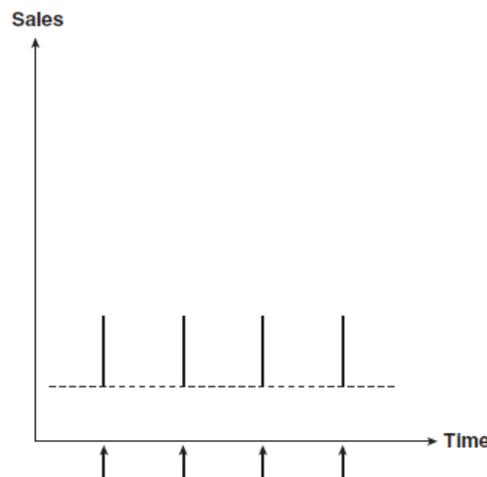


Figura 6: Efecto de respuesta inmediata (línea punteada corresponde a línea base de ventas)

(Fuente: Media Marketing Mix – Gerard J. Tellis)

2. Efecto de desgaste temporal (o Carryover Effect): este efecto se refiere a como la efectividad de una publicación disminuye a lo largo del tiempo, hasta el punto en que la publicidad ya no tiene incidencia sobre la cantidad de ventas u otra métrica de desempeño. El siguiente grafico esquematiza este efecto:

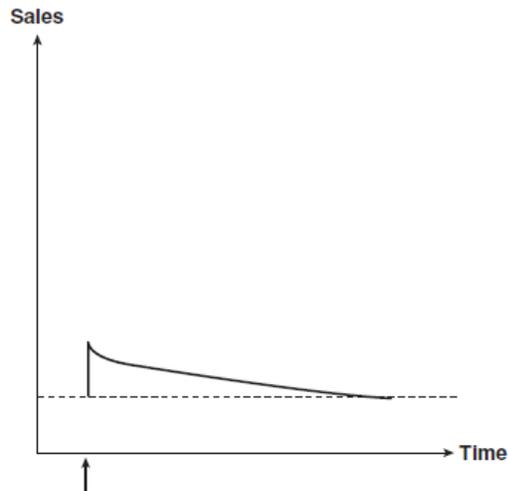


Figura 7: Efecto de desgaste temporal (línea punteada corresponde a línea base de ventas)

(Fuente: Media Marketing Mix – Gerard J. Tellis)

3. Efecto de forma: este efecto se refiere a cómo se comporta el desempeño de una publicación en función de la magnitud en el gasto en publicidad. A continuación, se muestra una gráfica que muestra diferentes opciones de comportamiento:

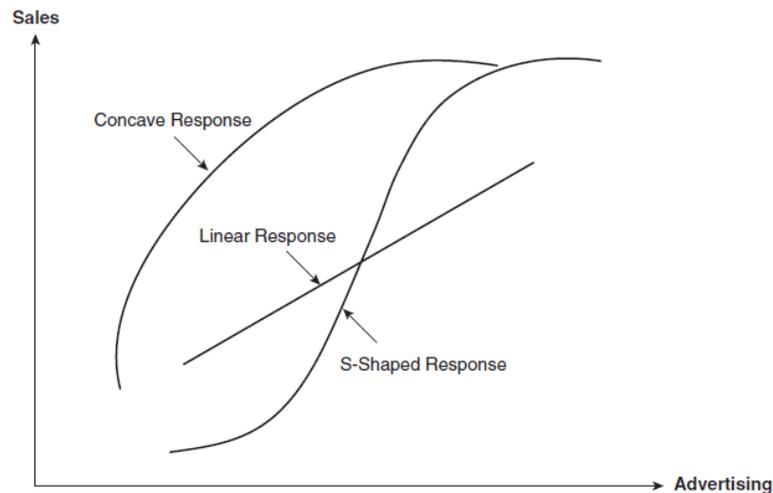


Figura 8: Efecto de forma (el eje horizontal representa el gasto en publicidad y el eje vertical la métrica de desempeño)

(Fuente: Media Marketing Mix – Gerard J. Tellis)

Este efecto dice principalmente que no necesariamente existe un efecto lineal entre el gasto que se hace en publicidad y la respuesta obtenida. Lo más común, es que el comportamiento sea en forma de S, es decir, a baja inversión de publicidad, se obtengan bajos retornos, y a partir de un punto, el desempeño crezca de manera exponencial, sin embargo, este crecimiento llegará hasta el límite de agotamiento del anuncio, punto en el cual el aumento de ventas marginal por efecto del aumento en gasto de publicidad será muy cercano a cero. Esto último da cuenta de un efecto muy importante a la hora de publicar anuncios, y es que en teoría la efectividad del anuncio tendrá un máximo de retorno al cual podrá llegar en función de gasto realizado

4. **Efecto de la competencia:** este efecto se refiere a como las acciones de la competencia inciden en el desempeño de los anuncios publicados. Este es uno de los efectos más complicados de modelar, ya que requiere integrar en el modelamiento variables de control que son exógenas al publicista. Una aproximación útil sugerida para la modelación de este efecto corresponde al considerar el marketshare del producto o servicio publicitado en el anuncio, o de la empresa que vende este [11]

5. **Efectos dinámicos o de estacionalidad:** este efecto corresponde a él como ciertos eventos de estacionalidad afectan el desempeño de un

anuncio. Estos efectos se refieren a por ejemplo a algunas épocas de año, periodos festivos, épocas de mayor o menos poder adquisitivo de la población, eventos puntuales como cyberday, o cualquier otro evento temporal que altere la efectividad de una publicación. Existen dos maneras en las que un efecto estacional puede afectar el desempeño, las cuales se pueden ver en el gráfico de a continuación:

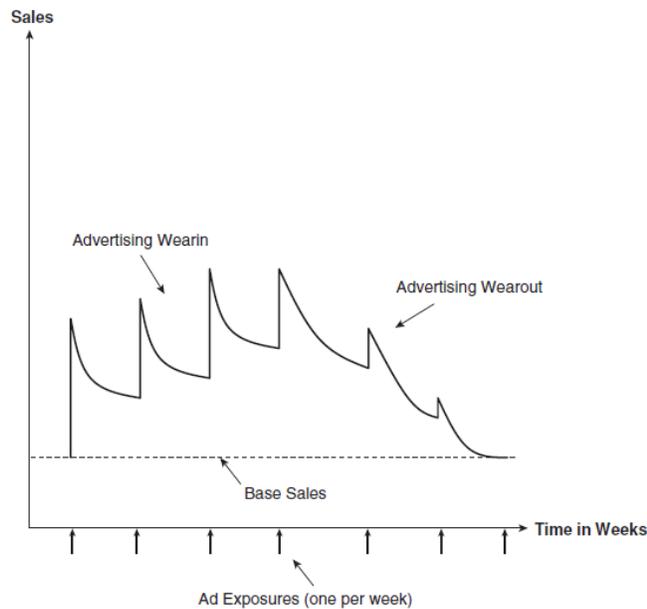


Figura 9: Efectos estacionales

(Fuente: Media Marketing Mix – Gerard J. Tellis)

La primera corresponde al efecto de Wearin de una publicación, donde el efecto estacional hace que de manera macro aumente el desempeño de los anuncios publicados. Por otro lado, existe el llamado efecto Wearout, que significa lo mismo, pero en sentido contrario, es decir, provoca una disminución a nivel general de las ventas. Se dice que es un efecto macro ya que es un efecto que es independiente a los efectos vistos anteriormente, es decir, este efecto se presenta en paralelo a los demás efectos descritos

6. Efectos de contenido: este efecto se refiere a como el contenido específico de cada anuncio afecta en el desempeño de este. Este último efecto, suele ser también uno de los más complicados de modelar a la hora de trabajar con un modelo de marketing mix, ya que es

relativamente complicado registrar variables que representen de manera fiel el contenido de cada uno de los anuncios

7. Efecto de plataforma o medio: este efecto se refiere a como la plataforma o canal de publicación elegido, afecta en el desempeño de una publicación. Este efecto es clave para el trabajo de memoria a desarrollar, ya que se busca entregar información acerca de que plataformas digitales pueden ser más o menos adecuadas para la publicación de los anuncios. Un punto muy importante a tratar en este contexto, corresponde a como cada plataforma digital mide su propio desempeño, aquí aparece el concepto de modelos de atribución de cada plataforma, el cual será descrito en detalle en la sección siguiente.

4.3. Importancia de los modelos de atribución para medir efectividad de una publicación digital

Un concepto muy importante a la hora de evaluar la efectividad de una plataforma de publicación de anuncios corresponde a como cada una de las plataformas se atribuye los "logros" de marketing. En este sentido, para medir cuantas ventas genera un anuncio publicado en alguna plataforma en particular, se usa un **modelo de atribución**, el cual suele ser único de la plataforma en particular en la cual se publicó el anuncio. Para entender en que consiste un modelo de atribución, es necesario primero conocer acerca del concepto de **embudo o funnel de compra**.

El embudo o funnel de compra representa las distintas etapas por las que pasa un comprador antes de finalmente llegar al último evento correspondiente a la compra de un producto o servicio. En la siguiente imagen se muestra un esquema de este proceso:



Figura 10: Esquema del funnel de compra (4A)

(Fuente: Polit, Emilio. Clases Curso Marketing Digital - Capitulo 4 – Otoño 2021)

Esta muestra que un comprador, antes de realizar la compra como tal, pasa por cuatro etapas generalmente: la etapa de conocimiento del producto o servicio, la etapa en la que se genera el atractivo o interés, la etapa en la cual el comprador realiza la acción de compra del producto o servicio y finalmente la etapa en la cual el comprador realiza efectivamente la compra. Los anuncios publicitarios siempre tienen el objetivo de atacar en alguna de estas etapas en específico, por ejemplo, pueden existir anuncios que no tengan una intención de generar una compra directa por parte del potencial comprador, si no que tengan la intención de generar conocimiento de marca (etapa de atención o conocimiento), como también pueden existir anuncios que tengan como objetivo la compra del producto, es decir, que estos se enmarquen en la etapa acción. Es importante mencionar que es este el funnel de compra teórico en el cual se basa el trabajo de memoria (4A), el cual no incluye una etapa en particular donde se considerará la generación de leads. Es importante tener en cuenta que este funnel ha evolucionado los últimos años, debido a la masificación de las redes sociales, las cuales les permiten a los consumidores investigar acerca del producto o servicio que quieren adquirir obteniendo incluso opiniones de otras personas acerca del producto o servicio. De esta

manera, el funnel de compra que es más adecuado a los nuevos contextos es el siguiente:



Figura 11: Esquema de funnel de compra (5A)

(Fuente: Polit, Emilio. Clases Curso Marketing Digital - Capítulo 4 – Otoño 2021)

Este funnel considera una etapa de averiguación entre la etapa de atractivo y la etapa de acción, en esta nueva etapa el potencial cliente busca información de manera activa sobre el producto o servicio e incluso puede llegar a registrarse en ciertos sitios especializados dejando registro de sus datos, lo lleva por lo tanto a la generación de un Lead. Es importante mencionar que estos funnel de compra solo entregan una guía teórica del proceso por el que pasa un cliente antes de generar una compra, en particular, para este trabajo de memoria estos funnel no tienen un gran impacto para los resultados finales obtenidos, ya que el enfoque corresponde a estudiar el CPC de un anuncio en cualquier etapa de este funnel. A pesar de esto último, se considera que el funnel más adecuado para este trabajo corresponde a la versión (4A), ya que no considera la etapa de generación de leads, la cual no es abordada en este trabajo de título. [23]

Teniendo en cuenta esta información de los funnel de compra, las plataformas por lo tanto deben responder la siguiente pregunta para poder determinar el desempeño de algún anuncio: ¿qué etapa dentro del funnel de compra es aquella que detona la conversión de un usuario? Para responder a esta última pregunta, es que los modelos de atribución entran en juego.

Un modelo de atribución permitirá a una plataforma poder atribuir logros de marketing a cada una de las etapas del funnel de compra, por ejemplo, a continuación, se muestra un tipo de modelo de atribución:



Figura 12: Modelo de atribución de ultima interacción

(Fuente: www.empresas.blogthinkbig.com/para-que-sirven-los-modelos-de-atribucion-google-analytics/)

En la figura se muestra que de una serie de eventos (exposiciones a publicidad), es la última exposición o el último anuncio que la persona vio el que se asume desencadenó o fue causante de la conversión. De esta manera, si se utiliza este modelo para atribuir causalidad de compra de un anuncio, aquellos que son dirigidos a etapas finales del proceso de compra (por ejemplo, los que son colocados ya en los sitios e-commerce de los clientes, es decir, en la etapa de acción dentro del funnel de compra), serán favorecidos con una mayor tasa de conversión y, por lo tanto, serán los que se mostrarán como más efectivos. Sin embargo, podría utilizarse también un modelo de atribución como el que se muestra en la siguiente imagen:



Figura 13: Modelo de atribución de primera interacción

(Fuente: www.empresas.blogthinkbig.com/para-que-sirven-los-modelos-de-atribucion-google-analytics/)

La cual indica que son las primeras exposiciones a publicidad las que desencadenan el proceso de compra. Por lo general los modelos de atribución que se utilizan no son los más extremos, por ejemplo, un modelo de atribución común es el mostrado en la siguiente imagen:



Figura 14: Modelo de deterioro en el tiempo

(Fuente: www.empresas.blogthinkbig.com/para-que-sirven-los-modelos-de-atribucion-google-analytics/)

Este modelo de atribución asume que, si bien los últimos anuncios son los que mayor impacto generan sobre la conversión, no trabaja bajo el supuesto de que los anuncios en etapas anteriores no tienen importancia en el proceso de compra, lo que es más realista en la mayoría de los casos.

Estos diferentes modelos de atribución jugarán un rol fundamental a la hora de comparar la efectividad de diferentes plataformas, ya que el uso de un modelo de atribución en particular puede llegar a sobreestimar o subestimar la efectividad real de la plataforma. Bajo esta misma línea, es importante considerar que no todas las métricas de desempeño de un anuncio se ven afectadas por estos modelos de atribución, ya que, por ejemplo, métricas como CTR, CPC o CPM no dependen del modelo de atribución de la plataforma.

4.4. Marco teórico asociado a algoritmos de aprendizaje automático

Los modelos de predicción desarrollados en este trabajo de título son todos correspondientes a algoritmos supervisados de aprendizaje automático, por lo que cada uno de estos tiene características particulares que lo definen, ventajas y desventajas. Esta sección tiene por finalidad describir de manera general las características de cada uno de estos algoritmos para entender a grandes rasgos su funcionamiento. Los modelos considerados se pueden dividir en las siguientes categorías:

- Algoritmos supervisados de clasificación
 - Decision Tree Classifier
 - Neural Network
 - RandomForestClassifier
 - XGBoost Classifier

- Algoritmos supervisados de predicción de variable continua
 - Decision Tree Regressor
 - XGBoost Regressor
 - Random Forest Regressor

4.4.1. Decision Tree Classifier

El funcionamiento de este algoritmo se basa en la elección de las mejores variables de entrada junto a su mejor valor para crear reglas de decisión que permitan subdividir el espacio de los datos hasta generar una heurística compuesta por una serie de reglas de decisión que permitirán clasificar nuevos registros en su clase más probable.

Un esquema que ejemplifica este funcionamiento puede encontrarse en la siguiente imagen:

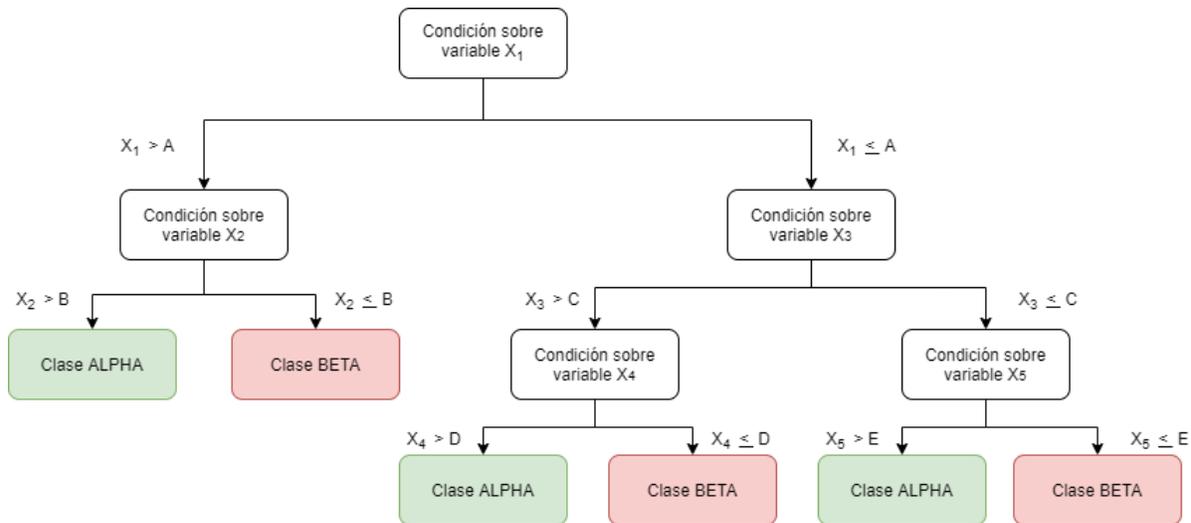


Figura 15: Esquema ejemplificativo del funcionamiento del algoritmo Decision Tree Classifier

(Fuente: elaboración propia)

El esquema muestra el resultado del entrenamiento del algoritmo a partir de una base de datos que contiene 5 covariables (X_1 , X_2 , X_3 , X_4 y X_5) y una variable de salida que determina la clase del registro (clase α o clase β). Se determina además que los umbrales de decisión elegidos corresponden a las constantes A, B, C, D y E, donde cada uno de estos umbrales determina la regla de división de algún nodo en particular del árbol. Este algoritmo, por lo tanto, permite a través de los datos históricos generar un "árbol" que permitirá determinar cuál será la clase de un nuevo registro.

Una de las principales ventajas de este modelo corresponde a su baja complejidad y por lo tanto también la fácil interpretabilidad de sus resultados, particularmente cuando la cantidad de covariables es pequeña. Por otro lado, otra de sus ventajas corresponde a que no es sensible a la escala de los datos que se usen como variables de entrada, y, por último, otra de sus ventajas corresponde a su facilidad de poder trabajar con distintos tipos de datos (categóricos, binarios, continuos, entre otros). En cuanto a sus desventajas principales se encuentran el hecho de que tiende por lo general a sobre ajustarse y además de que pequeños cambios en los datos producen alta varianza. [13]

4.4.2. Redes Neuronales

Las redes neuronales corresponden a algoritmos de aprendizaje automático que permiten clasificar registros por medio del uso y conexión de diferentes “neuronas”, las cuales corresponden, explicado de manera muy superficial, a operaciones matemáticas que son realizadas por diferentes funciones de activación. Existen diferentes tipos de redes neuronales, sin embargo, para efectos de este trabajo de título, se utilizan las conocidas como **redes neuronales profundas**. Este último tipo de red neuronal se caracteriza por usar una arquitectura como la mostrada en la siguiente figura:

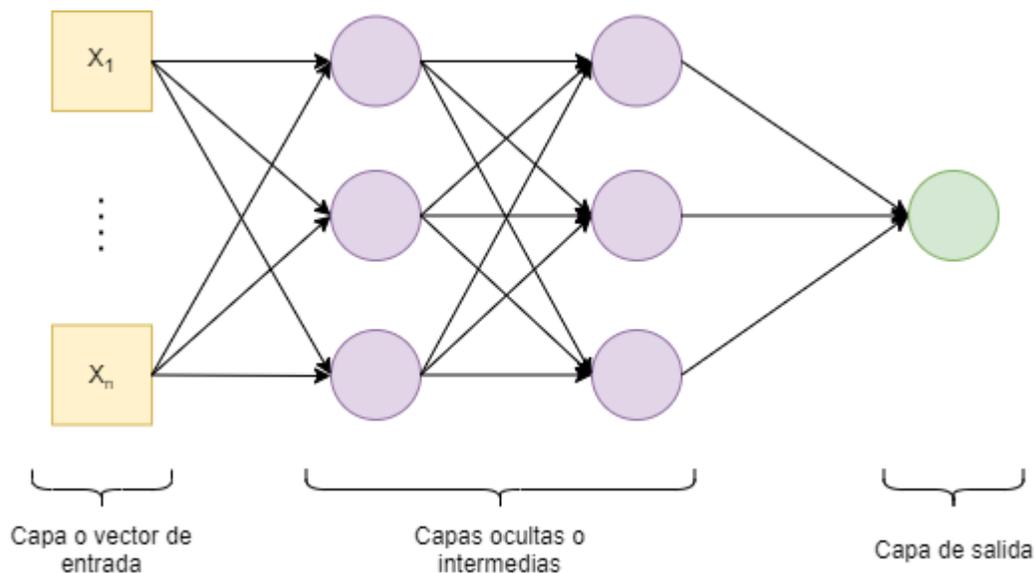


Figura 16: Ejemplo de estructura red neuronal profunda

(Fuente: elaboración propia)

La red de ejemplo mostrada en la imagen tiene una capa de entrada correspondiente a un vector de dimensión n , dos capas ocultas formadas por tres nodos o “neuronas” cada una, y finalmente, una capa de salida que corresponde a un vector con un único elemento. De manera general, una red neuronal profunda siempre tendrá al menos una capa de entrada, un conjunto de capas ocultas y una capa de salida.

El proceso de entrenamiento de este algoritmo consiste en el ajuste de los parámetros que determinan el cálculo que es realizado en cada nodo, todo esto por medio del procedimiento conocido como "backpropagation", el cual se basa principalmente en repartir el error encontrado de cada predicción entre los diferentes nodos, dándole una mayor relevancia a aquellos nodos que aportaron más en la determinación de ese resultado obtenido. De esta manera, el algoritmo es capaz de generar una estructura neuronal que logra modelar el comportamiento de los datos, y que, por lo tanto, permite la predicción de clase de nuevos registros con cierta probabilidad de éxito.

En cuanto a la estructura neuronal, deben establecerse los siguientes parámetros que son los encargados de darle forma:

- La cantidad de capas ocultas a utilizar
- La cantidad de nodos para cada una de las capas ocultas
- Las funciones de activación asociadas a cada una de las capas ocultas
- El tipo de capas ocultas a utilizar
- El orden en el que serán colocadas cada una de estas capas
- El tamaño de las submuestras que se utilizarán para entrenar la red o también conocidas como "batch size"

Un problema relativamente frecuente que se encuentra al utilizar este algoritmo tiene que ver con el posible sobreajuste de los datos en la etapa de entrenamiento del modelo. Esto provoca que el modelo sea muy preciso para predecir los registros con los cuales fue entrenado, sin embargo, que no sea preciso para predecir nuevos registros. Una manera de poder hacer frente a este problema es la incorporación de capas tipo "dropout" a la estructura, las cuales tienen la función de desactivar de manera aleatoria algún porcentaje de neuronas en el proceso de entrenamiento, evitando de esta manera el sobreajuste provocado por un aprendizaje muy ajustado a patrones particulares existentes en los datos de entrenamiento.

Una de las desventajas de este algoritmo es que es exigente computacionalmente, ya que la cantidad de parámetros a ajustar suele ser alta, en particular para estructuras con una gran cantidad de neuronas y capas. Por otro lado, otra desventaja es que la interpretabilidad de los resultados entregados es muy baja, ya que el gran número de operaciones que suelen realizarse dentro de la estructura neuronal complica demasiado la capacidad de entendimiento del algoritmo, por lo que, a los ojos de un usuario, su funcionamiento termina viéndose como una caja negra.

A pesar de esto, este algoritmo suele ser bastante bueno en cuando se entrena con grandes volúmenes de datos, y además debido a la combinación y la

posibilidad de uso de diferentes tipos de funciones, su forma de funcionar permite el uso de funciones no lineales, lo cual es una ventaja en contraste a otros algoritmos predictivos. [14]

4.4.3. Random Forest Classifier

Este algoritmo basa su funcionamiento en el algoritmo previamente descrito de Decision Tree Classifier, y puede considerarse como una modificación de este último algoritmo. La forma de funcionar del Random Forest Classifier consiste en generar diversas muestras aleatorias desde la base de datos original, donde para cada una de estas muestras, se hace funcionar el algoritmo de Decision Tree Classifier. De esta forma, se puede hacer una combinación de los errores de cada uno de los árboles generados para cada una de las muestras, generalizando de esta manera el resultado de predicción obtenido y disminuyendo por lo tanto la probabilidad de sobreajuste en el resultado de predicción final. [15]

4.4.4. XGBoost Classifier

XGBoost (Xtreme Gradient Boosting) es un popular algoritmo de aprendizaje supervisado, el cual puede ser usado para predecir registros de variables continuas por medio de su versión conocida como XGboost Regressor o también para la clasificación de registros a través de la versión conocida como como XGBoost Classifier.

A continuación, se muestra un esquema que gráfica de manera simplificada la forma general de funcionar del algoritmo (tanto para el problema de clasificación como para el de predecir variables continuas):

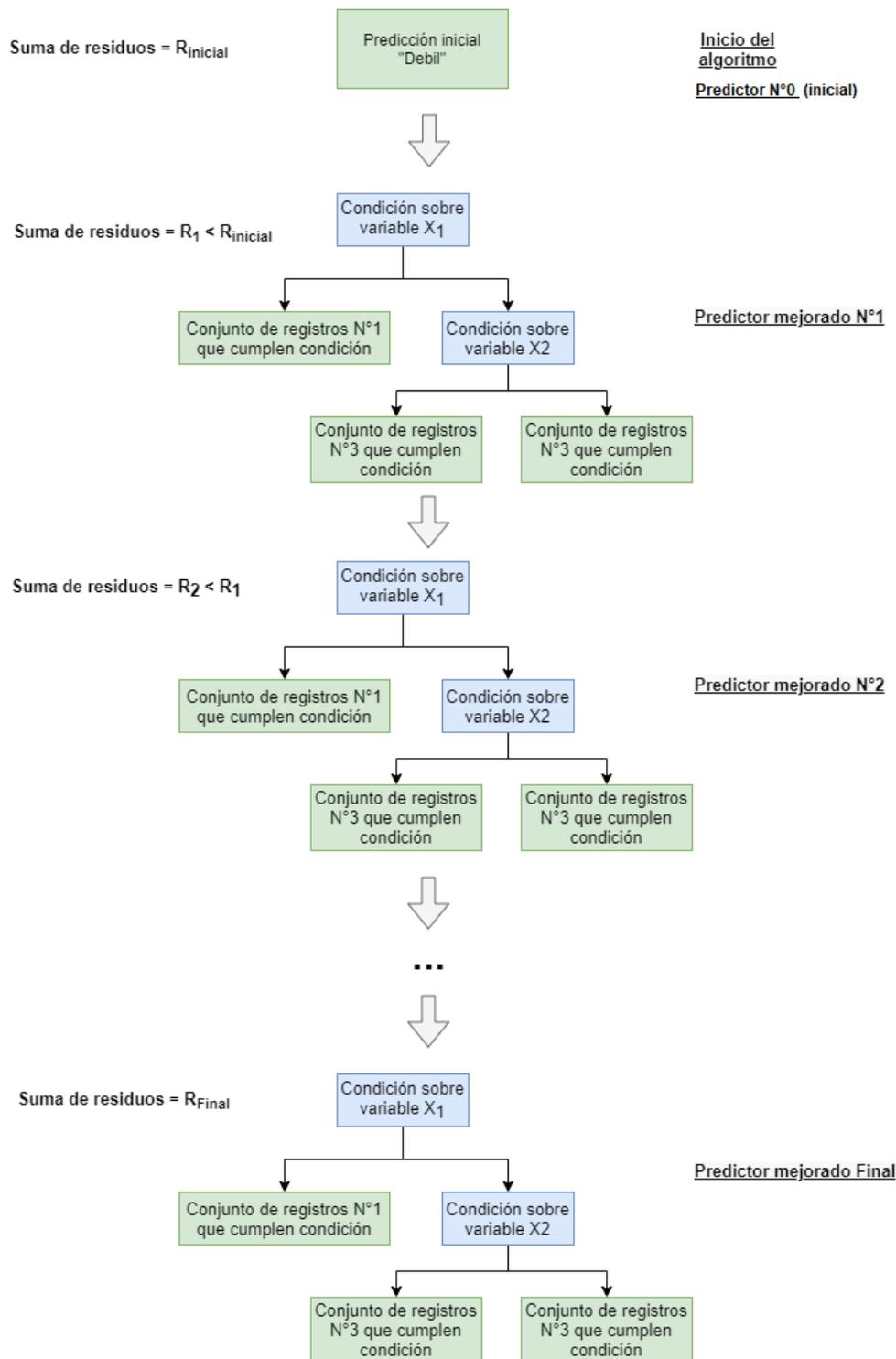


Figura 17: Esquema general algoritmo Xtreme Boost gradient

(Fuente: elaboración propia)

El funcionamiento del algoritmo se basa en mejorar iterativamente un predictor inicial, el cual suele ser un modelo de predicción trivial, por ejemplo, para un problema de clasificación binaria, podría ser un modelo que prediga el 50% de las veces una clase, y el otro 50% de las veces otra clase. Al predecir los registros por medio de este algoritmo inicial, se obtienen los llamados residuos de predicción, los cuales se entienden como una métrica de error predictivo correspondiente a la diferencia entre el valor obtenido y el valor predicho. El objetivo de los siguientes pasos del algoritmo es mejorar este predictor inicial reduciendo iterativamente la suma de los residuos obtenidos en cada una de estas iteraciones. El algoritmo avanza de esta forma hasta que se obtiene un valor de residuos muy bajo o hasta cuando se alcanza el límite de iteraciones establecido inicialmente.

En cuanto a la terminología relevante que se debe tener en cuenta para lograr un entendimiento a nivel general de este algoritmo, se encuentran en primer lugar los términos de similaridad y de ganancia.

En cada iteración se creará una estructura con forma de árbol, muy similar a la que se crea en el algoritmo de árboles de decisión, la cual guiará las predicciones según los valores de corte establecidos para algún conjunto de covariables. Cada hoja del árbol por lo tanto contendrá aquellos registros que cumplan con las condiciones que el árbol generado indica. En este contexto, en cada hoja se podrá obtener el factor de similaridad calculado utilizando los registros que componen esa hoja, el cálculo es realizado a través de la siguiente ecuación para el algoritmo de clasificación:

$$\text{Puntaje de similaridad} = \frac{\text{Suma de los residuos}^2}{\sum_i [P_i * (1 - P_i)] - \lambda} \quad (1)$$

Donde λ corresponde a un hiperparámetro de regularización y P_i corresponde a la probabilidad de que el registro i pertenezca a alguna clase en particular en la iteración anterior.

Este concepto de similaridad es importante ya que corresponde a un factor necesario para calcular la ganancia de una bifurcación dentro del árbol, la cual es calculada como:

$$Ganancia = Similaridad_{izquierda} + Similaridad_{derecha} - Similaridad_{Raiz} \quad (2)$$

La cual a su vez es utilizada por el algoritmo para determinar cuándo es eficiente crear una nueva división dentro del árbol, determinando de esta manera la estructura de árbol resultante que tendrá el algoritmo en esa iteración en particular.

Una vez creado el árbol de decisión para una iteración en particular, la predicción correspondiente para cada una de los registros se puede obtener de la siguiente forma:

$$Valor\ de\ predicción(X) = H_i(X) + \xi * H_{i+1}(X) \quad (3)$$

Donde $H_i(X)$ corresponde al valor del output para el registro X en la iteración i del algoritmo y ξ corresponde al learning rate.

En esta última ecuación, el valor del output de para cada registro en alguna iteración i se calcula como:

$$H_i(X) = \frac{\text{suma de residuos en la hoja correspondiente al registro } X}{\text{numero de residuos en la hoja correspondiente al registro } X + \lambda} \quad (4)$$

Lo que permite por lo tanto generar una nueva colección de residuos, obtenidos a partir de la predicción realizada utilizando este último árbol de decisión.

Por lo tanto, a modo de resumen, a través de este mecanismo es posible en cada iteración generar un "árbol de decisión" que permita predecir los registros obteniendo una suma de residuos menor que la obtenida utilizando el predictor de la iteración anterior, llegando de esta manera, a un buen predictor que fue construido mejorando de manera consecutiva el predictor inicial. [16][18]

4.4.5. Decision Tree Regressor

El funcionamiento de este algoritmo es en general equivalente al funcionamiento del Decision Tree Classifier, el cual fue explicado en detalle en la sección 4.3.1 de este informe. El Decision Tree Regressor logra diferenciarse principalmente por el hecho de que está construido con el objetivo de predecir variables continuas, y no categóricas como lo es para el algoritmo de clasificación.

A nivel de funcionamiento del algoritmo, esta diferencia se aprecia en la parte posterior a la construcción de la estructura del árbol de clasificación, en el momento en el cual debe establecerse la respuesta de predicción del algoritmo por medio de los registros de entrenamiento que fueron agrupados en cada una de las hojas del árbol. En este contexto, el algoritmo de Decision Tree Regressor establecerá que el resultado de predicción de un nuevo registro corresponderá al promedio del valor de los registros de entrenamiento que se encuentran en la hoja correspondiente a donde cae el nuevo registro, mientras que, por otro lado, el modelo de clasificación, dirá que el resultado de predicción corresponderá a la moda de los registros de entrenamiento que pertenecer a la hoja correspondiente del nuevo registro. [17]

4.4.6. Random Forest Regressor

El funcionamiento de este algoritmo es fácil de entender si se tienen en cuenta las explicaciones ya entregadas de los algoritmos de Decision Tree Regressor y Random Forest Classifier, esto último ya que la lógica de este algoritmo es equivalente a la de Random Forest Classifier, con la única diferencia que los árboles de decisión construidos en cada elección de muestra aleatoria, no corresponden a árboles de clasificación, si no que a árboles de regresión.

De esta manera, en pocas palabras, se puede entender el algoritmo de Random Forest Regressor como un algoritmo de Random Forest Classifier, pero cambiando los árboles de decisión de clasificación por arboles de decisión de predicción de variable continua.

4.4.7. XGBoost Regressor

Este algoritmo presenta en general la misma forma de funcionar que el XGBoost Classifier, sin embargo, difiere en el hecho de que la variable de respuesta es continua, y no categórica como lo es en el algoritmo de clasificación.

La diferencia a nivel de algoritmo se encuentra principalmente en la forma en la cual se calcula el puntaje de similaridad, a diferencia de cómo se obtiene para el problema de clasificación, en el caso de **predicción de variable continua** se calcula por medio de la siguiente ecuación [18]:

$$\text{Puntaje de similaridad} = \frac{\text{Suma de los residuos}^2}{\text{Cantidad de registros en la hoja} - \lambda} \quad (5)$$

4.5. Métricas de evaluación de desempeño de modelos predictivos

El objetivo de esta sección es describir aquellas métricas de desempeño que serán utilizadas a lo largo del trabajo para determinar la precisión de los distintos modelos predictivos construidos.

4.5.1. Accuracy

Esta métrica de desempeño es utilizada en el contexto de evaluar la precisión de un modelo predictivo de clasificación. Esta indica el porcentaje de registros que fueron predichos correctamente por el modelo predictivo, por lo que se puede calcular a través de una ecuación como la que se muestra a continuación:

$$Accuracy = \frac{\text{Cantidad de registros predichos correctamente}}{\text{Cantidad total de registros predichos}} \quad (6)$$

O también, escrita en términos de una matriz de confusión, puede expresarse de la siguiente manera:

$$Accuracy = \frac{VP + VN}{VP + VN + FP + FN} \quad (7)$$

Donde VP corresponde a la cantidad de verdaderos positivos, VN a la de verdaderos negativos, FP a la de falsos positivos y FN a la de falsos negativos.

El accuracy es una de las métricas más generales para determinar el desempeño de un algoritmo de clasificación, sin embargo, dependiendo de la importancia que le dé el investigador a las predicciones positivas o negativas, podría utilizar otras métricas de desempeño como lo son el recall o la precisión, sin embargo, debido a que en este trabajo no se presenta una necesidad por darle importancia predictiva a alguna clase en particular, se determina el accuracy como una métrica adecuada a utilizar.

4.5.2. Curva ROC (AUC)

Esta métrica de desempeño es utilizada también en el contexto de evaluar la precisión de modelos predictivos de clasificación. Tiene la ventaja de ser una buena métrica para comparar el desempeño de diferentes algoritmos de clasificación, ya que evalúa la capacidad de predicción considerando todos los escenarios de umbrales de predicción que pueden establecerse. En este contexto, el concepto de umbral de predicción hace referencia a aquel umbral que definirá a un registro como clase positiva o negativa a través de la probabilidad de pertenecer a una clase. Un ejemplo de umbral de predicción podría ser el establecer que un registro deba alcanzar una probabilidad de 0.7 o más de pertenecer a clase positivo para clasificarlo como tal, en caso de no cumplirse con esa condición, el registro será predicho como negativo.

La curva ROC se puede representar tal como se muestra en la siguiente figura:

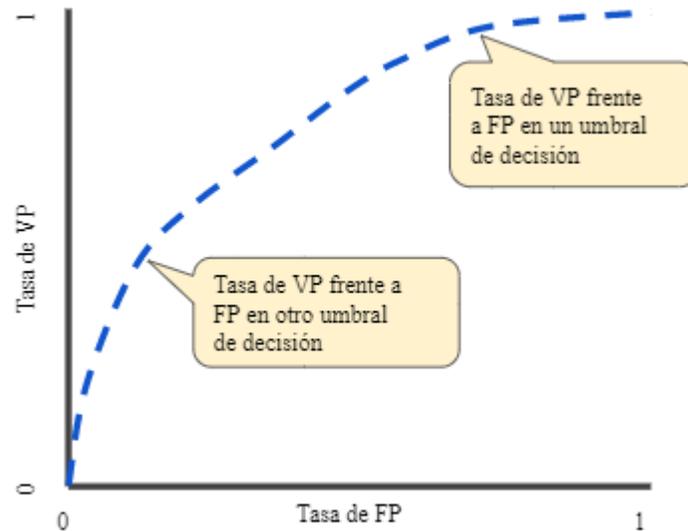


Figura 18: Curva ROC

(Fuente: <https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc?hl=es-419>)

Donde la tasa de verdaderos positivos (TPR) y la tasa de falsos positivos (FPR) se calculan como se muestra a continuación:

$$\text{Tasa de VP (TPR)} = \frac{VP}{VP + FN} \quad (8)$$

$$\text{Tasa de FP (FPR)} = \frac{FP}{FP + VN} \quad (9)$$

Esta curva es construida a partir de la tasa de VP y tasa de FP obtenidas al predecir un conjunto de testeo para cada uno de los umbrales posibles, de esta manera, y como se muestra en la figura, cada punto que forma la curva ROC representa el resultado de predicción de un umbral en particular representado en términos de tasa de VP y tasa de FP.

Un concepto importante que se desprende este análisis es el conocido como AUC o área bajo la curva ROC. Este último valor es relevante en los análisis ya que mide la calidad predictiva de un modelo considerando todos los posibles umbrales que se pueden establecer, de esta forma, el valor del AUC permite comparar la capacidad predictiva de distintos modelos de manera general. Esto último es particularmente relevante para este trabajo, ya que como se mencionaba anteriormente uno de los objetivos principales es comparar el desempeño de diferentes algoritmos predictivos.

Por último, el valor que puede tomar el AUC va desde 0 hasta 1, donde 0 significa un modelo que tiene muy mal resultado predictivo y 1 un resultado predictivo perfecto. Es importante mencionar también que un valor de AUC igual a 0.5 representa una predicción aleatoria. [19]

4.5.3. MAE (Mean Squared Error)

El MAE (Mean Absolute Error) corresponde a una métrica de desempeño de modelos predictivos de variable continua. Esta se caracteriza por ser una de las más simples de entender y de calcular, su ecuación se presenta a continuación:

$$MAE = \frac{1}{n} \sum |e_t| \quad (10)$$

Donde n corresponde a la cantidad de registros predichos y e_t error de la predicción entendido como la diferencia entre el valor real y el valor predicho.

Esta métrica se puede entender como el error absoluto promedio que se obtiene al predecir un conjunto de registros. A diferencia de otras métricas de medición tales como el MAPE, esta se caracteriza por darle igual importancia a todos los errores obtenidos, indistintamente si el error es obtenido en predicciones que tienen un alto o bajo valor base, es decir, no normaliza por el valor real de la predicción. Por otro lado, tampoco les da una mayor importancia a errores más grandes por sobre los errores pequeños, como si lo hace el RMSE, en cual se explica en la sección inmediatamente posterior. [20]

4.5.4. RMSE (Root Mean Squared Error)

El RMSE corresponde también a una métrica de desempeño de modelos predictivos de variable continua. Este se puede obtener a través de la siguiente ecuación de cálculo:

$$RMSE = \sqrt{\frac{1}{n} \sum e_t^2} \quad (11)$$

Donde n corresponde a la cantidad de registros predichos y e_t error de la predicción entendido como la diferencia entre el valor real y el valor predicho, al igual que para el MAE.

Una de las características principales de esta métrica tiene que ver con el hecho de que los errores son elevados al cuadrado dentro de su cálculo, operación que provoca que los errores que en magnitud son más grandes se amplifiquen más que lo que se amplifican los errores pequeños en magnitud. De esta manera, el RMSE es una buena métrica de desempeño cuando los errores grandes son particularmente perjudiciales para el investigador. [20]

5. Metodología

5.1. Metodología general utilizada: CRISP-DM

La metodología que se utiliza en general para este trabajo de memoria, corresponde a la llamada CRISP-DM (Cross-Industry Standard Process for Data Mining), la cual permite ordenar el trabajo en una serie de etapas, las cuales se muestran en el siguiente esquema:

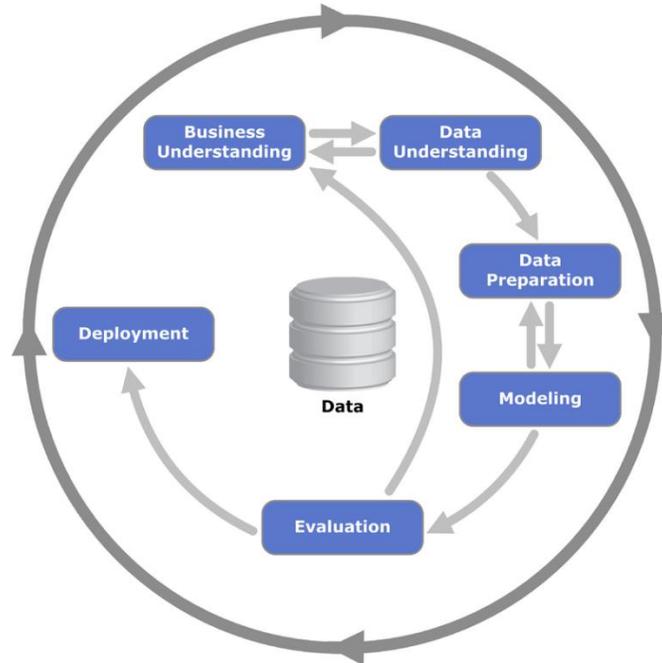


Figura 19: Metodología CRISP-DM

(Fuente: <https://www.datascience-pm.com/crisp-dm-2>)

La principal ventaja del uso de esta metodología corresponde a la posibilidad de iterar entre cada una de sus etapas, ya que permite la implementación de retroalimentación a partir de los resultados que se van obteniendo, permitiendo mejorar de manera constante los resultados hasta llegar a la mejor versión de los modelos.

La primera etapa, de conocimiento del negocio, tiene por finalidad el entender el contexto actual en el cual participa la agencia, teniendo en cuenta aspectos del mercado relevante, servicios entregados, perfil de los clientes, métodos de pricing, entre otras variables. Un aspecto clave en esta etapa, corresponde al entender cómo el problema en cuestión afecta a la agencia, y también, los beneficios que entregará el trabajo de memoria a la empresa.

Las siguientes cuatro etapas de esta metodología, correspondientes a el entendimiento de los datos, preparación de los datos, modelación y evaluación, forman parte del proceso del desarrollo de los modelos predictivos. Estas etapas se consideran cíclicas entre sí, ya que, una vez se llega a la etapa de evaluación de los modelos, es posible volver a comenzar desde el entendimiento del negocio u otra etapa posterior a esta, permitiendo

incorporar los aprendizajes obtenidos hasta el momento, e iterar el proceso mejorando el desempeño en cada iteración.

5.2. Metodología en detalle utilizada para el desarrollo de modelos de aprendizaje automático

Los algoritmos predictivos que se evalúan en este trabajo de título se pueden dividir en dos grandes grupos, en primer lugar, un conjunto de algoritmos de clasificación y, por otro lado, un conjunto de algoritmos de predicción de variable continua, en ambos casos, corresponden a algoritmos supervisados derivados del Machine Learning. Mas concretamente, los algoritmos que se evalúan son los siguientes:

- Algoritmos supervisados de clasificación
 - Decision Tree Classifier
 - Neural Network
 - Random Forest Classifier
 - XGBoost Classifier

- Algoritmos supervisados de predicción de variable continua
 - Decision Tree Regressor
 - XGBoost Regressor
 - Random Forest Regressor

El principal objetivo a evaluar para cada uno de estos algoritmos corresponde a la precisión con la cual puedan predecir el CPC de un anuncio digital publicado a nivel semanal.

5.2.1. Metodologías generales para todos los algoritmos

Se utilizan técnicas metodológicas que forman parte de la etapa previa al desarrollo de los modelos predictivos, ya que estas forman parte del proceso de transformación, selección y preprocesamiento de los datos que posteriormente serán utilizados como materia prima para el entrenamiento de los diferentes modelos. A continuación, se describen en detalle cada una de estas técnicas:

- Se convierten a variables binarias (dummies) aquellas variables categóricas que se utilizarán como variables independientes, estas variables corresponden a el cliente asociado de la publicación, la plataforma de publicación, el dispositivo de impresión y los diferentes eventos temporales que caracterizan a la publicación. Cada uno de estos eventos temporales son descritos en detalle en la sección 6.1 de este documento
- Se normaliza la variable de costo de publicación del anuncio a través de un escalamiento de mínimos y máximos
- Para todos los modelos desarrollados, la predicción de desempeño es realizada a partir de datos agregados a nivel semanal, por lo tanto, cada uno de los registros de la base de datos corresponde a un resumen de estadísticas semanales de cada una de los anuncios publicados por la agencia
- La variable de desempeño a predecir por cada uno de los algoritmos (variable dependiente) corresponde al CPC que alcanza un anuncio a nivel semanal. Los algoritmos de clasificación tienen por objetivo el predecir si un anuncio tendrá un CPC que estará sobre o bajo un umbral establecido, y los algoritmos de predicción variable continua tienen por objetivo predecir cual será el valor del CPC alcanzado por ese anuncio esa semana
- Para los modelos de clasificación, por lo tanto, se crea una variable binaria que representa si el CPC esta sobre o bajo el umbral establecido, en concreto, esta variable tomará el valor de 1 si el CPC resultante es menor al umbral establecido y toma el valor 0 en caso contrario

- Cada uno de los algoritmos utilizados requiere para su funcionamiento seleccionar una serie de hiperparámetros que determinan la capacidad de predicción de los modelos. Para elegir el valor de cada uno de estos es que se utiliza la metodología conocida como búsqueda de grilla o “grid search”, la cual consiste en automatizar el proceso de probar la efectividad de cada una de las combinaciones posibles que pueden existir de hiperparámetros, encontrando de esta manera la combinación que provoca el mejor desempeño predictivo del algoritmo. Es necesario considerar el hecho de que esta técnica requiere en algunos casos una gran capacidad computacional para funcionar, lo que puede restringir la cantidad de combinaciones que pueden testearse a través de esta metodología. Los detalles de implementación de esta técnica se pueden encontrar en el anexo 11.3 de este informe, donde se muestran los distintos conjuntos de hiperparámetros considerados para cada algoritmo, a excepción del algoritmo de redes neuronales, en el cual se usa una metodología diferente de optimización de hiperparámetros la cual es descrita en la sección 5.2.2 de este informe
- La evaluación de desempeño de cada uno de los modelos se realiza sobre una base de testeo generada aleatoriamente, la cual corresponde a un 25% del total de los registros disponibles en la base de datos, lo cual quiere decir que el entrenamiento de los modelos es realizado considerando el 75% de los datos disponibles
- La manera de medir el desempeño varía según se trate de modelos de clasificación o modelos de predicción de variable continua. Para los algoritmos de clasificación, estos son evaluados según su desempeño medido a través de principalmente el accuracy y AUC alcanzado. En cuanto a los algoritmos de predicción de variable continua, se evalúa su capacidad de predicción por medio de las métricas de desempeño del MAE y RMSE.

5.2.2. Metodología particular para el desarrollo de modelos de redes neuronales

Un caso particular de algoritmo predictivo corresponde a las redes neuronales, que no solo requieren establecer una combinación de hiperparámetros, sino

que también una estructura de red neuronal para su funcionamiento. El establecimiento de esta estructura se logra por medio de un proceso de estructuración manual. El punto de partida de esta corresponde a una red neuronal básica, la cual es mostrada en la siguiente figura:

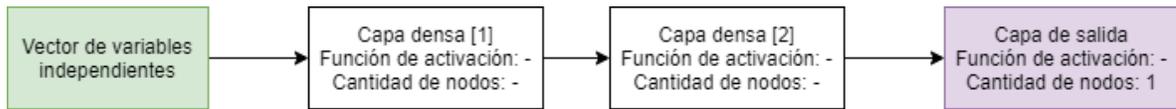


Figura 20: Estructura inicial de la red neuronal utilizada para clasificar (los guiones representan los hiperparametros que deben se optimizados por medio del uso de búsqueda de grilla)

(Fuente: elaboración propia)

El procedimiento para llegar a la estructura final corresponde a la prueba y error, donde se agregan, eliminan y reordenan capas en cada iteración, cada una de estas iteraciones es evaluada por medio del accuracy que logra esa estructura neuronal en la predicción de una base de testeo generada aleatoriamente, de esta manera, la estructura neuronal que alcanza un mayor accuracy y que por tanto es elegida finalmente, es la que se muestra en la siguiente figura:

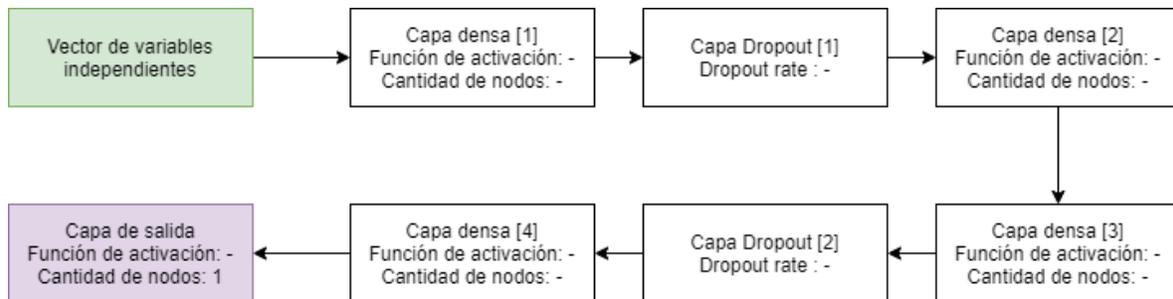


Figura 21: Estructura final de la red neuronal utilizada para clasificar (los hiperparametros a ser ajustados se muestran a través de un "-")

(Fuente: elaboración propia)

Una vez definida esta estructura base, comienza el proceso de optimización de hiperparámetros a través del procedimiento de búsqueda de grilla, sin embargo, es implementado de manera diferente en este algoritmo ya que la exigencia de capacidad de procesamiento llega a superar la capacidad computacional disponible para la realización de este trabajo de título.

La forma de optimizar hiperparámetros se hace en este caso a través de un procedimiento de búsqueda de grilla "por partes", de esta forma, se disminuye la cantidad de combinaciones que deben ser evaluadas, permitiendo que la optimización de hiperparámetros sea factible para este algoritmo. El concepto de "por partes" significa en la práctica que se comienza a partir de seleccionar un valor inicial para cada uno de los hiperparámetros, luego se optimiza por medio de búsqueda de grilla solo un subconjunto de hiperparámetros, dejando el resto constantes, por lo que una vez realizada esta optimización, se fijan los hiperparámetros ya optimizados y se sigue optimizando otro subconjunto de hiperparámetros dejando constante los demás, incluyendo a el conjunto que acaba de ser optimizado, este procedimiento sigue hasta haber pasado por todos los hiperparámetros. Los resultados en detalle y los hiperparámetros considerados se muestran en el anexo 11.4 de este informe.

5.2.3. Herramientas y tecnologías utilizadas para el desarrollo de los modelos

En cuanto a las tecnologías que se utilizan para el desarrollo de estos modelos se considera el lenguaje de programación Python junto con el uso de sus diversas librerías de ciencia de datos, entre las más importantes se encuentran la librería Pandas, Sckit-Learn, Numpy, Tensorflow, entre otras. Para la limpieza y manejo de datos también se utilizan estas tecnologías recién mencionadas, pero, además, se complementan con el uso del software Excel y Tableau, este último también utilizado principalmente para la visualización y entendimiento de los datos.

6. Entendimiento de los datos

6.1. Descripción de los datos disponibles

Las métricas de desempeño de las publicaciones realizadas históricamente por la agencia son registradas por las mismas plataformas en las cuales se publican estos anuncios, para efectos de este trabajo de memoria, los datos serán obtenidos desde Facebook ADS y Google ADS, los cuales registran los datos de las plataformas en las cuales la agencia realiza sus publicaciones.

Los datos de estas plataformas no son entregados de manera directa por estas tecnologías, sino que es necesario utilizar un software ETL (Extract, Transform and Load) para obtener un base de datos completa con los registros históricos de desempeño. De esta manera, una vez realizado el proceso de extracción, se obtiene una base de datos que posee las siguientes características:

- Se obtienen 17.532 registros de desempeño semanal, donde cada registro corresponde a el desempeño de una publicación en una semana, plataforma, y dispositivo de impresión en particular
- Estos datos corresponden a las campañas históricas que comprenden desde inicios del 2018 hasta finales del diciembre del año 2020
- Los datos consideran la información de los 4 clientes de e-commerce más importantes de la empresa
- Los datos consideran la información de las 7 plataformas elegidas y que forman parte del alcance de este trabajo de título
- Las variables independientes que se tienen de cada publicación son:
 - La semana del año en la que se realizó la publicación
 - Los días de la semana en los cuales la publicación estuvo en línea
 - Un contador que indica el número de semanas que el anuncio ya ha sido publicado en el pasado
 - El cliente asociado a la publicación
 - La plataforma de publicación
 - El dispositivo de impresión del anuncio
 - Indicador que muestra la cantidad adicional de plataformas que son utilizadas por anuncios de un mismo cliente en la semana correspondiente del registro
 - Un conjunto de indicadores que dan cuenta si el anuncio fue publicado en algún evento temporal especial como puede ser

época navideña, época de Cyberday, día de los enamorados, entre otras épocas relevantes, las cuales se describen más adelante

- Las variables dependientes (métricas de desempeño) que se tienen de cada publicación son:
 - Cantidad de clics logrados
 - Cantidad de impresiones realizadas
 - Costo de publicación del anuncio
 - CTR
 - CPC
 - CPM
 - ROAS

En relación al detalle de los indicadores de eventos estacionales, estos son los siguientes:

- Indicador de época navideña: se activa cuando el anuncio es publicado algún día entre el 4 de diciembre y 24 de diciembre de algún año
- Indicador de día del niño: se activa cuando el anuncio es publicado algún día entre el 16 de julio y el 5 de agosto para el año 2018, entre el 23 de julio y el 11 de agosto para el año 2019 y entre el 27 julio y el 16 de agosto para el año 2020
- Indicador del día de los enamorados: se activa cuando el anuncio es publicado algún día entre el 26 de enero y el 14 de febrero de algún año
- Indicador del día de la madre: se activa cuando el anuncio es publicado algún día entre el 24 de abril y el 13 de mayo para el año 2018, entre el 23 de abril y el 12 de mayo para el año 2019 y entre el 21 de abril y el 10 de mayo para el año 2020
- Indicador del día del padre: se activa cuando el anuncio es publicado algún día entre el 30 de mayo y el 17 de junio para el año 2018, entre el 29 de mayo y el 16 de junio para el año 2019 y entre el 2 de junio y el 21 de junio para el año 2020
- Indicador de época de vuelta a clases: se activa cuando el anuncio es publicado algún día entre el 9 de febrero y el 7 de marzo de algún año
- Indicador de primer retiro de las AFP: se activa cuando el anuncio es publicado algún día entre el 4 de julio y el 9 de agosto del año 2020

- Indicador de segundo retiro de las AFP: se activa cuando el anuncio es publicado algún día entre el 4 de diciembre del 2020 y el 9 de enero del 2021
- Indicador de COVID-19: se activa cuando el anuncio es publicado algún día posterior al 1 de marzo del año 2020
- Indicador de época cyber, se activa cuando el anuncio es publicado algún día dentro de los siguientes intervalos de tiempo:
 - Entre el 15 de mayo y el 30 de mayo del año 2018
 - Entre el 25 de septiembre y el 10 de octubre del año 2018
 - Entre el 15 de mayo y el 29 de mayo del año 2019
 - Entre el 20 de noviembre y el 2 de diciembre del año 2019
 - Entre el 12 de mayo y el 25 de mayo del año 2020
 - Entre el 24 de septiembre y el 10 de octubre del año 2020
 - Entre el 14 de noviembre y el 27 de noviembre del año 2020

Es importante además explicar el porqué se incorpora dentro de las variables independientes el indicador que muestra la cantidad adicional de plataformas que son utilizadas por anuncios de un mismo cliente en alguna semana en particular. Esto último nace de la necesidad de modelar el hecho de que cuando un anuncio es publicado en más de una plataforma al mismo tiempo, la publicación en una plataforma podría alterar el resultado de la publicación en la otra plataforma. Actualmente la empresa no cuenta con una disponibilidad de datos para poder modelar este fenómeno a nivel de anuncio en particular, por lo que se modela a partir de una aproximación, la cual trabaja bajo el supuesto de que los anuncios publicados de un mismo cliente son similares entre si independiente de su contenido. De esta forma se crea esta variable que indica cuantas plataformas adicionales a la que hace referencia el registro son utilizadas por ese cliente en esa semana en particular, de esta manera, como en este trabajo se considera un total de 7 plataformas diferentes, este indicador toma valores enteros que van desde 1 hasta 6. Adicionalmente, este fenómeno es de relevancia de modelar ya que en la mayoría de los casos la agencia publicita anuncios de un cliente en particular en diferentes plataformas al mismo tiempo. Esto último se puede evidenciar a través de la siguiente tabla:

Tabla 4: Datos relacionados a la simultaneidad de uso de plataformas para un mismo cliente

Cliente	Porcentaje de anuncios que son publicados en simultaneo con anuncios del mismo cliente, pero en plataformas diferentes
Cliente 1	93.46%
Cliente 2	80.65%
Cliente 3	81.63%
Cliente 4	4.17%

(Fuente: elaboración propia)

Por otro lado, un análisis muy relevante para este trabajo de tesis tiene que ver con la disponibilidad de variables que tiene la agencia para modelar el problema y, por lo tanto, para predecir el desempeño de un anuncio publicado de manera digital. En la sección 4.1 del marco teórico de este informe se detallan 9 efectos que en teoría deben ser considerados para modelar correctamente el desempeño de una publicación digital, en este sentido, la empresa cuenta con información para modelar todos estos efectos a excepción de dos, los cuales son:

1. Efecto de contenido
2. Efecto de la competencia

De esta manera, el trabajo entonces se enmarca en el contexto de evaluar la capacidad de predicción de los modelos teniendo en consideración que no se cuenta con la información relacionada al contenido del anuncio ni tampoco del contexto de competitividad existente al momento de la publicación.

6.2. Análisis descriptivo de datos

Una primera información relevante para comenzar la construcción de modelos predictivos del CPC corresponde al saber cómo ha sido la distribución de esta métrica a lo largo de la historia considerada, el siguiente gráfico corresponde a un histograma que nos muestra esta distribución:

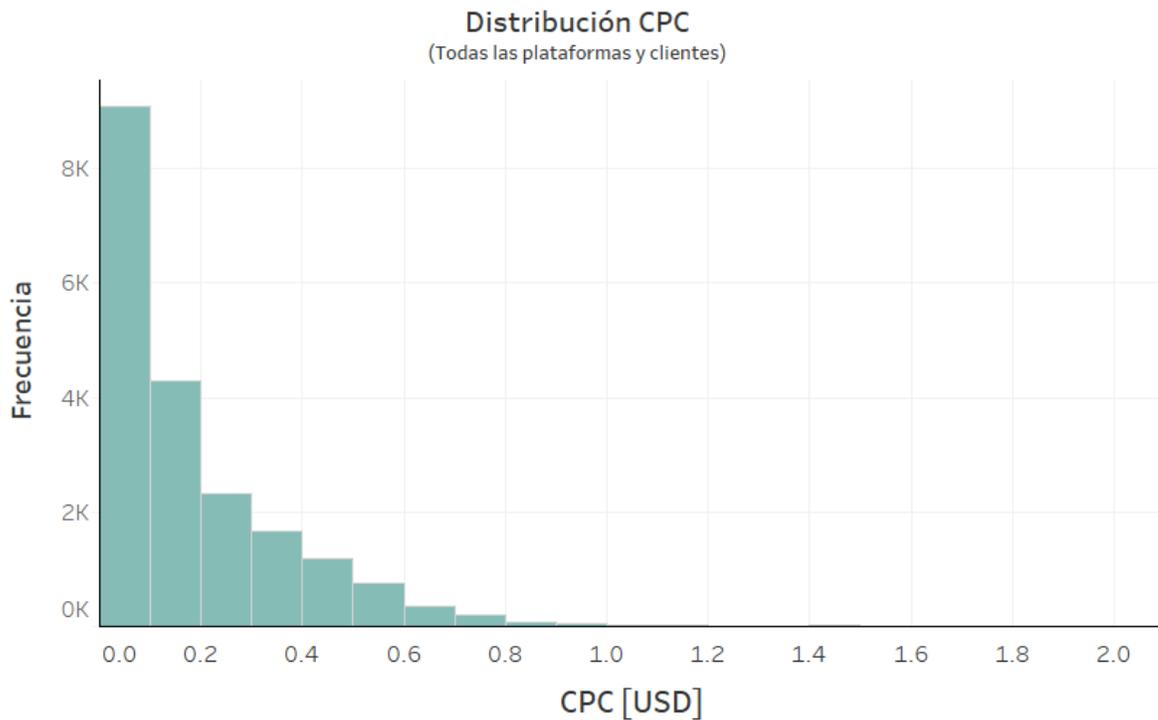


Figura 22: Magnitud histórica del CPC agregado a nivel semanal

(Fuente: elaboración propia)

De esta grafica se puede ver que la mayoría de los resultados a nivel semanal terminan teniendo un costo por clic menor a 1 [USD], en particular, una gran cantidad de estos se encuentran acumulados en valores menores a 0.5 [USD].

Otra información relevante tiene que ver con la posible heterogeneidad de resultados al analizar por diferentes clientes y también diferentes plataformas,

a continuación, se muestran dos gráficas que permiten observar estos comportamientos:

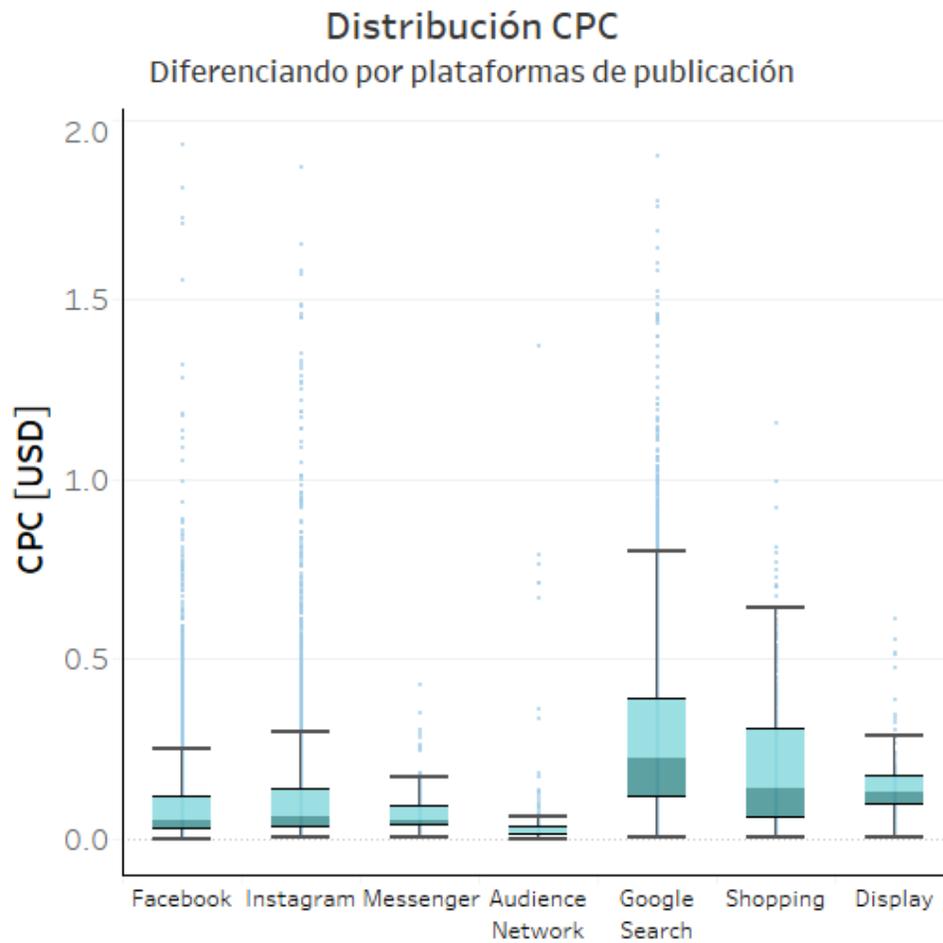


Figura 23: Distribución del CPC para cada una de las plataformas de publicación

(Fuente: elaboración propia)

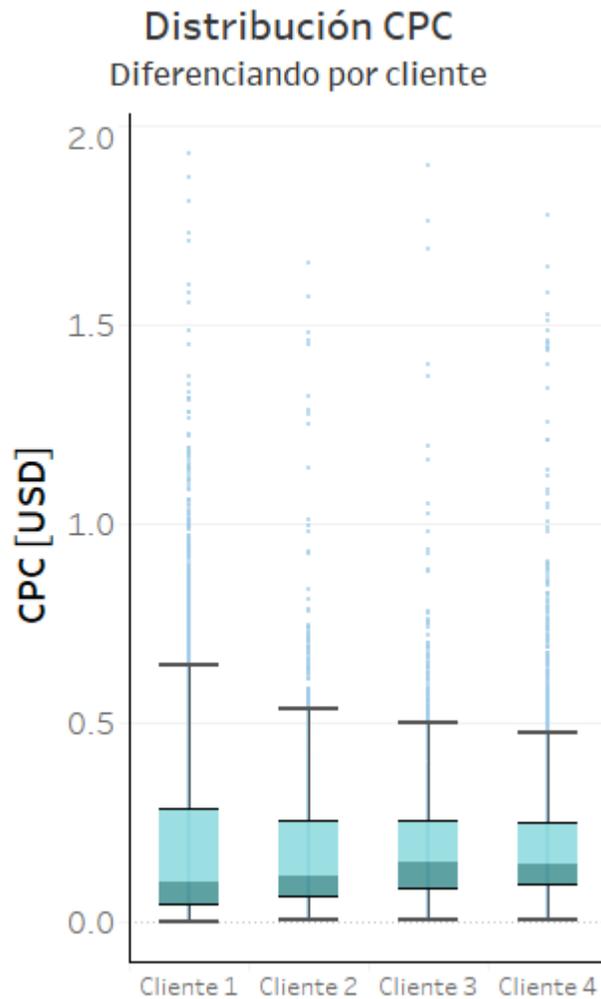


Figura 24: Distribución del CPC para cada una de los clientes considerados
(Fuente: elaboración propia)

En la figura 22, se puede observar que a nivel histórico las plataformas de Google ADS han tenido un CPC mayor en comparación a las plataformas de Facebook ADS, en particular aquellas plataformas que funcionan en base a motores de búsqueda como lo son Google Search o Google Shopping. Esto último es relevante ya que como se verá más adelante, influirá de manera directa en la capacidad de predicción de los modelos de clasificación que requieren el establecer un umbral de predicción que determina que clase es positiva y que clase es negativa.

Por otro lado, en la figura 23 se puede ver que los resultados a nivel de CPC no varían de manera significativa cuando se analiza para cada cliente en particular, lo que sugiere que no existe evidencia de una alta heterogeneidad entre los comportamientos de cada uno de los clientes y que, por lo tanto, es baja la probabilidad de que este sea un factor que aporte ruido a las predicciones realizadas por los modelos.

Otra información relevante que hay que tener en cuenta tiene que ver con la cantidad de información que existe para cada una de las plataformas. A continuación, se muestra la cantidad de registros a nivel semanal que se tiene para cada una de estas:

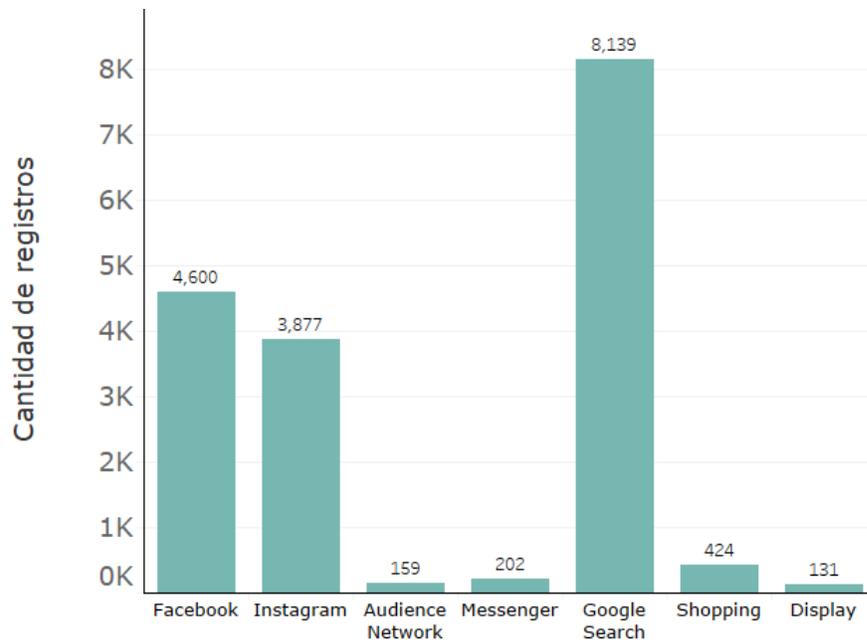


Figura 25: Cantidad de registros disponibles para cada una de las plataformas

(Fuente: elaboración propia)

De aquí se puede apreciar que la gran mayoría de los registros pertenecen a las plataformas de Facebook, Instagram y Google Search. En comparación a estas últimas, el resto de las plataformas tienen una cantidad de registros mucho menor, lo cual podría eventualmente alterar la capacidad de predicción de los modelos para estas plataformas. Esto último es parte de la evaluación en la etapa de desarrollo de los modelos predictivos.

7. Desarrollo y evaluación de modelos predictivos

7.1. Modelos de clasificación

La primera decisión clave que se debe tomar para comenzar con el desarrollo de estos modelos tiene que ver con cuál será el umbral de predicción que permitirá diferenciar entre clase positiva o negativa. En este trabajo se estudian distintos umbrales, los cuales corresponden a valores límites del CPC, donde a magnitudes mayores de ese valor límite se categorizará el resultado como de clase negativa (resultado insuficiente) y para el caso contrario como clase positiva (resultado suficientemente bueno). El objetivo será entonces la evaluación de distintos modelos de predicción para distintas elecciones de este umbral, en detalle los umbrales a evaluar serán los siguientes:

1. Un umbral de predicción único para todas las plataformas correspondiente a 0.14 [USD]
2. Un umbral de predicción diferenciado para cada una de las plataformas, el cual corresponderá a la mediana histórica del CPC para cada una de estas

Tabla 5: Umbrales de predicción establecidos tomando como referencia la mediana histórica de CPC para cada plataforma

Plataforma	Umbral de predicción (Mediana histórica de CPC) [USD]
Facebook	0.06
Instagram	0.087
Messenger	0.051
Audience Network	0.011
Google Search	0.237
Shopping	0.153
Display	0.132

(Fuente: elaboración propia)

Es importante mencionar que el umbral de 0.14 [USD] no es elegido arbitrariamente, ni tampoco tiene que ver con los resultados históricos de la agencia, sino que es establecido a través de un benchmark realizado a nivel de industria que considera información de la región en la que participa la agencia, industria de las empresas cliente consideradas y, por último, de las plataformas que son consideradas en este trabajo de título. [5][6][7]

7.1.1. Primer modelo: clasificación de CPC por medio de un umbral único para todas las plataformas

Este modelo tiene como finalidad predecir si es que una publicación a nivel semanal tendrá un resultado de CPC mayor o menor a 0.14 [USD]. La clase positiva en este caso corresponde a un CPC menor o igual a 0.14 [USD] y la clase negativa a un CPC mayor a 0.14 [USD]. La siguiente tabla muestra los resultados generales de cada uno de los diferentes algoritmos evaluados:

Tabla 6: Resultados generales de predicción para el primer modelo evaluado

Algoritmo de predicción	Accuracy	AUC
Decision Tree	78.78%	0.88
Neural Network	81.09%	0.80
Random Forest	77.39%	0.85
XGBoost Classifier	83.60%	0.92

(Fuente: elaboración propia)

A partir de estos resultados es que se elige el algoritmo XGBoost Classifier para analizar en más detalle su capacidad de predicción, ya que fue el algoritmo que entregó mejores resultados de predicción en todas las métricas consideradas.

Lo primero que se estudia es la capacidad de predicción de clase positiva y negativa, diferenciando entre cada una de las plataformas. La siguiente figura muestra esta información:

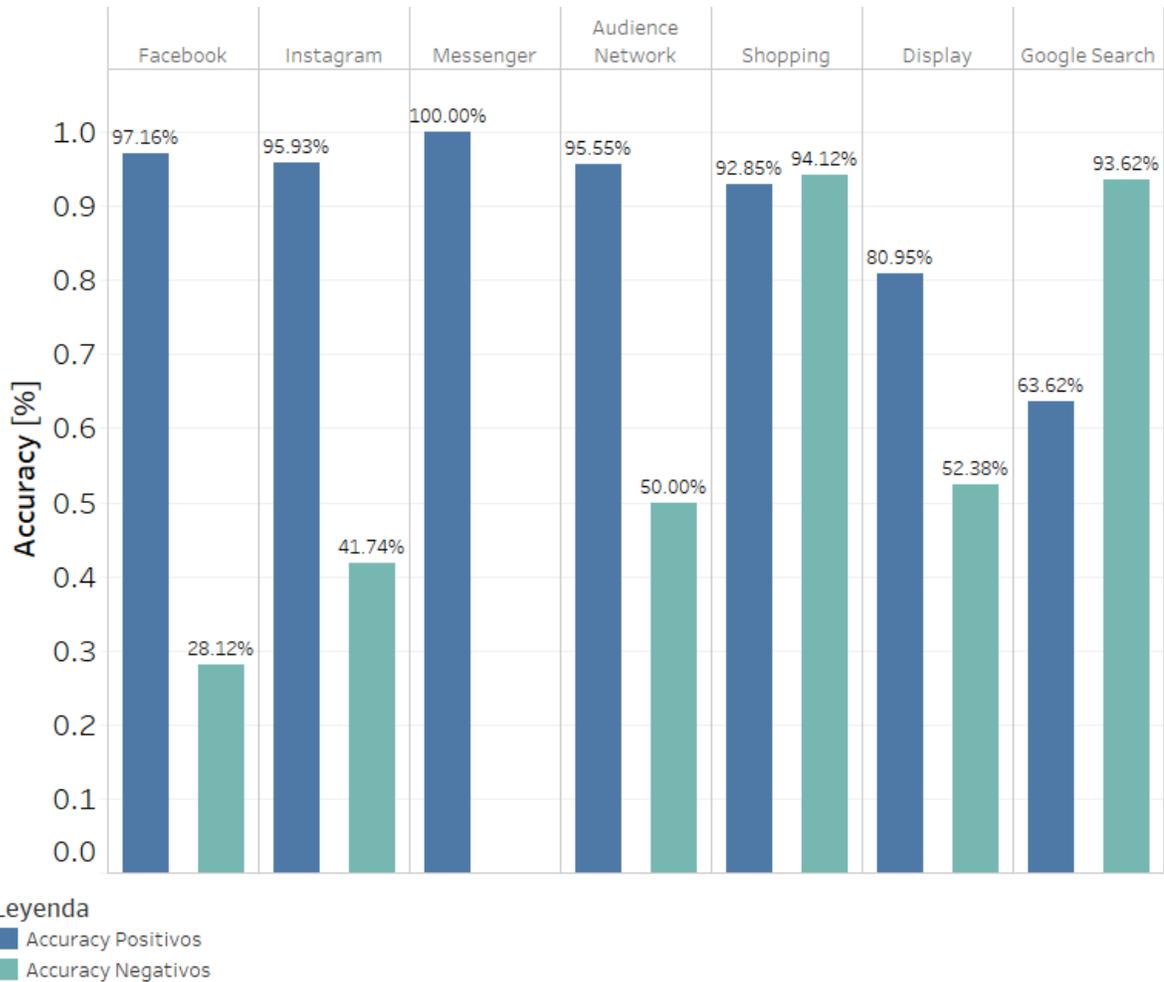


Figura 26: Accuracy detallado para cada una de las clases y cada una de las plataformas

(Fuente: elaboración propia)

Se puede apreciar que esta forma de modelar varía enormemente su desempeño dependiendo de la plataforma y clase de la observación. Lo más importante que se puede notar a partir de estos resultados es que:

1. Existe una diferencia de comportamiento importante entre las plataformas de Facebook ADS (Facebook, Instagram, Messenger y Audience Network) y Google ADS (Search, Shopping y Display)
2. Para las plataformas de Facebook ADS pareciera que el modelo es muy bueno para predecir registros de clase positiva y muy malo para predecir clase negativa, llegando incluso al caso extremo de tener un

100% de accuracy para predecir clase positiva y 0% para clase negativa en la plataforma Messenger

De esta manera, con la finalidad de indagar más en profundidad sobre estos últimos puntos, es que a continuación se muestra dos gráficas que describirán las características de la base de testeo sobre la cual se realizó esta predicción:

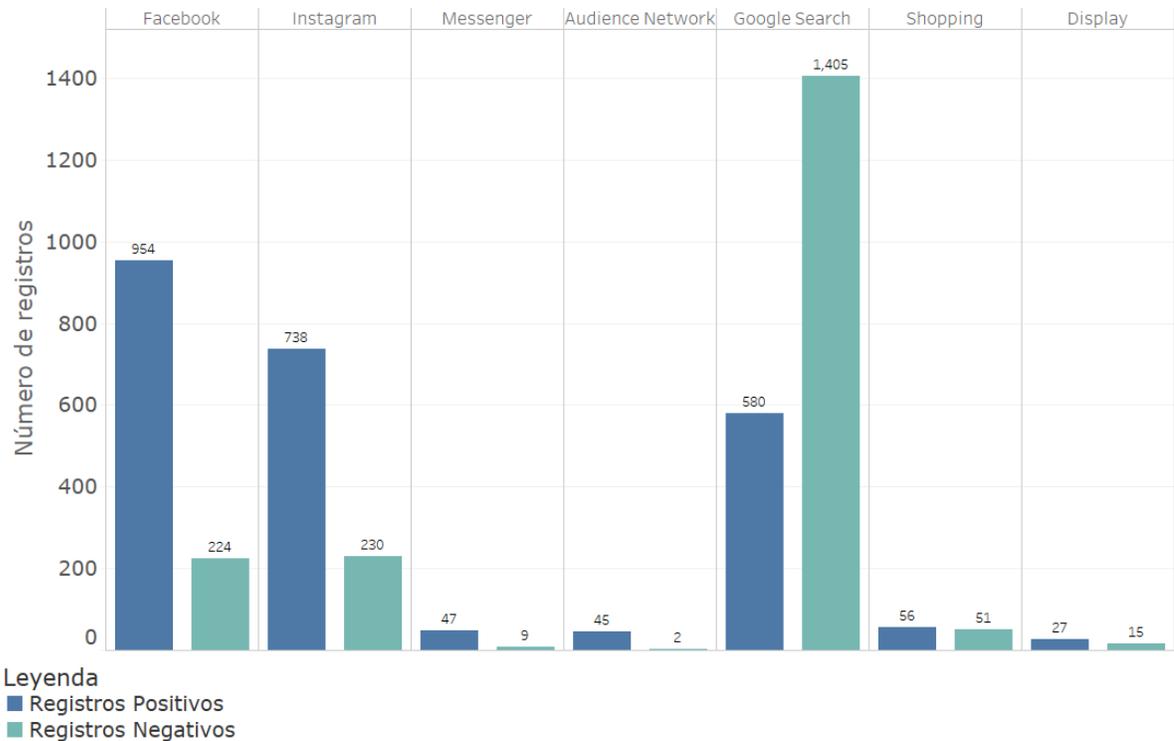


Figura 27: Cantidad de registros con clase positiva y negativa de cada plataforma en la base de testeo

(Fuente: elaboración propia)

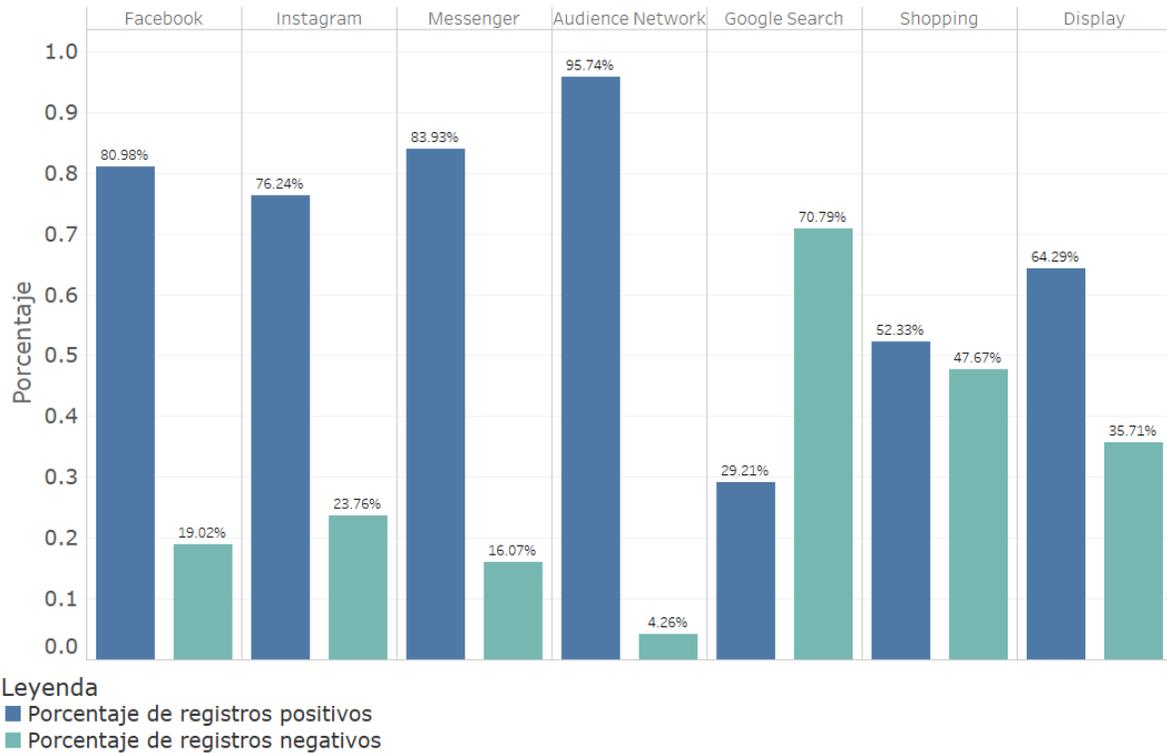


Figura 28: Porcentaje de registros con clase positiva y negativa de cada plataforma en la base de testeo

(Fuente: elaboración propia)

De los gráficos recién mostrados se pueden concluir dos informaciones relevantes:

1. Existe un desbalance de clases no menor para las plataformas del grupo de Facebook ADS
2. En cuanto a cantidad de registros, existen plataformas con un muy pocas observaciones en la base de testeo, llegando incluso al extremo donde la plataforma Audience Network llega a tener solo 2 registros de clase negativa

El primer punto entrega una de las razones del porqué el modelo presenta diferencias en la capacidad de predicción cuando se estudian las clases positiva y negativa por separado. Se puede ver que en general la proporción que existe de clase positiva en la base de testeo es muy superior para las plataformas de Facebook ADS, lo que termina generando que el modelo prediga mejor esta

clase y que también por otro lado, el modelo tenga cierta tendencia a dar como respuesta de predicción una clase positiva, lo que termina perjudicando la capacidad de predicción del modelo para clase negativa.

Este desbalance de clases puede explicarse debido al umbral de predicción elegido para este modelo, el cual fue fijado en 0.14 [USD], esto se puede entender mucho mejor a través de la siguiente gráfica:

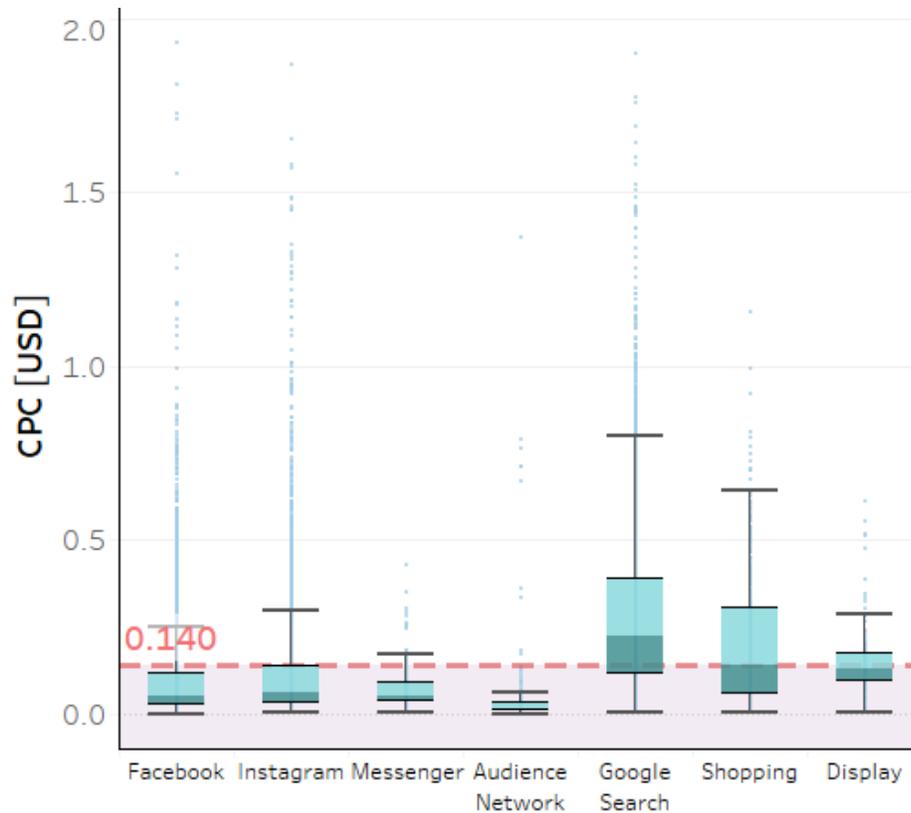


Figura 29: Distribución del CPC para cada una de las plataformas de publicación en contraste con el umbral de predicción de 0.14 [USD]

(Fuente: elaboración propia)

Se puede ver que el umbral de predicción elegido tiene relación con el desbalance de clases producido, ya que este provoca que la mayoría de las observaciones para plataformas de Facebook ADS se encuentren bajo el umbral de predicción, lo que no sucede para las plataformas de Google ADS, lo que explica porqué este problema no se presenta para este último grupo de plataformas.

Es importante de todas formas hacer el análisis de cuál es el resultado predictivo del modelo de umbral único de predicción cuando se deja fuera del análisis las plataformas correspondientes al grupo de Facebook ADS, es decir, se dejan afuera aquellas plataformas que terminaron teniendo un alto desbalance de clases debido a la elección de 0.14 [USD] como umbral de predicción. Al hacer esto, quedan solo las plataformas de Google ADS correspondientes a Google Search, Shopping y Display, en este contexto, el modelo logra un accuracy promedio de 84.53%, es decir, el modelo es capaz de predecir con un 84.53% de exactitud si es que un anuncio tendrá un CPC mayor o menor a 0.14 [USD] para este grupo de plataformas. Por otro lado, también es importante mencionar que el modelo logra predecir con un 94.17% de exactitud aquellos registros de clase negativa para este grupo de plataformas, lo cual es relevante, ya que esto significa que el modelo es capaz de predecir correctamente en un 94.17% de los casos si es que un anuncio publicado tendrá un resultado deficiente en cuando a CPC se trata.

7.1.2. Segundo modelo: clasificación del CPC por medio de un umbral diferenciado para cada una de las plataformas

Esta forma de modelar tiene como finalidad diferenciar el umbral de predicción para cada una de las plataformas, de esta manera, la definición de clase positiva y negativa se ajusta a los valores históricos de cada plataforma en particular, evitando así los posibles desbalances de clases que pueden terminar alterando negativamente el resultado de predicción. Este umbral establecido corresponderá a la mediana histórica de CPC para cada una de las plataformas, los valores de estos umbrales se detallan en la siguiente tabla:

Tabla 7: Umbrales de predicción establecidos tomando como referencia la mediana histórica de CPC para cada plataforma

Plataforma	Umbral de predicción (Mediana histórica de CPC) [USD]
Facebook	0.06
Instagram	0.087
Messenger	0.051
Audience Network	0.011
Google Search	0.237
Shopping	0.153
Display	0.132

(Fuente: elaboración propia)

Esto quiere decir que ahora por cada una de las plataformas se desarrolla un modelo diferenciado, donde cada uno de estos definirá de diferente manera lo que corresponde a una clase positiva y negativa por medio de su propio umbral de predicción. A continuación, se entregan los resultados utilizando esta metodología a través de una tabla que muestra el mejor algoritmo de predicción para cada plataforma, en conjunto con sus métricas de desempeño (los resultados detallados de cada uno de los algoritmos se detallan en el anexo 11.1 de este informe):

Tabla 8: Resultados generales de predicción para cada una de las plataformas

Plataforma	Algoritmo de predicción más eficiente	Accuracy	AUC
Facebook	XGBoost Classifier	78.96%	0.88
Instagram	XGBoost Classifier	79.28%	0.88
Audience Network	XGBoost Classifier	75.00%	0.86
Messenger	XGBoost Classifier	74.51%	0.81
Google Search	XGBoost Classifier	82.85%	0.91
Display	Random Forest Classifier	63.64%	0.70
Shopping	XGBoost Classifier	93.40%	0.97

(Fuente: elaboración propia)

Esta tabla se construyó considerando como algoritmo más eficiente a aquel algoritmo que tiene un mayor AUC para predecir los datos de esa plataforma en particular. En los casos en los que dos o más algoritmos tengan un valor de AUC muy similar, se elegirá aquel que tenga un mayor accuracy.

A continuación, se muestra también el balance de clases que existió en las bases de testeo sobre las cuales se realizaron las predicciones:

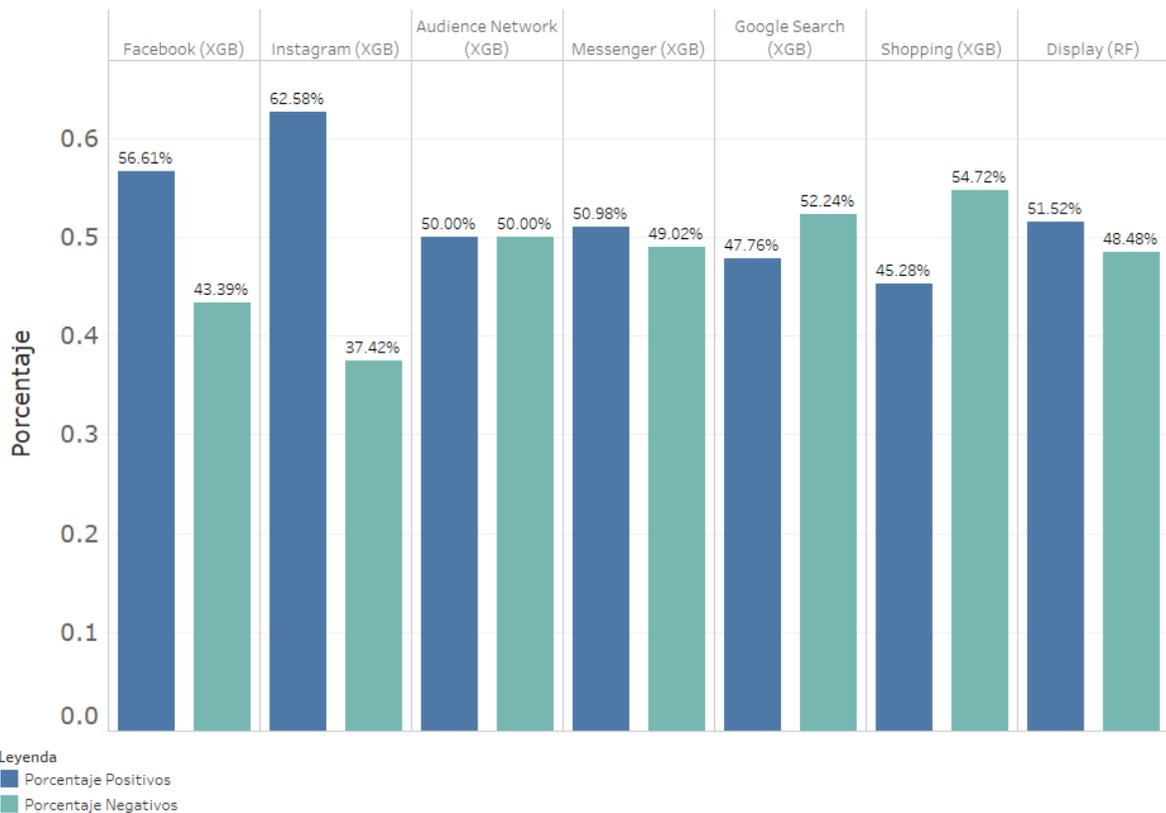


Figura 30: Porcentaje de registros con clase positiva y negativa de cada plataforma en la base de testeo

(Fuente: elaboración propia)

Es importante también tener en cuenta la cantidad de registros por clase que existió en la base de testeo, esto se muestra en la siguiente figura:

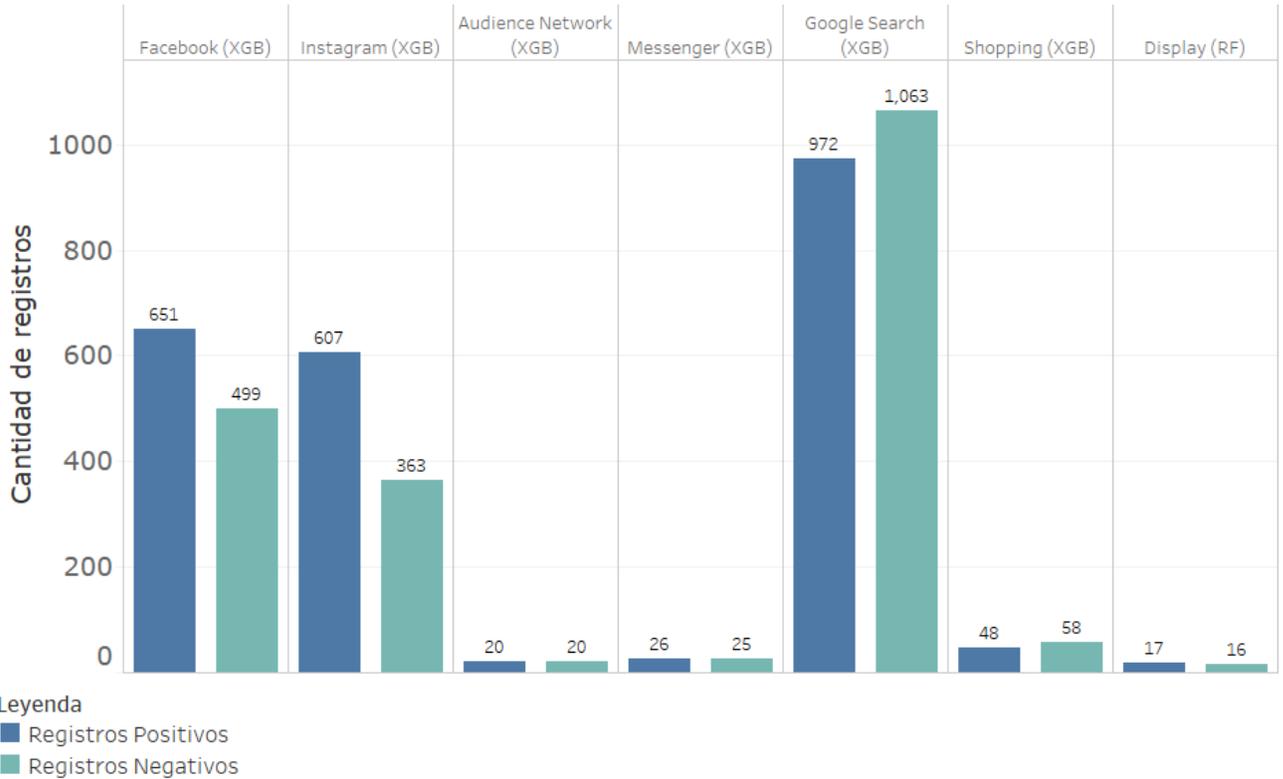


Figura 31: Cantidad de registros por clase en la base de testeo dividiendo por plataforma

(Fuente: elaboración propia)

Y, por último, se muestra a continuación la capacidad de predicción para cada una de las plataformas, pero diferenciando entre capacidad de predicción de clase positiva y negativa:

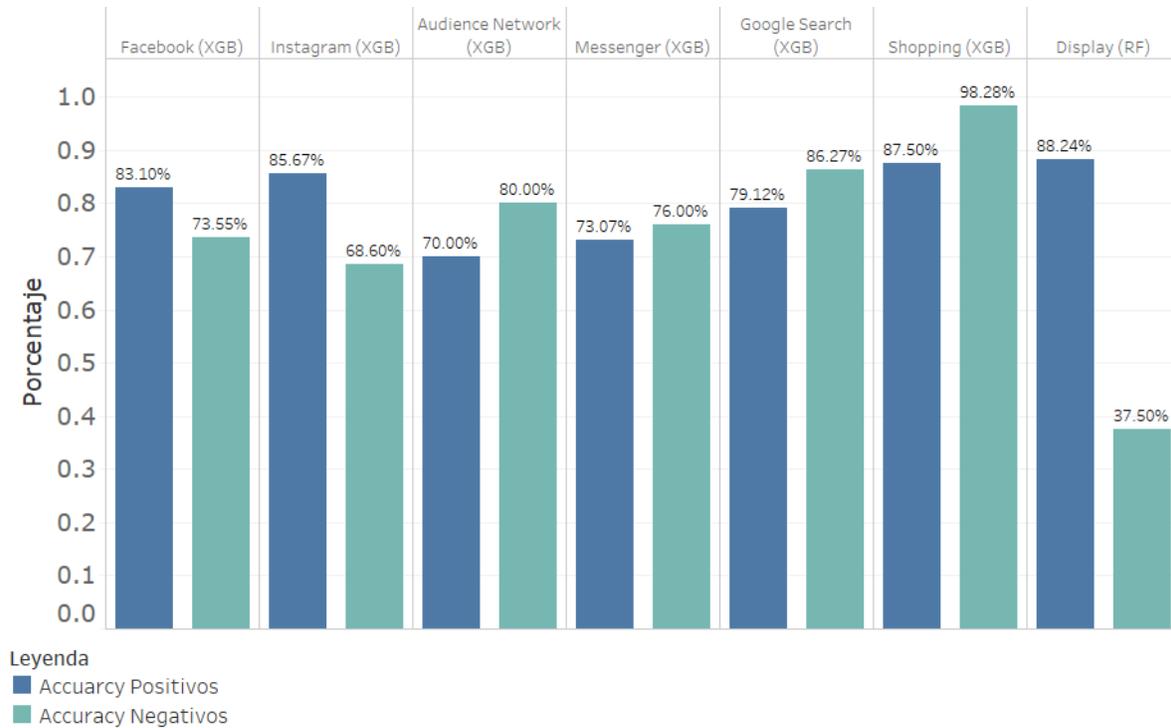


Figura 32: Capacidad de predicción (accuracy) diferenciando por plataforma y por clase

(Fuente: elaboración propia)

Es importante mencionar que la evaluación de la precisión mostrada en estos últimos dos gráficos es realizada por medio de una predicción de la base de testeo utilizando el algoritmo de mejor desempeño para cada una de las plataformas.

De los resultados obtenidos se desprenden las siguientes informaciones relevantes:

1. El mejor algoritmo de predicción no es único este caso, y varía dependiendo de la plataforma, sin embargo, en la mayoría de los casos XGBoost Classifier entrega mejores resultados
2. La magnitud de los resultados varía de manera significativa dependiendo de la plataforma, donde el menor accuracy a nivel general se alcanza en la plataforma Display con un 63.64% y el mayor en la plataforma Shopping con un 93.4%

3. A nivel general, la capacidad de predicción es menor en comparación al modelo que considera un umbral único de predicción
4. La elección de la mediana histórica como umbral de predicción para cada plataforma logra solucionar el problema de desbalance de clases provocado por la elección de un umbral único de predicción
5. Las plataformas de Audience Network, Display, Messenger y Shopping tienen una cantidad de registros muy baja en su base de testeo, lo que termina provocando que la métrica accuracy pueda no estar representando la real capacidad predictiva

En cuanto a la capacidad de predicción a nivel general de este modelo se obtiene un accuracy del 80.93%, considerando todas las plataformas al mismo tiempo. Por otro lado, si se calcula este mismo accuracy, pero solo para las plataformas Facebook, Instagram y Search (aquellas que tienen una alta cantidad de registros), se obtiene un accuracy del 77.05%, lo que indica que, al agrupar los resultados de predicción de todas las plataformas, aquellas que tienen una cantidad de registros baja suelen impactar en cerca de 4 puntos porcentuales a los resultados a nivel agregado.

7.1.3. Análisis de importancia de features

Un análisis interesante corresponde al estudio de la importancia que tienen cada una de las variables independientes utilizadas en los modelos. Esto quiere decir que a continuación se entenderá que variables fueron más relevantes que otras para efectos de alcanzar los resultados predictivos mostrados anteriormente.

Este análisis se basa en el uso de "Shap Values", los cuales corresponden a una metodología derivada de la teoría de juegos, la cual permite determinar que variables tienen mayor incidencia sobre el resultado de clasificación, entregando información de la importancia de cada covariable tanto en magnitud como en la dirección en la cual fue esta importancia, es decir, si es que la covariable aportó en la dirección de clasificar el registro como positivo o como negativo. Se aplica este análisis para los modelos entre a través del algoritmo XGBoost Classifier (umbral diferenciado por plataforma) para las plataformas Facebook, Instagram y Google Search, ya que son estos los que poseen una mayor confianza estadística por la alta cantidad de registros que se tienen de estas plataformas.

En el anexo 11.7, se pueden encontrar los resultados para cada uno de los modelos. Lo que muestran estos gráficos corresponde a la importancia de cada una de las covariables por medio del cálculo de la importancia promedio que tuvo esa covariable para cada uno de los registros predichos. Es importante mencionar que los gráficos muestran el valor absoluto promedio de la importancia, es decir, que entrega información acerca de la importancia en general de la variable independiente sin considerar si aportó en la clasificación positiva o negativa.

El principal resultado relevante corresponde a que por lejos el costo de publicación semanal del anuncio es la variable más importante en el proceso de clasificación, tanto para la plataforma de Facebook, Instagram y Google Search. Por otro lado, llama la atención el hecho de que el cliente asociado la publicación también tiene una alta importancia, en este sentido los diversos clientes importan de manera diferente en el proceso de clasificación, destacando el cliente categorizado como 1 por ser el cliente que mayor importancia tiene en el proceso de clasificación. Por último, de este análisis se puede ver que la cantidad de semanas que lleva publicado un anuncio es importante respecto al CPC alcanzado y también, de manera sorpresiva, que el día de la semana es también relevante, en particular, el hecho de que una publicación esté en línea un día domingo tiene importancia respecto al CPC alcanzado por la publicación.

7.1.4. Conclusiones generales de los modelos de clasificación desarrollados

Una de las conclusiones más importantes que pueden obtenerse a partir de los resultados recién descritos, es el hecho de que el umbral de predicción elegido es muy relevante respecto al desempeño predictivo que logran los modelos de clasificación. Se logra apreciar que la elección de un umbral único para todas las plataformas no logra entregar buenos resultados de predicción a nivel general, esto debido a que, dada la heterogeneidad de comportamiento en cuanto a CPC de cada una de las plataformas, hace que este umbral sea adecuado solo para algunas de estas en particular, dejando afuera a aquellas plataformas que en general tienen niveles de CPC lejanos al umbral elegido. Por lo tanto, si lo que se busca es precisión a nivel general considerando todas las plataformas al mismo tiempo, la versión del modelo que diferencia el umbral de predicción para cada una de las plataformas sería por lo tanto la opción más adecuada.

A pesar de esto, es importante mencionar que el modelo que considera un umbral único de predicción igual a 0.14 [USD] si bien no entrega buenos resultados a nivel general, si logra entregar una buena precisión para las plataformas pertenecientes al conjunto de Google ADS, es decir, Google Search, Shopping y Google Display. Esto último indica que los modelos de umbral único por plataforma podrían ser útiles para contextos particulares de predicción.

Por otro lado, se puede concluir que, al diferenciar el umbral de predicción para cada una de las plataformas por medio de su mediana histórica de CPC, se suele beneficiar a los modelos en cuanto a un mejor balance de clases, sin embargo, los resultados de predicción suelen ser menos precisos. Esto último puede explicarse teniendo en cuenta de que cuando se establece un umbral único de predicción, existen plataformas que de manera histórica suelen alejarse de ese umbral, lo que provoca que el modelo pueda de manera más o menos fácil predecir la clase de las observaciones de esas plataformas. Si bien esto termina provocando una mayor precisión predictiva, por otro lado, disminuye el aporte en la reducción de la incertidumbre para la empresa, ya que, de establecerse un umbral de predicción demasiado alejado del desempeño histórico de alguna plataforma, también es más sencillo para el tomador de decisión poder anticiparse a esos resultados con una simple observación a los datos históricos.

En cuanto a las magnitudes de precisión que son logradas por estos modelos, se puede ver que dependen del tipo de modelo a utilizar (umbral único o diferenciado) y también de la plataforma a la cual pertenezcan los registros, a pesar de esto, en general se obtienen precisiones que van desde el 63.64% de accuracy hasta un 93.4%, donde en su mayoría las precisiones se encuentran en un intervalo entre el 75% y 85% dependiendo de la plataforma. Un insight relevante encontrado en esta sección corresponde a que las plataformas que en general lograron mayor desempeño en cuanto a precisión fueron las plataformas que funcionan asociadas a un motor de búsqueda, las cuales son Google Search y Shopping.

Es importante además tener en cuenta el hecho de que en este análisis se contó con muy pocos datos de las plataformas Audience Network, Messenger, Shopping y Display en contraste con el resto de las plataformas, por lo que es posible que no se esté logrando representar fielmente el comportamiento de estas plataformas por medio de los datos disponibles.

Por último, un aprendizaje importante obtenido es que existen variables claves que determinan el resultado de CPC de un anuncio publicado a nivel semanal, entre ellas están el más importante que es el costo del anuncio y menor medida

está el cliente asociado a la publicación, la cantidad de semanas consecutivas de la publicación y por último el día de la semana publicado, donde el día domingo tiene una especial relevancia en contraste al resto de los días de la semana.

7.2. Modelos de predicción del CPC como variable continua

Como se explicaba ya previamente, estos modelos tienen por objetivo predecir el valor que tendrá el CPC para la publicación de un anuncio a nivel semanal. La configuración de los modelos evaluados en esta sección son los siguientes:

1. Predicción continua del CPC considerando información de todas las plataformas al mismo tiempo
2. Predicción continua del CPC generando un proceso de entrenamiento distinto para cada una de las plataformas

7.2.1. Predicción continua del CPC considerando todas las plataformas al mismo tiempo

Los resultados de predicción obtenidos por cada uno de los algoritmos se muestran en la siguiente tabla:

Tabla 9: Resultados de predicción de CPC continuo considerando todas las plataformas

Algoritmo de predicción	MAE [USD]	RMSE [USD]
Decision Tree Regressor	0.115	0.171
XGBoost Regressor	0.077	0.128
Random Forest Regressor	0.084	0.136

(Fuente: elaboración propia)

De estos algoritmos el que mejor precisión alcanza en cuanto a MAE y RMSE corresponde al XGBoost Regressor. Para poder saber si este error es grande o pequeño, es necesario tener en cuenta los resultados históricos de la agencia. De esta forma, un MAE igual a 0.077 [USD] representa un 41.82% del promedio histórico de CPC en la base de testeo.

Una posible explicación del porqué no se logra una mejor precisión en la predicción, tiene que ver con la alta variabilidad que existe en la base de datos, donde la heterogeneidad existente en los niveles de CPC para las distintas plataformas puede estar alterando negativamente la precisión del modelo.

Es por esto que el siguiente modelo separa el entrenamiento y la evaluación por cada una de las distintas plataformas, esto con la finalidad de disminuir el ruido se pueda estar generando por la inclusión de diferentes plataformas a un mismo modelo predictivo.

7.2.2. Predicción continua del CPC diferenciando por plataforma

Al igual que en los modelos de clasificación, como se entrenan distintos algoritmos por cada una de las plataformas, se obtiene un mejor algoritmo predictivo para cada una de estas, de esta forma, los resultados de este proceso de entrenamiento y evaluación:

Tabla 10: Resultados predictivos de modelo de predicción continua que diferencia por cada una de las plataformas

Plataforma	Algoritmo de predicción más eficiente	MAE [USD]	Porcentaje que representa el MAE del promedio histórico de la plataforma
Facebook	XGBoost Regressor	0.058	62.25%
Instagram	Random Forest Regressor	0.070	62.05%
Audience Network	XGBoost Regressor	0.027	38.87%
Messenger	XGBoost Regressor	0.037	45.57%
Google Search	XGBoost Regressor	0.092	34.36%
Display	Random Forest Regressor	0.067	42.00%
Shopping	XGBoost Regressor	0.068	29.96%

(Fuente: elaboración propia)

De estos resultados se puede obtener las siguientes informaciones relevantes:

1. El algoritmo que mejores resultados de predicción entrega es en general el XGBoost Regressor, a excepción de las plataformas Instagram y Display, donde el algoritmo Random Forest Regressor fue el que mejor resultados de predicción entregó. Es importante también mencionar que en ninguna plataforma el algoritmo de Decision Tree Regressor entregó los mejores resultados
2. La precisión alcanzada para cada una de las plataformas varía mucho entre las diferentes plataformas, ya que los errores pueden ir desde el 29,96% hasta el 62.25% dependiendo de estas
3. En general los resultados de predicción son poco precisos, ya que 6 de estas 7 plataformas llegan a tener un error promedio absoluto que representa más del 30% del promedio histórico de CPC de cada una de estas plataformas

7.2.3. Conclusiones de los modelos predictivos del CPC como variable continua

Una de las primeras conclusiones importantes que se pueden obtener a partir del análisis de estos modelos de predicción de variable continua, es que no es adecuado entrenar un modelo que considerando todas las plataformas al mismo tiempo. Esto último se explica por el hecho de que al entrenar modelos que diferencian entre cada una de las plataformas, los resultados muestran que la precisión que puede ser alcanzada difiere mucho de plataforma en plataforma, por lo que sería un error agrupar los resultados y decir que existe una métrica de desempeño predictivo que representa a todas las plataformas al mismo tiempo.

Por otro lado, es importante mencionar que, si bien el error predictivo es alto a nivel general, existen plataformas que tienen un nivel de error que es considerado bajo en contraste a las demás plataformas, como por ejemplo para la plataforma Shopping, la cual tiene un MAE de 0.068 [USD], lo que representa un 29.96% del promedio histórico obtenido por esa plataforma.

Al igual que en los modelos de clasificación, algunas de estas plataformas tienen una cantidad de registros muy baja (Audience Network, Messenger, Display y Shopping), por lo que, si se omiten estas plataformas del análisis, la predicción con mejor desempeño corresponde a la plataforma Google Search, la cual es aproximadamente el doble más precisa que las plataformas de Facebook e Instagram cuando se miden a partir del porcentaje que el MAE representa del promedio histórico de la plataforma, de esta forma, una conclusión relevante corresponde a que Google Search logra un muy buen desempeño bajo este contexto.

8. Contraste de efectividad de modelos contra decisiones tomadas a criterio experto

Hasta el momento se han desarrollado y evaluado modelos que con cierta precisión predicen el resultado de un anuncio digital publicado a nivel semanal, lo cual permite a la agencia contar con herramientas estimativas adicionales a la de criterio experto. De esta forma, una pregunta relevante tiene que ver con el hecho de si es que los modelos desarrollados logran ser más precisos que las posibles estimaciones realizadas a criterio experto. Para responder a esto es que se realiza un diseño experimental que tiene por finalidad comparar la precisión de los modelos de clasificación con la precisión predictiva que puede llegar a tener un(a) experto(a).

El experimento consiste en predecir un conjunto de registros representativos de la base de datos por medio de los modelos de clasificación, y, por otro lado, predecir el mismo conjunto de registros, pero utilizando la metodología de criterio experto. De esta forma, las características que definen el experimento se describen a continuación:

- De manera aleatoria, se extraen de la base de datos 92 registros. Donde 10 registros corresponden a la plataforma Facebook, 17 a Instagram, 16 a Audience Network, 9 a Messenger, 8 a Google Search, 16 a Shopping y 16 a Display. Estos registros son elegidos aleatoriamente desde la base de datos de testeo, es decir, no formaron parte de los registros que entrenaron a los modelos
- La forma de predecir los registros por parte de los expertos es por medio de un cuestionario que simula diversos escenarios de inversión, donde cada una de estas situaciones corresponde a un registro de la base de datos. Una situación ejemplo de este cuestionario puede encontrarse en el anexo 11.5
- Los modelos de clasificación utilizados para predecir corresponden a los desarrollados en la sección 7.1.2 de este informe, es decir, los modelos de clasificación que utilizan la mediana de cada plataforma como umbral que define clase positiva y negativa
- El cuestionario fue resuelto por dos expertos pertenecientes al equipo de medios, donde las situaciones se repartieron de manera equitativa y de forma aleatoria
- La manera de concluir este experimento es a partir de un test de diferencia de proporciones de dos muestras, donde la proporción considerada corresponde al accuracy que alcanza una metodología en particular luego de la predicción
- Las hipótesis alternativa y nula establecidas para este test son las siguientes:

$$H_0: Accuracy_{modelos} = Accuracy_{expertos} \quad (11)$$

$$H_1: Accuracy_{modelos} > Accuracy_{expertos} \quad (12)$$

Los resultados de predicción muestran que el accuracy alcanzado es de un 77.18% para los modelos de aprendizaje automático, considerando de manera agregada todas las plataformas. Por otro lado, el accuracy alcanzado por los expertos al predecir la misma muestra es de un 47.83%.

De esta manera, dados los resultados obtenidos, a un nivel de significancia menor al 0.1% ($\alpha \approx 0.001$), se puede rechazar la hipótesis nula propuesta en este test de diferencia de proporciones, concluyendo el que el accuracy alcanzado por los modelos es mayor respecto al accuracy alcanzado por los expertos.

Es importante tener en consideración que lo que se prueba en este test es la diferencia en la capacidad predictiva de los modelos en comparación a los expertos, **pero bajo las mismas condiciones de disponibilidad de datos**. Es posible que los expertos consideren variables exógenas a la base de datos disponible para estimar resultados de inversión. Esto es importante, ya que el accuracy de los(las) expertos(as) podría variar dependiendo de la disponibilidad de datos que le son entregados. Esto último podría explicar en parte el bajo desempeño predictivo por parte de los expertos en este experimento. Teniendo en cuenta esto último, la conclusión de experimento es acotada y debe entenderse como que **ante restricciones de variables de control** los modelos desarrollados son más eficaces para predecir el desempeño de un anuncio en cuanto a su CPC alcanzado. En la siguiente sección se proponen en detalle aspectos de mejora de este test para trabajos futuros.

9. Recomendaciones de implementación y propuestas de mejora para trabajos futuros

Esta sección tiene por finalidad mostrar aquellas informaciones y conclusiones relevantes que se obtienen del trabajo realizado de cara a entregar valor a la empresa. De esta manera, se desprenden diferentes subsecciones de relevancia, las cuales serán expuestas a continuación.

9.1. Recomendaciones de uso de modelos para realizar predicciones

En cuanto a los modelos de predictivos de clasificación, una de las primeras indicaciones tiene que ver con el hecho de que no es recomendable predecir todas las plataformas a través del entrenamiento de un modelo único. Como se vio en la evaluación de los resultados predictivos, el agregar todas las plataformas se provoca que el modelo tenga buenos resultados solo para las plataformas que se ajustan bien al umbral de predicción elegido. De esta manera se recomienda entrenar un modelo diferente con los datos de cada plataforma en particular, y elegir el umbral de predicción de cada plataforma según la utilidad que le otorgue a la agencia para la reducción de la incertidumbre. En los siguientes párrafos se hablará más en detalle acerca de la elección de estos umbrales.

Para la elección de estos umbrales de predicción, hay que tener en cuenta de que esta decisión se va a traducir en un trade-off entre precisión y ganancia de información para la agencia. Esto último debido a que por ejemplo la predicción por medio de umbrales como la mediana histórica, le otorga una alta exigencia al modelo, lo que implica por lo tanto una baja en la precisión alcanzada. Sin embargo, estas predicciones pueden ser muy útiles para efectos del negocio, ya que los resultados obtenidos corresponden a predicciones no triviales de realizar sin el uso de modelos predictivos que consideran los datos a nivel histórico de la agencia. De esta manera, si el usuario de estos modelos es capaz de tolerar un mayor riesgo de error asociado, es recomendable usar umbrales de predicción que se basen en los resultados históricos tales como la mediana o la media histórica.

Relacionado también con el párrafo anterior se desprende otra recomendación relevante para la agencia, y es el hecho de que el establecimiento de umbrales de predicción alejados del comportamiento histórico, si bien terminan generando predicciones que en la mayoría de los casos no aportan gran información a la agencia, si se usan correctamente, podrían llegar entregar valor a la compañía. Un ejemplo claro de esto podría ser la utilización de modelos con umbrales de predicción muy altos o muy bajos para detectar con una alta precisión publicaciones con resultados extremos, es decir, o muy malos o muy buenos. De esta manera, se recomienda el establecimiento de umbrales extremos para evitar muy malos resultados (CPC muy alto) o detectar posibles resultados considerados sobresalientemente buenos (CPC muy bajo).

Por último, un punto importante también a mencionar es que es recomendable utilizar en general el algoritmo XGBoost Classifier para realizar predicciones de clase por sobre otros algoritmos, ya que en la mayoría de las situaciones evaluadas fue este el algoritmo el que entregó mejores resultados de ajuste.

En cuanto a los modelos de predicción continua, la primera recomendación se relaciona con que en general estos modelos entregan resultados muy poco precisos, de esta manera, se recomienda primero que nada el uso de modelos predictivos de clasificación por sobre el intentar predecir el CPC como variable continua. Sin embargo, si por algún motivo el usuario de estos modelos prefiere predecir de manera continua a pesar de los errores encontrados, la recomendación a nivel general es parecida a la de modelos de clasificación, entrenar separadamente modelos por cada una de las plataformas, ya que esta forma se obtendrán mejores resultados predictivos en cuanto a precisión.

Por último, cabe destacar el buen desempeño en cuanto a precisión que tiene la plataforma Shopping por sobre las demás, tanto en los modelos predictivos de clasificación como en los de predicción de variable continua. Estos resultados indican que los datos de esta plataforma responden bien a los modelos desarrollados, por lo que se recomienda en particular el uso de estos modelos para predecir el desempeño de los anuncios pertenecientes a esta plataforma.

9.2. Recomendaciones relacionadas a la disponibilidad de datos

9.2.1. Implementación de proceso de registro de datos y consideración de nuevas variables de control

En esta sección las recomendaciones se basan en dos ideas principalmente, la primera corresponde al establecer procesos de almacenamiento de datos de desempeño que faciliten la implementación y mejora continua de los modelos predictivos, y, en segundo lugar, el comenzar también a registrar nuevas variables que no estuvieron disponibles para este trabajo de título y que en teoría son útiles para determinar el desempeño de un anuncio publicitario.

Respecto al primer punto, muchas de las variables utilizadas para el desarrollo y evaluación de los modelos predictivos debieron ser trabajadas y obtenidas a través de complejos procesos de limpieza y transformación de variables.

Muchas de estas variables pueden ser registradas por la empresa de manera directa para su almacenamiento, a continuación, se describirán las oportunidades de mejora para cada una de estas variables:

- **Días de la semana en los cuales la publicación estuvo en línea:** esta variable está actualmente de manera implícita dentro de los datos de la empresa, y para obtenerla es necesario extraer los datos desde un software ETL considerando dos niveles de agregación, el nivel diario y el nivel semanal. Solo a través de la base de datos a nivel diaria es posible obtener esta variable, por lo que es necesario crear la variable en este nivel de agregación y luego cruzarla con la base de datos a nivel semanal. Esto será necesario realizarlo cada vez que quiera entrenarse nuevamente los modelos con nuevos datos, por lo que establecer algún proceso de registro de datos que facilite esta tarea podría ser fundamental para el proceso de mejora continua de los modelos
- **Contador de número de semanas que el anuncio lleva publicado:** esta variable también está de manera implícita en los datos, y es necesaria obtenerla por medio de la aplicación de lógicas de manejo de datos que toman un tiempo en ser desarrolladas, lo cual podría ser un impedimento para establecer un proceso de mejora continua de los modelos
- **Indicador que muestra la cantidad adicional de plataformas que son utilizadas por anuncios de un mismo cliente en la semana correspondiente del registro:** este indicador nace con la finalidad de modelar de manera aproximada el efecto que puede tener un anuncio publicado en una plataforma sobre el mismo anuncio, pero publicado en otra plataforma en el mismo periodo de tiempo. Este indicador corresponde solo a una aproximación y fue necesario utilizarlo por la imposibilidad de poder saber cuándo un anuncio es publicado en más de una plataforma a la vez, esto ya que los identificadores de los anuncios no son iguales cuando son publicados en distintas plataformas. De esta manera, una recomendación muy importante es la de asignar identificadores únicos a nivel de anuncio para poder obtener información respecto a un mismo anuncio que es publicado en diferentes plataformas

En general, las recomendaciones en este punto se pueden resumir en el hecho de que la empresa no cuenta actualmente con un procedimiento establecido de almacenamiento de datos, si no que los datos con los que se cuenta son solo aquellos que quedan registrados dentro de las plataformas en las cuales se hacen las publicaciones, por lo que el formato y el acceso a ellos no es

amigable para la implementación de modelos de aprendizaje automático. Debido a esto último, es recomendable establecer un proceso de registro de datos que sea independiente de las plataformas de publicación, y que le permita a la empresa llevar un registro histórico con las variables de interés que permita una fácil implementación y mejora continua de los modelos predictivos.

Respecto al segundo punto, los modelos que se evaluaron en este trabajo de título fueron desarrollados bajo un contexto de disponibilidad limitada de variables, de hecho, como ya se mencionaba en la sección 6.1 de entendimiento de datos, al menos dos variables que en teoría se sabe que determinan el desempeño de un anuncio no fueron incluidas debido a que no existía registro de ellas. De esta forma se recomienda el registro histórico de variables que modelen los siguientes efectos:

- **Efectos de contenido:** este efecto se puede modelar por medio de variables estandarizadas que determinen y caractericen el contenido de un anuncio. Por ejemplo, variables que determinen cual es el segmento objetivo al cual va dirigido el anuncio, si es que tiene presencia de la imagen de personajes conocidos (influencers), si es que corresponde a un descuento de tipo monetario, si corresponde a una oferta del tipo 2x1 o variaciones, entre diversas otras variables que caractericen el contenido de una publicación digital
- **Efectos de competencia:** este efecto no es fácil de modelar debido a que requiere la información de las acciones de la competencia, información a la cual en la mayoría de los casos no es posible acceder. Sin embargo, es posible modelar este efecto por medio de aproximaciones. Una recomendación para poder modelar este efecto es tener en cuenta la participación de mercado de la empresa o de la marca del producto o servicio que se está publicitando, de esta manera, se puede integrar a los modelos el hecho de que las acciones de la competencia tienen cierta influencia sobre el desempeño de un anuncio digital [11]

Por último, una última recomendación relacionada al registro de datos tiene que ver con el que la empresa no cuenta con un registro formal y accesible de los pronósticos de desempeño que determinan los expertos para cada uno de los anuncios publicados. El registro de estos pronósticos le podría permitir a la empresa poder evaluar de mejor manera las decisiones tomadas a criterio experto y también decisiones basadas en la aplicación de los modelos

predictivos, de esta forma, el registro de estos pronósticos puede ser beneficioso para la compañía en el futuro.

9.2.2. Consideración de un mayor universo de datos

El presente trabajo se basa en los datos históricos de una agencia en particular, de un conjunto de plataformas en particular y de un conjunto de clientes en particular, esto hace que la validez externa de las conclusiones obtenidas se limite de cierto modo. En esta sección se analiza que variables son extrapolables a otros contextos y cuales son las recomendaciones para que en un futuro se puedan obtener resultados que sean más generalizables.

En cuanto a los datos históricos considerados estos comprenden desde inicios del 2018 hasta finales del 2020. Por otro lado, es importante mencionar que los clientes a los cuales corresponden los datos históricos son todas tiendas de comercio electrónico. Por último, es necesario mencionar que las publicidades tienen como público objetivo a consumidores que residen en un lugar geográfico en particular, el cual corresponde a todo el territorio nacional de Chile. Estos datos son mencionados debido a que es en este contexto en particular en el cual se desarrolla el presente trabajo, por lo que la validez de las conclusiones solo se puede asegurar dentro de este contexto. A continuación, se detallan que conclusiones pueden generalizarse a otros contextos y cuales no:

1. Resultados podrían ser extrapolables a más clientes manteniendo las demás variables del contexto

La figura 24 de este informe muestra que la distribución histórica de CPC de cada uno de los 4 clientes no tiene variaciones significativas, a pesar de que cada uno de estos clientes posee diferencias relevantes entre sí. Estos clientes se diferencian entre sí respecto a su público objetivo, tipo de productos/servicios ofrecidos, porcentaje del mercado relevante obtenido, entre otras diferenciaciones. De esta manera, se puede pensar que para otros clientes e-commerce que operen dentro del territorio nacional podrían obtener resultados similares si es que mantienen la misma distribución de CPC que los clientes analizados para este trabajo

2. Resultados por plataforma difícilmente generalizables a nuevas plataformas

Del análisis realizado se pudo ver que cada plataforma tenía características particulares que definían su comportamiento, de hecho, una de las conclusiones relevantes de este de este trabajo es que el entrenamiento de los modelos predictivos debe ser diferente por cada plataforma, por la misma razón recién descrita. De esta forma, es que los resultados obtenidos no necesariamente se replicarán en nuevas plataformas diferentes a las consideradas en este trabajo

3. Resultados no necesariamente se replicarán a nuevas regiones

De las investigaciones realizadas para encontrar el valor del CPC estándar para la región, se pudo dar cuenta de que esta métrica varia su nivel estándar de región en región, un ejemplo de esto es que por ejemplo el CPC en Norte América es mucho más alto que para Latinoamérica, esto último da indicios de diferencias en el comportamiento del CPC de región en región, lo que dificulta que los resultados de este trabajo sean homologables a publicidades pertenecientes a otras regiones

Todas estas variables dan cuenta de la validez externa de las conclusiones obtenidas en este trabajo. Se recomienda por lo tanto aumentar la cantidad de plataformas, tipos de clientes y regiones de publicación para en un futuro entrenar modelos que sean generalizables a contextos más diversos.

9.3. Propuestas para trabajos futuros

Esta parte del informe tiene como finalidad entregar diferentes indicaciones y recomendaciones para trabajos futuros que vengán a complementar lo realizado en este trabajo de tesis.

Una de las principales propuestas de mejora tiene que ver con volver a desarrollar y evaluar los modelos predictivos desarrollados, pero considerando una mayor cantidad de variables de control. Las variables más importantes a ser implementadas como complemento a las que ya fueron usadas son las descritas en la sección 9.2 de este informe.

En conjunto con el agregar nuevas variables de control al análisis, también existen propuestas de mejora en cuanto a los algoritmos utilizados. Por un lado, los algoritmos que fueron utilizados en este trabajo pueden mejorarse a través de la prueba de más combinaciones de hiperparámetros, o para el algoritmo de redes neuronales, también por medio de la prueba de diferentes estructuras neuronales. También, sería beneficioso el probar nuevos algoritmos predictivos que no hayan sido considerados en este trabajo. Estas propuestas de mejora tienen como fin el perfeccionar los resultados en cuanto a la precisión que logran las predicciones.

Por otro lado, un trabajo futuro complementario a este trabajo y que puede ser de utilidad para la agencia, corresponde a la evaluación y desarrollo de modelos predictivos, pero ahora considerando más plataformas de las que se consideraron en este trabajo, lo cual le pueda permitir a la empresa ampliar en universo de posibilidades y, por lo tanto, tener más opciones para mejorar sus resultados.

Es importante también mencionar que algunas de las plataformas evaluadas en este trabajo tuvieron al momento del desarrollo de los modelos muy pocas observaciones en comparación a las demás. En concreto, las plataformas de Audience Network, Messenger, Google Display y Shopping tuvieron muy pocos datos en comparación a Facebook, Instagram y Google Search, de esta manera, una recomendación para trabajos futuros tiene que ver con el evaluar modelos predictivos con una mayor cantidad de registros para estas plataformas, de esta manera, se podrá representar de manera más fiel su comportamiento, y, por lo tanto, entregar resultados de predicción más útiles y confiables a la agencia.

Por último, algo que quedó fuera de los alcances de este trabajo tiene que ver con la implementación de estos modelos predictivos en las operaciones de la empresa. Un trabajo futuro que considere esta etapa es valioso ya que entregará información de cómo funcionan estos modelos en la realidad, cuál es su desempeño en tiempo real y cómo se comportan los modelos ante posibles variaciones en el contexto de la publicación de anuncios en línea.

10. Bibliografía

- [1] Chafflet, Dave; Russell Paul. 2002. **E-Marketing Excellence: planning and optimizing your digital Marketing**
- [2] Carlos Andrés Cosming Gonzáles. 2017. **Plan de negocios para la creación de una agencia de marketing digital en Chile**
- [3] Asociación Chilena de publicidad; IPSOS. 2018. **Radiografía de las agencias**
- [4] ¿Qué es Conar? [en línea]. <<https://www.conar.cl/sobre-conar/que-es-conar/>> [consulta: 14 julio 2021]
- [5] What is the average CTR, CPC, CPM of Facebook ads? [en línea]. <<https://e27.co/what-is-the-average-ctr-cpc-cpm-of-facebook-ads-20150121/>> [consulta: 14 julio 2021]
- [6] ¿Cuánto es un Buen CTR en Google AdWords? [en línea]. <<https://yopagoporclick.com/buen-ctr-en-adwords/>> [consulta: 14 julio 2021]
- [7] Google Ads ¿Cuánto cuestan sus anuncios en Chile? [en línea]. <<https://impulsados.cl/presupuesto-anunciarse-google-ads-chile/>> [consulta: 14 julio 2021]
- [8] David Chan; Michael Perry. 2017. **Challenges and Opportunities in Media Mix Modeling**
- [9] Yuxue Jin; Yueqing Wang; Yunting Sun; David Chan; Jim Koehler. 2017. **Bayesian Methods for Media Mix Modeling with Carryover and Shape Effects**
- [10] Yong Liu; Jorge Laguna; Matt Wright; Hua He. 2014. **Media Mix Modeling – A Monte Carlo Simulation Study**
- [11] Gerard J. Tellis. 2006. **Modeling Marketing Mix**
- [12] ¿Qué es el remarketing? Funcionamiento, tipos y ventajas [en línea]. < <https://www.inboundcycle.com/diccionario-marketing-online/remarketing> > [consulta: 14 julio 2021]
- [13] Pros and cons of various Machine Learning algorithms [en línea]. <<https://towardsdatascience.com/pros-and-cons-of-various-classification-ml-algorithms-3b5bfb3c87d6>> [consulta: 14 julio 2021]

- [14] Qué son las redes neuronales y cuál es su aplicación en el marketing [en línea]. < <https://artyco.com/que-son-las-redes-neuronales-y-cual-es-su-aplicacion-en-el-marketing/> > [consulta: 14 julio 2021]
- [15] Random Forest, el poder del Ensamble [en línea]. <<https://www.aprendemachinelearning.com/random-forest-el-poder-del-ensamble/>> [consulta: 14 julio 2021]
- [16] ¿Como funciona XGBoost? [en línea]. <<https://ichi.pro/es/como-funciona-xgboost-128143693994154>> [consulta: 15 julio 2021]
- [17] Árboles de Decisión - Parte I [en línea]. <<https://bookdown.org/content/2031/arboles-de-decision-parte-i.html#arboles-de-regression-vs.arboles-de-clasificacion>> [consulta: 15 julio 2021]
- [18] Canal YouTube: StatQuest with Josh Starmer, Lista de reproduccion: XGboost [en línea]. <https://www.youtube.com/watch?v=OtD8wVaFm6E&list=PLblh5JKOoLULU0irPgs1SnKO6wqVjKUsQ&ab_channel=StatQuestwithJoshStarmer> [consulta: 15 julio 2021]
- [19] Clasificación: ROC y AUC [en línea]. <<https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc?hl=es-419>> [consulta: 15 julio 2021]
- [20] Forecast KPIs: RMSE, MAE, MAPE & Bias [en línea]. <<https://towardsdatascience.com/forecast-kpi-rmse-mae-mape-bias-cdc5703d242d>> [consulta: 15 julio 2021]
- [21] XGBoost Documentation [en línea]. <https://xgboost.readthedocs.io/en/latest/python/python_api.html> [consulta: 15 julio 2021]
- [22] Scikit Learn: User Guide [en línea]. <https://scikit-learn.org/stable/user_guide.html> [consulta: 15 julio 2021]
- [23] Emilio Polit. 2021. Clases Curso Marketing Digital, Capitulo 4. Ingeniería Civil Industrial, Universidad de Chile

11. Anexos

11.1. Resultados predictivos de modelos de clasificación que consideran un umbral de predicción diferente por plataforma

Tabla 11: Resultados generales de predicción para la plataforma Facebook

Algoritmo de predicción	Accuracy	AUC
Decision Tree	73.91%	0.82
Neural Network	76.61%	0.77
Random Forest	71.65%	0.80
XGBoost Classifier	78.96%	0.88

(Fuente: elaboración Propia)

Tabla 12: Resultados generales de predicción para la plataforma Instagram

Algoritmo de predicción	Accuracy	AUC
Decision Tree	77.11%	0.86
Neural Network	78.14%	0.76
Random Forest	69.18%	0.79
XGBoost Classifier	79.28%	0.88

(Fuente: elaboración Propia)

Tabla 13: Resultados generales de predicción para la plataforma Audience Network

Algoritmo de predicción	Accuracy	AUC
Decision Tree	72.50%	0.79
Neural Network	77.50%	0.77
Random Forest	75.00%	0.86
XGBoost Classifier	75.00%	0.86

(Fuente: elaboración propia)

Tabla 14: Resultados generales de predicción para la plataforma Messenger

Algoritmo de predicción	Accuracy	AUC
Decision Tree	64.70%	0.76
Neural Network	58.80%	0.58
Random Forest	66.67%	0.82
XGBoost Classifier	74.51%	0.81

(Fuente: elaboración propia)

Tabla 15: Resultados generales de predicción para la plataforma Google Search

Algoritmo de predicción	Accuracy	AUC
Decision Tree	79.85%	0.88
Neural Network	77.05%	0.77
Random Forest	74.50%	0.83
XGBoost Classifier	82.85%	0.91

(Fuente: elaboración propia)

Tabla 16: Resultados generales de predicción para la plataforma Shopping

Algoritmo de predicción	Accuracy	AUC
Decision Tree	92.45%	0.93
Neural Network	91.51%	0.92
Random Forest	88.68%	0.88
XGBoost Classifier	93.40%	0.97

(Fuente: elaboración propia)

Tabla 17: Resultados generales de predicción para la plataforma Display

Algoritmo de predicción	Accuracy	AUC
Decision Tree	66.67%	0.65
Neural Network	63.64%	0.63
Random Forest	63.64%	0.70
XGBoost Classifier	57.58%	0.67

(Fuente: elaboración propia)

11.2. Resultados en detalle de modelos de predicción del CPC como variable continua

Tabla 18: Resultados de predicción de CPC continuo considerando solo Facebook

Algoritmo de predicción	MAE [USD]	RMSE [USD]
Decision Tree Regressor	0.070	0.119
XGBoost Regressor	0.058	0.107
Random Forest Regressor	0.059	0.106

(Fuente: elaboración propia)

Tabla 19: Resultados de predicción de CPC continuo considerando solo Instagram

Algoritmo de predicción	MAE [USD]	RMSE [USD]
Decision Tree Regressor	0.084	0.140
XGBoost Regressor	0.070	0.123
Random Forest Regressor	0.070	0.121

(Fuente: elaboración propia)

Tabla 20: Resultados de predicción de CPC continuo considerando solo Audience Network

Algoritmo de predicción	MAE [USD]	RMSE [USD]
Decision Tree Regressor	0.082	0.189
XGBoost Regressor	0.027	0.087
Random Forest Regressor	0.074	0.163

(Fuente: elaboración propia)

Tabla 21: Resultados de predicción de CPC continuo considerando solo Messenger

Algoritmo de predicción	MAE [USD]	RMSE [USD]
Decision Tree Regressor	0.039	0.059
XGBoost Regressor	0.037	0.055
Random Forest Regressor	0.038	0.059

(Fuente: elaboración propia)

Tabla 22: Resultados de predicción de CPC continuo considerando solo Google Search

Algoritmo de predicción	MAE [USD]	RMSE [USD]
Decision Tree Regressor	0.135	0.183
XGBoost Regressor	0.092	0.142
Random Forest Regressor	0.102	0.152

(Fuente: elaboración propia)

Tabla 23: Resultados de predicción de CPC continuo considerando solo Shopping

Algoritmo de predicción	MAE [USD]	RMSE [USD]
Decision Tree Regressor	0.085	0.129
XGBoost Regressor	0.068	0.109
Random Forest Regressor	0.075	0.117

(Fuente: elaboración propia)

Tabla 24: Resultados de predicción de CPC continuo considerando solo Display

Algoritmo de predicción	MAE [USD]	RMSE [USD]
Decision Tree Regressor	0.071	0.120
XGBoost Regressor	0.073	0.111
Random Forest Regressor	0.067	0.116

(Fuente: elaboración propia)

11.3. Hiperparámetros de cada algoritmo optimizados a través de búsqueda de grilla (sin considerar redes neuronales)

Tabla 25: Conjunto de hiperparámetros optimizados por búsqueda de grilla para algoritmo Decision Tree Classifier [22]

Hiperparámetros	Valores testeados
Criterion	Gini, Entropy
Max Depth	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12
Splitter	Best, Random
Min Samples Split	0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4
Max Features	Auto, sqrt, log2, None

(Fuente: elaboración propia)

Tabla 26: Conjunto de hiperparámetros optimizados por búsqueda de grilla para algoritmo Random Forest Classifier [22]

Hiperparámetros	Valores testeados
Criterion	Gini, Entropy
Max Features	Auto, sqrt, log2, None

(Fuente: elaboración propia)

Tabla 27: Conjunto de hiperparámetros optimizados por búsqueda de grilla para algoritmo XGBoost Classifier [21]

Hiperparámetros	Valores testeados
Max Depth	3, 4, 5, 7
Learning Rate	0.01, 0.05, 0.1
Gamma	0, 0.25, 1
Scale Pos Weight	1, 3, 5

(Fuente: elaboración propia)

Tabla 28: Conjunto de hiperparámetros optimizados por búsqueda de grilla para algoritmo Decision Tree Regressor [22]

Hiperparámetro	Valores testeados
Splitter	best, random
Max Depth	1, 3, 5, 7, 9, 11, 12
Min Samples Leaf	5, 6, 7, 8, 9, 10
Min Weight Fraction Leaf	0.1, 0.4, 0.5
Max Features	Auto, log2, sqrt, None
Max Leaf Nodes	None, 10, 20, 30, 90

(fuente: elaboración propia)

Tabla 29: Conjunto de hiperparámetros optimizados por búsqueda de grilla para algoritmo XGBoost Regressor [21]

Hiperparámetro	Valores testeados
Learning rate	0.05, 0.07, 0.3
Max Depth	5, 6, 7
Subsample	0.7, 0.9
Colsample Bytree	0.7, 0.9
N Estimators	200, 500

(fuente: elaboración propia)

Tabla 30: Conjunto de hiperparámetros optimizados por búsqueda de grilla para algoritmo Random Forest Regressor [22]

Hiperparámetro	Valores testeados
Max Depth	80, 90, 100, 100
Max Features	2, 3
Min Samples Leaf	3, 4, 5
Min Samples Split	8, 10, 12
N_estimators	100, 200, 300, 1000

(fuente: elaboración propia)

11.4. Procedimiento de optimización de hiperparametros para el algoritmo de redes neuronales

Tabla 31: Conjunto de hiperparametros considerados y sus valores testeados

Hiperparametros	Valores testeados
Batch Size	20, 60, 100, 150
Cantidad de épocas	50, 100, 150, 200, 250
Nodos de capa neuronal 1	20, 60, 100
Nodos de capa neuronal 2	20, 60, 100
Nodos de capa neuronal 3	20, 60, 100
Nodos de capa neuronal 4	20, 60, 100
Dropout Rate 1	0.1, 0.2, 0.5
Dropout Rate 2	0.1, 0.2, 0.5
Función de activación 1	Softmax, Softplus, Softsign, ReLu, Tanh, Sigmoid, Hard Sigmoid, Linear
Función de activación 2	Softmax, Softplus, Softsign, ReLu, Tanh, Sigmoid, Hard Sigmoid, Linear
Función de activación 3	Softmax, Softplus, Softsign, ReLu, Tanh, Sigmoid, Hard Sigmoid, Linear
Función de activación 4	Softmax, Softplus, Softsign, ReLu, Tanh, Sigmoid, Hard Sigmoid, Linear
Optimizer	SGD, RMSprop, Adagrad, Adadelata, Adam, Adamax, Nadam

(Fuente: elaboración propia)

Tabla 32: Orden de optimización de hiperparametros por medio de procedimiento de búsqueda de grilla

Número de iteración	Hiperparametros optimizados por búsqueda de grilla en esa iteración
1	Batch Size, Cantidad de épocas
2	Funciones de activación 1, 2, 3 y 4
3	Nodos de capa neuronal 1, 2, 3 y 4
4	Dropout Rate 1 y 2

(Fuente: elaboración propia)

11.5. Situación de ejemplo expuesta en cuestionario para expertos

7. ID: 97

¿Cuánto pronostica que tendrá de **costo por clic (CPC)** un anuncio publicado a nivel semanal que tiene las siguientes características?

Cliente: ██████████

Plataforma de publicación: **Instagram**

Dispositivo de impresión: **Smartphone**

Cantidad de plataformas simultaneas utilizadas para el cliente la misma semana de publicación: **1**

Número de semana consecutiva en la cual en anuncio será publicado: **3**

Días de la semana publicado: **Lunes - Martes**

Publicación corresponde a época post-covid (después del mes de marzo del 2020): **Si**

Indicador Cyber: **No**

Indicador Retiro AFP: **No**

Indicador Navidad: **No**

Indicador Día del(la) Niño(a): **No**

Indicador Día de la Madre: **No**

Indicador Día del Padre: **No**

Indicador Vuelta a Clases: **No**

Indicador San Valentin: **No**

Costo de publicación: **7.23 USD (5413 CLP)**

*

Mayor a 0.087 [USD] (65 CLP)

Menor o igual a 0.087 [USD] (65 CLP)

Figura 33: Ejemplo de situación expuesta en cuestionario para ser resuelta por algún(a) experto(a)

11.6. Ejemplos de publicaciones para cada una de las plataformas

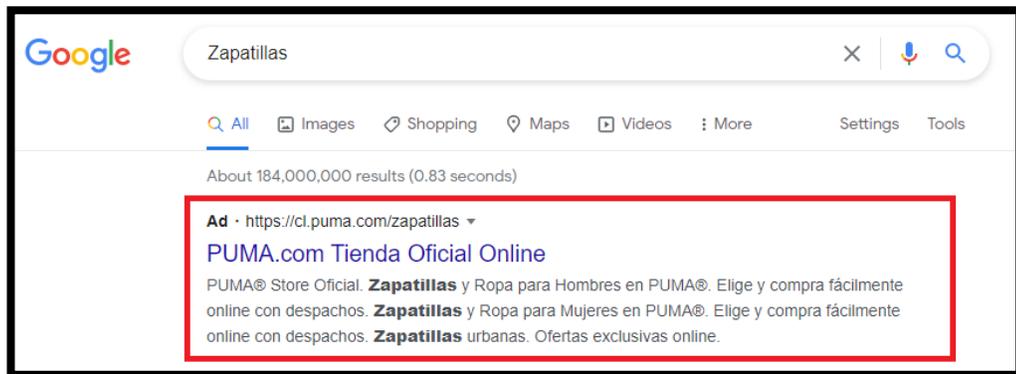


Figura 34: Anuncio de zapatillas en Google Search
(Fuente: Google.com)



Figura 35: Anuncio de educación impreso en el Feed de Facebook en un navegador web

(Fuente: Facebook.com)

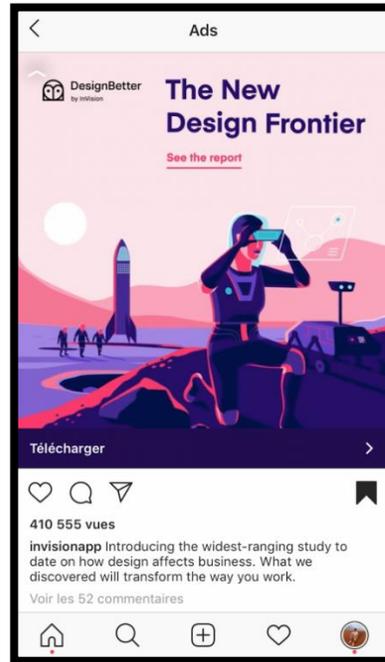


Figura 36: Anuncio impreso en el feed de la aplicación móvil Instagram
(Fuente: aplicación de Instagram para Smartphone)

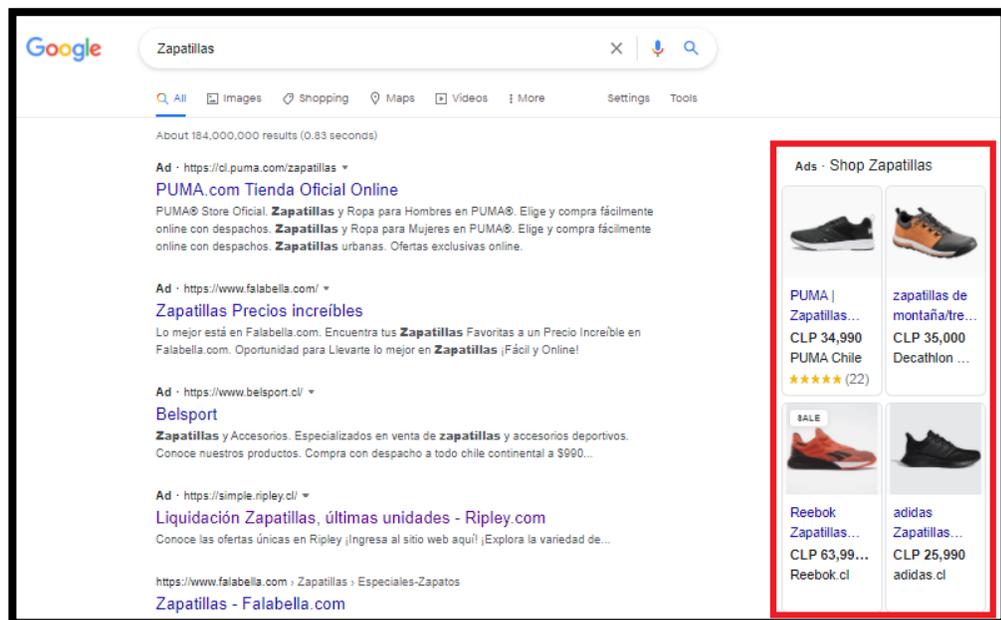


Figura 37: Anuncio de zapatillas en Shopping
(Fuente: Google.com)



Figura 38: Anuncio de guitarras desplegado en alguna parte de un sitio web
(Fuente: Anuncio público insertado en un sitio web)

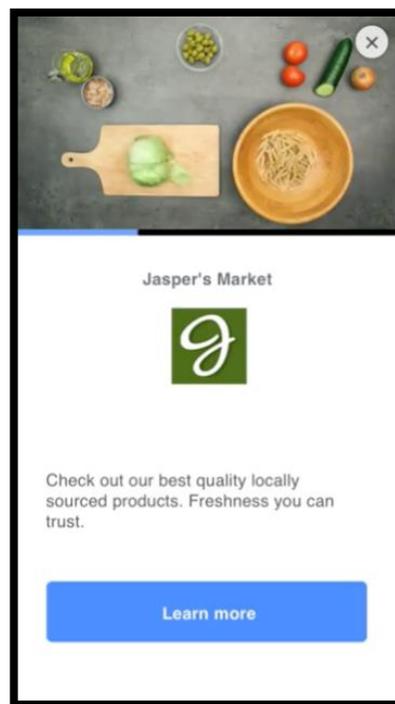


Figura 39: Anuncio impreso en medio del uso de una aplicación que admite publicidad
(Fuente: Anuncio público insertado en una aplicación)

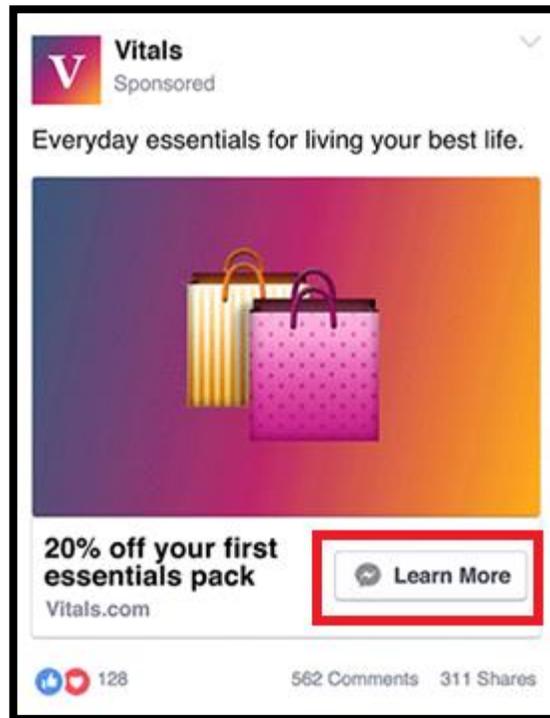


Figura 40: Anuncio impreso en el feed de la red social Facebook, en cual desencadena una conversación al hacer clic

(Fuente: Facebook.com)

11.7. Resultado de Shapely Values para modelos XGBoost Classifier (umbral diferenciado)

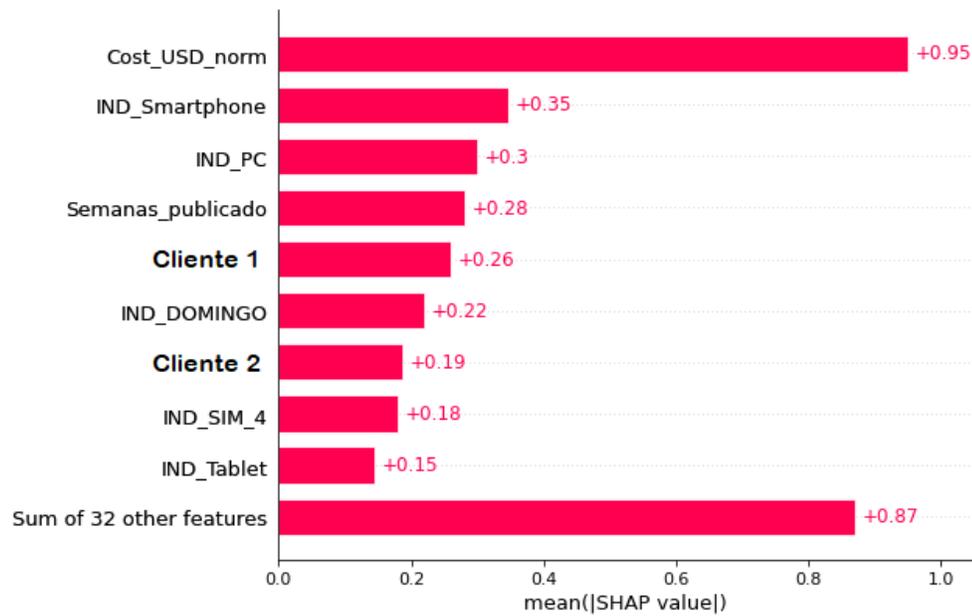


Figura 41: Shapely Values para la plataforma Facebook

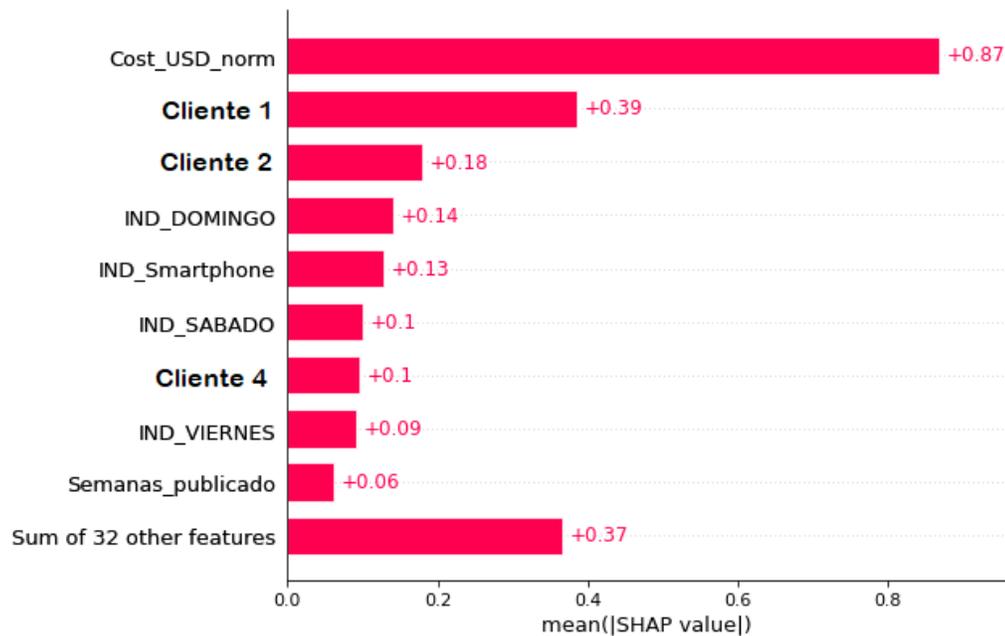


Figura 42: Shapely Values para la plataforma Instagram

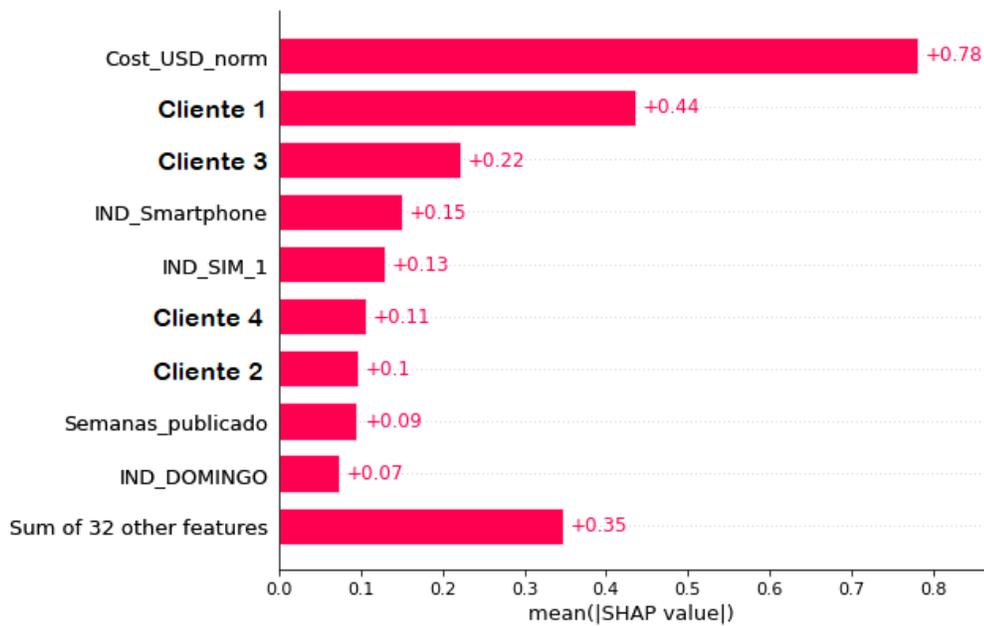


Figura 43: Shapely Values para la plataforma Google Search