

Predicción del Índice de Promotores Neto (*Net Promoter Score*) de Netflix en la población hispanohablante de Google Play Store mediante *Machine Learning* y Raspado Web (*Web Scraping*) para el primer semestre de 2024

Elaborado por:

Daniel Santiago González Citelly

Ivonne Maritza Martínez Colmenares

Mayra Dahian Rosero Franco

Andrés Marín Valencia

Universidad EAN

Especialización en Machine Learning

Especialización en Gerencia de Proyectos

Bogotá

02/12/2024

Tabla de contenido

Resumen.....	4
Planteamiento de Investigación	6
Descripción del problema	7
Pregunta de investigación	8
Objetivos	9
Objetivo General.....	9
Objetivos específicos	9
Justificación	11
Marco Teórico.....	13
Índice de Promotores Neto (Net Promoter Score)	13
Modelo: Representaciones de Codificador Bidireccional de Transformadores (BERT, Bidirectional Encoder Representations from Transformers)	17
Estado Del Arte.....	19
Marco Institucional	23
Marco Legal	27
Metodología	31
Enfoque, alcance y diseño de la investigación.....	31
Análisis y discusión de los resultados.....	38
1. Recolección y análisis de datos.....	38
2. Validación y Evaluación del Desempeño del Modelo	48
3. Discusión del modelo y resultados	

4.Contribución al campo y futuras aplicaciones
57

Conclusiones59

Referencias.....60

Resumen

En este trabajo de investigación presenta el análisis del *Net Promoter Score* (NPS) de Netflix en la población hispanohablante de Google Play Store, utilizando técnicas de *Machine Learning* sobre las reseñas y calificaciones de usuarios durante el primer semestre de 2024. La investigación aborda el problema de la satisfacción y fidelización del cliente mediante la clasificación de comentarios y el análisis de sentimiento, empleando herramientas como *Web Scraping* y *TextBlob* para la extracción y el análisis de datos. Las teorías fundamentales se basan en modelos de regresión y técnicas de Procesamiento de Lenguaje Natural (NLP, por sus siglas en inglés). Para llevar a cabo este análisis, se aplican métodos de *Scraping* para la recolección de reseñas de usuarios, que posteriormente se analizan utilizando el modelo “Representaciones de Codificadores Bidireccionales a partir de Transformadores” (BERT, por sus siglas en inglés), para optimizar la interpretación de las emociones y el contexto de las opiniones. El modelo predictivo desarrollado muestra un alto rendimiento, con un puntaje *F1* de 87.24% y un *AUC* superior a 0.9, lo que permite una precisa determinación del NPS y facilita la identificación de patrones de satisfacción y lealtad. Los resultados destacan la efectividad de las técnicas de *Scraping* y el análisis de sentimientos con BERT, evidenciando herramientas robustas para medir la satisfacción del cliente y proporcionando valiosos *insights* para la toma de decisiones en el entorno empresarial, con aplicaciones potenciales en otras plataformas y sectores.

Palabras clave: Índice de Promotores Neto (*Net Promoter Score*), análisis de sentimiento, *Machine Learning*, Netflix, Raspado Web (*Web Scraping*), Procesamiento de Lenguaje Natural.

Abstract

This research paper presents the analysis of the Net Promoter Score (NPS) of Netflix in the Spanish-speaking population of Google Play Store, using Machine Learning techniques on user reviews and ratings during the first semester of 2024. The research addresses the issue of customer satisfaction and loyalty through the classification of comments and sentiment analysis, employing tools such as Web Scraping and TextBlob for data extraction and analysis. The foundational theories are based on regression models and Natural Language Processing (NLP) techniques. To carry out this analysis, scraping methods are applied for the collection of user reviews, which are subsequently analyzed using the "Bidirectional Encoder Representations from Transformers" (BERT) model to optimize the interpretation of emotions and the context of the opinions. The developed predictive model shows high performance, with an F1 score of 87.24% and an AUC greater than 0.9, allowing for precise determination of the NPS and facilitating the identification of satisfaction and loyalty patterns. The results highlight the effectiveness of Scraping techniques and sentiment analysis with BERT, showcasing robust tools for measuring customer satisfaction and providing valuable insights for decision-making in the business environment, with potential applications in other platforms and sectors.

Palabras clave: Índice de Promotores Neto (*Net Promoter Score*), Análisis de sentimiento (*Sentiment analysis*), Aprendizaje automático (*Machine learning*), Netflix, Raspado web (*Web Scraping*), Procesamiento de lenguaje natural (*Natural Language Processing*).

Planteamiento de Investigación

El siglo XXI ha sido símbolo incuestionable de revolución tecnológica y consiguientemente ha traído una nueva forma de ver a la industria del entretenimiento, sufriendo una metamorfosis, para ponerlo en términos kafkianos, en el consumo del séptimo arte que ha migrado de forma masificada a la pantalla chica, cambio ratificado más aún con las condiciones socioeconómicas durante y después de la pandemia del COVID-19, donde cobró relevancia la irrupción de plataformas de *streaming* bajo demanda como Netflix, HBO (Max), Amazon Prime, Disney Plus, entre otras, para los hogares a nivel global.

Dichas plataformas no solo han cambiado la forma en que los consumidores acceden al contenido e intensifican las exigencias de estos servicios, sino que también han replanteado el cómo las empresas pueden mejorar tanto la calidad del producto como su posicionamiento y lograr responder a los diferentes formatos donde los usuarios expresan su experiencia como un factor esencial diferenciador en la actualidad.

Uno de los sistemas utilizados para evaluar la satisfacción del cliente es el *Net Promoter Score* (NPS), que mide la probabilidad de que un cliente recomiende un servicio o una marca a otra u otras personas (Zendesk, 2023). El enfoque propuesto aborda un desafío clave en el contexto digital actual, caracterizado por el elevado número de interacciones y experiencias de los usuarios: desarrollar una herramienta que permita capturar y analizar los datos de manera más precisa, facilitando un análisis profundo de la experiencia del usuario y su lealtad hacia la marca (Bocchino, 2023; Inesdi Business School, 2022; Budiu, 2023).

Descripción del problema

Actualmente, es clara la necesidad de contar con una experiencia de usuario *Customer Experience* (CX) bien estructurada en plataformas *streaming*. La marca objetivo de este estudio es Netflix, dado que cuenta con una gran variedad de público y opiniones sobre la industria e intereses segmentados, basados en una apreciación subjetiva, por parte del espectador, de lo que constituye un contenido audiovisual o entretenimiento en sí.

Aunque el proyecto de investigación enfocará sus esfuerzos en la marca Netflix, el desarrollo de un modelo así será un insumo extrapolable a cualquier área correspondiente a la experiencia de los usuarios, ya que mostrará un formato personificado sobre la percepción que decantará con mayor afinidad las segmentaciones de los clientes y proporcionará un entendimiento global de los comportamientos y sentimientos de estos.

Es importante señalar que el NPS se ha fundamentado principalmente en una clasificación de tres rangos: detractores (0-6), neutros (7-8) y promotores (9-10). Sin embargo, esta clasificación presenta limitaciones significativas, como lo expone NN Group (2023) en su video *Downsides of the Net Promoter Score*. El problema radica en el alto margen de error al evaluar la experiencia y la lealtad de los usuarios hacia una marca. Por ejemplo, una reseña con una calificación de 6 se considera un detractor, al igual que una reseña con una calificación de 0. Esta falta de matices en la clasificación puede distorsionar la comprensión real de la lealtad del cliente. Por lo tanto, se

promueve la búsqueda de soluciones, como el uso de *Machine Learning*, para estimar el NPS con mayor precisión y mejorar la efectividad de los resultados.

Pregunta de investigación

¿Cómo implementar técnicas de Raspado Web (*Web Scraping*) para recopilar las opiniones de los usuarios hispanohablantes de la aplicación de Netflix con el fin de comprender su nivel de satisfacción y lealtad hacia la marca, utilizando el Índice de Promotores Neto (*Net Promoter*) *Score* como indicador del grado de recomendación?

Objetivos

Objetivo General

Implementar técnicas de Raspado Web (*Web Scraping*) para recopilar las opiniones de los usuarios hispanohablantes de la aplicación de Netflix en Google Play Store el primer semestre del 2024, con el fin de comprender su nivel de satisfacción y lealtad hacia la marca, utilizando el Índice de Promotores Neto (*Net Promoter Score*) como indicador del grado de recomendación.

Objetivos específicos

1. Implementar técnicas de Raspado Web (*Web Scraping*) que permitan recolectar opiniones por parte de los usuarios de Netflix disponibles en la Play Store de Google, para lograr una vista con mayor detalle sobre el nivel de satisfacción y lealtad de los usuarios con la marca.
2. Desarrollar un modelo predictivo de estimación de la satisfacción de los usuarios de Netflix, empleando técnicas avanzadas de Aprendizaje Automático (*Machine Learning*) y seleccionando las variables más pertinentes a partir de los hallazgos obtenidos en investigaciones previas.

3. Evaluar, mediante métricas de rendimiento, la efectividad del modelo de estimación del NPS para determinar su grado de viabilidad para la toma de decisiones por parte de Netflix.

Justificación

Esta investigación está fundamentada según los lineamientos de la Universidad EAN desde sus diferentes estadios, a saber:

Conveniencia de la Investigación: ofreciendo una herramienta predictiva crucial para Netflix, capaz de estimar el *Net Promoter Score* (NPS) a partir de reseñas y calificaciones de usuarios en Google Play Store. Esto no solo facilitará la identificación de patrones en el *feedback* de los usuarios, permitiendo a la plataforma abordar proactivamente procesos de mejora en la experiencia del cliente, sino que también los resultados de este proyecto podrían ser aplicados a otras plataformas similares, proporcionando una alternativa valiosa para evaluar la satisfacción y lealtad de sus usuarios.

Relevancia Social: enfocado en la satisfacción del cliente como estrategia fundamental para la fidelización en el sector de **streaming** bajo demanda. Ofreciendo a su vez un análisis detallado del NPS y su correlación con las opiniones de los usuarios apunta a elevar la calidad del servicio y mejorar la satisfacción general de los consumidores.

Implicaciones Prácticas: La metodología empleada, que incluye **Web Scraping** y análisis de sentimientos, representa un enfoque innovador en la gestión de la experiencia del usuario y los resultados de éste podrían convertirse en un insumo esencial para la toma de decisiones estratégicas en Netflix quien, tomando como perspectivas clave (*insights*) los hallazgos sobre la percepción del

cliente, podrá implementar cambios que incrementen la retención de suscriptores y, por ende, su rentabilidad.

Valor Teórico: Esta investigación enriquecerá la literatura sobre el NPS y el análisis de sentimientos en el contexto del *Machine Learning*, proponiendo modelos y teorías que pueden ser aplicables en diversas industrias.

Utilidad Metodológica: La combinación de técnicas de análisis de datos con estudios del comportamiento del consumidor y sentimentología, aporta valor a la investigación actual y establece un marco metodológico replicable en futuros estudios.

Este trabajo se enmarca en el campo de Ciencia, tecnología e innovación, en el grupo de investigación de Ciencias Básicas bajo la línea de investigación Estadística Aplicada y Ciencia de Datos. Su contribución se enfoca en el análisis de datos a través de técnicas avanzadas, como machine learning y web scraping, para evaluar la experiencia del cliente y proporcionar conocimiento práctico sobre el comportamiento del consumidor. También es relevante en campos como el emprendimiento digital y el marketing inteligente.

Fuera del marco institucional, el enfoque de este proyecto de investigación se puede llevar a áreas de Administración y Marketing, complementando la investigación de Comportamiento del Consumidor, y de esta manera contribuir al conocimiento teórico y práctico en la experiencia del cliente y el *machine learning*.

Marco Teórico

Las investigaciones preliminares relacionadas con este estudio son diversas y presentan hallazgos significativos que ofrecen un marco conceptual robusto. Este marco no solo facilita la comprensión, sino que también valida la necesidad de explorar cómo las emociones de los usuarios en diferentes plataformas sociales pueden determinar niveles de recomendación. Este enfoque permitirá la implementación de estrategias efectivas para atraer y retener clientes, enriqueciendo el marco teórico con enfoques y metodologías que se aplicarán en el estudio. Esto proporcionará una base sólida y un conocimiento más profundo sobre el tema.

Índice de Promotores Neto (Net Promoter Score)

El *Net Promoter Score* (NPS) o, en español, Índice de Promotores Neto, es una métrica que permite evaluar el nivel de satisfacción del cliente mediante encuestas que capturan tanto aspectos cualitativos como cuantitativos. Su propósito es investigar la probabilidad de que los usuarios recomienden productos y servicios a otros, lo que proporciona una estimación de su lealtad hacia la marca (Sosa, 2023). Esta metodología fue introducida por Reichheld (2003) en su artículo "*The One Number You Need to Grow*", publicado en Harvard Business Review.

La métrica NPS es fundamental para el seguimiento y la comprensión de los factores que influyen en la puntuación. Las preguntas bien estructuradas son útiles tanto para clientes satisfechos como para aquellos insatisfechos, facilitando así la identificación de las causas de las experiencias negativas. Esta información permite a los departamentos correspondientes

implementar estrategias específicas para mejorar la satisfacción del cliente, un aspecto crítico en el desarrollo de modelos de *Machine Learning* enfocados en la retención y fidelización del usuario (Jiménez Zapata & Llaury Sihuincha, 2021).

Las plataformas *streaming*, utilizan encuestas sencillas que buscan medir el nivel de satisfacción y de este modo mejorar su oferta. La pregunta generalmente es “¿Con qué probabilidad recomendarías el servicio a un amigo y se pone en una escala de 0 –10?” el resultado se clasifica en **Promotores** de (9- 10) **pasivos** de (7-8) y **detractores** de (0-6) y el cálculo **NPS será % de promotores - % Detractores** (Inesdi, 2021).

El resultado del cálculo NPS explica que, cuanto mayor sea el NPS, mayor será la proporción de promotores en relación con los detractores, entonces la experiencia de la mayoría de los clientes es satisfactoria y que hay una alta probabilidad de recomendación por parte de ellos (González & Pérez, 2022).

Extracción de Datos Web (Web Scraping)

El *Web Scraping* es una técnica utilizada para la recolección de información, esencial para el cálculo del *Net Promoter Score* (NPS). Esta técnica permite obtener datos de diferentes sitios web, y en el contexto de esta investigación, está centrado en las mediciones realizadas por Google Play Store. Este proceso implica el uso de programas o *scripts* que analizan el contenido de las páginas web y extraen datos específicos (Smith, 2020). Su principal objetivo, como lo establece Johnson & Lee (2021) es obtener información que no está disponible a través de bases de datos o

APIs, las cuales han sido desarrolladas e integradas en diversas aplicaciones o plataformas utilizadas por diferentes áreas de las empresas.

Existen dos tipos principales de *Web Scraping*: el *scraping* estático, que se aplica a páginas *web* cuyo contenido no cambia con frecuencia y está en formato HTML, y el *scraping* dinámico, que se utiliza en sitios *web* donde la información es generada dinámicamente mediante JavaScript (Davis, 2022). Entre las herramientas más comunes para desarrollar *Web Scraping* se encuentran lenguajes de programación como Python, utilizando bibliotecas como BeautifulSoup y Scrapy (Martin, 2023).

A pesar de sus ventajas, *Web Scraping* presenta varios desafíos. Muchas aplicaciones bloquean las solicitudes si detectan actividad de *scraping*, lo que puede retrasar el desarrollo de dichas solicitudes (Thompson, 2022). Además, el marco legal se ha vuelto más estricto en relación con los servicios y solicitudes de datos, variando en diferentes países (Wang & Chen, 2023).

Asimismo, existen aplicaciones más simples que recopilan información de manera común, como lo es la librería de Python Google Play Scraping (Olano, n.d.), simbolizando una entrada eficiente a información de dicha tienda virtual sin vulnerar las políticas de información sensible.

Machine Learning

El *Machine Learning* es un subconjunto de la inteligencia artificial (IA) donde las computadoras aprenden de los datos y mejoran con la experiencia sin ser programadas

explícitamente (ZHOU, 2021). El uso de *machine learning* en el análisis de NPS en plataformas como Netflix, ha ganado relevancia ya que le permite profundizar en la comprensión de la lealtad de sus suscriptores y mejorando las estrategias de fidelización.

Las plataformas de transmisión bajo demanda (*streaming*) podrían hacer uso del análisis de sentimientos por parte de los suscriptores o usuarios, extrayendo comentarios como insumo cualitativo y las puntuaciones como insumo cuantitativo para la medición y predicción del NPS.

Las aplicaciones del *Machine Learning* permiten desarrollar modelos tanto de estimación como predictivos, como los modelos de clasificación en categorías de positivo, negativo o neutro. Estos modelos, cuyo desarrollo se debe a las contribuciones de pioneros como Arthur Samuel (1959) y Frank Rosenblatt (1958), permiten analizar la satisfacción y lealtad al servicio a través de puntuaciones NPS y características del usuario. Con algoritmos de series temporales, es posible identificar patrones en la puntuación NPS, lo que permite a las plataformas anticipar cambios en la lealtad de los suscriptores.

Modelo: Representaciones de Codificador Bidireccional de Transformadores (BERT, Bidirectional Encoder Representations from Transformers)

Las Representaciones de Codificador Bidireccional de Transformadores (BERT) son un modelo fundamental en el campo del procesamiento del lenguaje natural (NLP) que ha revolucionado la comprensión del lenguaje humano por parte de las máquinas. Desarrollado por Google en 2018, BERT emplea una arquitectura de transformadores que permite al modelo captar el contexto de las palabras de manera bidireccional. Esto significa que el modelo considera el contexto de una palabra no solo a partir de las palabras que la preceden, sino también de aquellas que la siguen, lo que resulta en una representación más precisa del significado contextual de las palabras (Devlin et al., 2019).

Una innovación clave de BERT es su enfoque de preentrenamiento, que se basa en dos tareas principales: el enmascaramiento de palabras y la predicción de la siguiente oración. A través de este preentrenamiento en un corpus de texto extenso, BERT aprende representaciones contextuales que pueden ser transferidas a diversas tareas de NLP. Esto permite que el modelo sea fácilmente ajustado para aplicaciones específicas utilizando una cantidad reducida de datos etiquetados, lo que lo convierte en una herramienta práctica y eficiente para tareas como el análisis de sentimientos y la traducción automática (Devlin et al., 2019).

La adopción de BERT ha establecido nuevos estándares en el rendimiento de tareas de NLP, demostrando mejoras significativas en varias competiciones y *benchmarks*. Su capacidad para comprender y generar lenguaje humano ha ampliado las posibilidades de interacción entre

humanos y máquinas, influyendo en el desarrollo de modelos posteriores que continúan construyendo sobre su arquitectura. La relevancia de BERT radica no solo en su eficacia, sino también en su contribución al avance de la inteligencia artificial en el lenguaje natural (Ghazvininejad et al., 2019).

Estado Del Arte

Hernández, J. (2023), realizó un estudio cuantitativo cuyos objetivos principales fueron desarrollar modelos y métodos para identificar y analizar series temporales relacionadas con las noticias en redes sociales. Este trabajo se enfocó en los usuarios de Twitter en México, y la investigación utilizó una muestra de tweets recopilados en marzo y abril de 2020. Usando las técnicas de *Web Scraping* y análisis de sentimientos, el artículo pretende establecer una correlación entre las expresiones de emociones en Twitter y los cambios en los indicadores económicos. Así, el objetivo general de la investigación consistió en desarrollar modelos predictivos sobre cómo las tendencias de las series temporales están afectadas por las emociones. Los resultados del artículo demostraron que las emociones expresadas en los *tweets* pueden influir en las tendencias del mercado y sugirieron que los modelos predictivos se vuelven más precisos, al considerar las emociones como factores que influyen en la toma de decisiones de un consumidor. El artículo está relacionado con la investigación actual ya que los dos trabajan en la dirección de cómo la actitud de los usuarios en las plataformas digitales influye en el mercado, y el resultado de Hernández Padilla evidenció que es crucial incluir los datos emocionales al hacer los modelos predictivos para ganar ventajas significativas al comprender la dinámica del mercado actual. Eso puede ser útil al medir la saturación o la retención de clientes de Netflix en Colombia.

Franco Santoliquido (2021) llevó a cabo un estudio cuantitativo centrado en la predicción de Índice de Promotores Neto (Net Promoter Score) (NPS) en una plataforma de comercio electrónico. Este estudio involucra un grupo de vendedores activos, seleccionando una muestra de 2,085 participantes en función de su volumen de transacciones. Usando encuestas y variables

transaccionales para alimentar un modelo de Ridge, se predice NPS incluso cuando las encuestas no han recopilado un número suficiente de respuestas. Los resultados mostraron que, el modelo elaborado, supera la eficacia de las soluciones existentes. Esta herramienta valoraría diariamente los valores de NPS, dando a los empleados una señal para identificar desviaciones en la experiencia del cliente y tomar decisiones informadas. La metodología operativa fue la siguiente: Extraer, transformar y cargar los datos, seguido de un análisis detallado de las variables relevantes. Estos dos informes ponen de relieve la necesidad de usar los datos a los que se puede acceder para predecir NPS y la necesidad de métricas precisas y ayudar para mejorar la experiencia honesta del cliente. El trabajo sugiere que la integración de modelos de Machine Learning puede producir soluciones en situaciones donde los datos son escasos.

Orozco Puello y Gómez Estrada (2019) en su trabajo de grado titulado “Desarrollo de un prototipo de aplicación web que permita la extracción de las ofertas laborales de las principales plataformas que postulan empleos en la región Caribe, usando la técnica *Web Scraping*”, realizado en la Universidad de la Costa en 2019, desarrollaron una investigación cuantitativa acerca de la población de búsqueda de empleo en la región Caribe de Colombia. Con la información publicada en las distintas plataformas mediante la técnica de scraping, los autores calcularon la disponibilidad de ofertas laborales en la región. El objetivo general fue crear un prototipo que permitiera recoger la información sobre las ofertas laborales y presentarla al usuario en un solo lugar. Los resultados indican que la aplicación del prototipo reduce sustancialmente su tiempo de exposición. La metodología abarca un diseño en el que se contemplan el levantamiento de requerimientos, el análisis y diseño del sistema, el desarrollo de la plataforma web y las pruebas. De este modo, la aplicación satisface los requisitos funcionales y no funcionales seleccionados por

los autores quienes concluyen que la centralización de datos y los modelos predictivos son una excelente forma de mejora.

La consolidación de datos y el desarrollo de modelos predictivos permiten explorar cómo estos productos digitales pueden beneficiar a la sociedad. Mediante el análisis de las emociones de los usuarios y la aplicación de técnicas de machine learning, se busca no solo estimar el NPS, sino también identificar oportunidades para mejorar la experiencia del cliente, contribuyendo así a plantear estrategias efectivas de fidelización y retención.

López, Martínez y Rodríguez (2020), llevaron a cabo un estudio denominado “Análisis de sentimientos en reseñas de aplicaciones móviles: Un enfoque de Machine Learning”, el cual se centró en el análisis de sentimientos a través de las reseñas a diversas aplicaciones móviles y su relevancia sobre indicadores de experiencia del usuario, utilizaron técnicas de aprendizaje supervisado para recopilar comentarios de usuarios en la Google Play Store, confirmando que, fundamentado en las opiniones, se logra predecir de una manera efectiva, la satisfacción de los usuarios y, asimismo, la probabilidad de recomendación. Este método es relevante para la investigación, ya que sirve de insumo para entender más profundamente el NPS mediante técnicas similares al análisis de reseñas de Netflix en la Play Store.

Por su parte, González y Pérez (2022), exploraron cómo el análisis de sentimientos puede utilizarse para evaluar la satisfacción del usuario en plataformas de streaming, en su artículo “Evaluación de la satisfacción del usuario en plataformas de streaming a través de análisis de sentimientos”. En este caso, el estudio validó en diversas aplicaciones de streaming y empleando

técnicas de *Web Scraping* para la recolección de datos a través de las reseñas de los usuarios, el grado de correlación entre las emociones expresadas a través de las reseñas publicadas en las páginas web y la probabilidad de recomendación de dichas aplicaciones. Los resultados concluyeron que el análisis de sentimientos proporciona una visión significativa sobre la satisfacción del usuario, lo que permite establecer planes de acción por parte de las plataformas de streaming para así optimizar las estrategias enfocadas al mejoramiento de la experiencia del cliente, haciéndolo pertinente para el proyecto, ya que expone y soporta la utilidad del análisis de datos de usuarios para mejorar la comprensión del NPS en contextos (plataformas de entretenimiento streaming) similares a Netflix con la extracción de información en la web.

Marco Institucional

En los últimos veinticinco años, Netflix ha revolucionado la manera en que se produce y consume cine y televisión a nivel global. Desde su lanzamiento en 1997 en Estados Unidos, la plataforma ha sabido adaptarse a cada cambio del mercado, alterando las dinámicas de producción, distribución y consumo en el sector audiovisual. Este cambio ha cuestionado los modelos tradicionales de la industria del cine y la televisión, estableciendo un nuevo paradigma de consumo de contenido bajo demanda que ha transformado los hábitos de visualización de millones de personas en todo el mundo (Bocchino, N., 2023). Debido a su relevancia para la industria y su impacto en el desarrollo de la innovación se consideró como la empresa objeto de estudio del presente trabajo.

El objetivo de este marco es exponer los aspectos relevantes de Netflix utilizando un enfoque metodológico cualitativo.

Nombre y Ubicación

Netflix es una empresa de entretenimiento digital con sede en Los Gatos, California, Estados Unidos. Desde su fundación en 1997 por Reed Hastings y Marc Randolph, Netflix se ha convertido en la plataforma digital de streaming líder en el mundo (Netflix, n.d).

Sector Económico y CIIU

La actividad que realiza Netflix se encuentra dentro de las “actividades de programación y transmisión.” En Colombia su actividad CIIU es: 5911 Actividades de producción de películas cinematográficas, videos, programas, anuncios y comerciales de televisión (Netflix, n.d).

El sector económico al cual pertenece es el de entretenimiento y medios. Su modelo de negocio se basa en las suscripciones, en donde los clientes pagan un valor mensual por una amplia gama de contenido de video y audio; esto incluye series, películas y programas de televisión. Con presencia en más de 190 países (Organigrama10, 2024).

Nichos de Mercado y Productos

Netflix es una compañía enfocada en la categoría de la actividad de entretenimiento transmitida por Internet. Sus grupos objetivos incluyen: suscriptores de servicios de transmisión; fans de contenido original y exclusivo, consumidores que desean obtener la experiencia de visualización más conveniente (Organigrama10, 2024).

Los productos de la plataforma son series, películas, documentales, contenido para niños; licenciado y producido originalmente por Netflix (Organigrama10, 2024).

Procesos Clave

Adquisición y producción de contenidos, monitoreo y análisis de los datos con recomendaciones, desarrollo y mantenimiento de la plataforma, gestión de las suscripciones y la facturación (Netflix Tech Blog, 2020).

Estructura Organizacional

Netflix opera con una estructura organizativa "plana y minimalista" que permite contar con una cadena de mando relativamente corta que a la vez facilita una comunicación abierta y directa entre todos los niveles jerárquicos de la empresa. Esta compañía es claramente una organización con un fuerte sesgo por el enfoque integrado, centralizando, en su nivel superior, generando un importante grado de diversificación entre los departamentos que la componen. Algunas de las figuras presentes en el nivel superior son, el director ejecutivo de la compañía Reed Hastings y los Codirectores Ejecutivos Ted Sarandos y Greg Peters (Organigrama10, 2024).

Valor de la Cultura de Organización

La cultura de Netflix se basa en la libertad y la responsabilidad. Algunos de los valores y prácticas importantes son: Apoyar las decisiones independientes de los empleados, comunicación de información abierta, transparente y específica, énfasis en adquirir y retener empleados talentosos, evitar reglas, procedimientos innecesarios (Netflix Tech Blog, 2020).

Esta cultura ha sido un gran motor del éxito de Netflix, animando a la innovación y a la creatividad, pero también siendo fundamental en la capacidad para adaptarse a nuevos entornos, sin ninguna duda indispensable en un sector que evoluciona tan rápidamente (Netflix Tech Blog, 2020).

Se puede indicar que Netflix es líder en el campo del entretenimiento digital y el streaming, con una cultura corporativa enfocada en la autonomía, la responsabilidad y el desarrollo del talento. Su estructura ágil y sus procesos innovadores le permiten ofrecer contenido personalizado y de alta calidad a millones de usuarios en todo el mundo.

Marco Legal

El desarrollo de proyectos que involucran la recolección y análisis de datos mediante técnicas de *Web Scraping* plantea diversos desafíos legales, especialmente cuando se trabaja con información de plataformas como la Google Play Store, donde las reseñas de los usuarios pueden contener datos protegidos por normas de privacidad. En Colombia, la protección de datos personales la regula la Ley 1581 de 2012, conocida como la Ley de Habeas Data, que establece lineamientos para tratar de forma correcta la información personal. Adicionalmente, la propiedad intelectual relacionada con los contenidos digitales está protegida por la Ley 23 de 1982. Este marco legal tiene como objetivo garantizar que las actividades de *scraping* y análisis de datos en el presente proyecto cumplan con las regulaciones nacionales e internacionales vigentes, promoviendo el respeto a la privacidad y los derechos de los usuarios, así como la legalidad en el uso de las herramientas tecnológicas involucradas.

Protección de Datos Personales en Colombia (Ley 1581 de 2012)

En Colombia, la Ley 1581 de 2012 establece el marco normativo para tratar datos personales, para proteger los derechos de los titulares de la información y garantizar que su uso sea adecuado y respetuoso con su privacidad. En proyectos como el presente, que utiliza técnicas de *Web Scraping* para recolectar reseñas y calificaciones de usuarios de la Google Play Store, es fundamental comprender las implicaciones legales de manejar datos potencialmente sensibles. Aunque las reseñas y puntuaciones pueden parecer información pública, es crucial asegurarse de que no se recopile ningún dato personal identificable sin el debido consentimiento del usuario.

Artículo 3: Definición de datos personales

Este artículo define los datos personales como cualquier información que identifique o pueda identificar a una persona natural, ya sea directamente (nombre, cédula) o indirectamente (correo electrónico, número de teléfono asociado a otra base de datos) (Ley 1581, 2012). En proyectos de *scraping*, como en este caso, es fundamental que se recopilen únicamente datos no identificables públicamente, tales como reseñas anónimas y puntuaciones que no permitan vincular a un usuario específico

Artículo 8: Derechos de los titulares de los datos

El artículo 8 de la ley establece que los titulares de los datos personales tienen derecho a conocer, actualizar y rectificar la información que se recoja sobre ellos (Ley 1581, 2012). En el contexto de *Web Scraping*, esto implica que, si por error se llegara a recopilar información personal, los titulares de los datos deben ser informados y se les debe permitir ejercer sus derechos. Para evitar conflictos legales, el *scraping* debe configurarse de manera que recoja únicamente datos anonimizados o no sensibles, como el contenido textual de las reseñas y las valoraciones numéricas, sin asociarlas directamente a usuarios específicos.

Artículo 17: Obligaciones de los responsables del tratamiento de datos

El artículo 17 impone obligaciones a quienes gestionan datos personales, incluyendo la obligación de obtener autorización expresa de los titulares antes de procesar sus datos, y de garantizar su seguridad para evitar accesos no autorizados o fraudes (Ley 1581, 2012). En proyectos de scraping, es esencial asegurarse de que no se extraigan datos personales sin un consentimiento previo claro. La información que se recolectará (reseñas y puntuaciones) debe estar delimitada y usarse solo para análisis de satisfacción y lealtad, sin comprometer la identidad de los usuarios.

Ley 23 de 1982: Derechos de Autor.

El *web Scraping*, o extracción automatizada de datos de sitios web, se ha vuelto una práctica común en la recolección de información digital. Sin embargo, esta técnica presenta desafíos legales, especialmente en lo que respecta a la propiedad intelectual y los derechos de autor. En Colombia, la Ley 23 de 1982 protege los contenidos digitales, incluidos textos y reseñas publicadas en plataformas como la Google Play Store, lo que implica que su uso debe alinearse con las disposiciones legales establecidas.

Propiedad Intelectual en la Play Store (Ley 23 de 1982)

La Ley 23 de 1982 establece un marco legal para la protección de los derechos de autor en Colombia, y es particularmente relevante para los contenidos generados por los usuarios en plataformas digitales.

Artículo 11: Derechos de autor sobre contenidos originales

Este artículo reconoce que los autores de obras originales gozan de derechos sobre sus creaciones, lo que incluye textos, imágenes y reseñas publicadas en plataformas como la Google Play Store (Ley 23, 1982). En el caso de las reseñas de usuarios, aunque estas sean generadas por los mismos usuarios, la ley protege su expresión y los derechos de autor asociados a estas contribuciones.

Artículo 15: Adaptación y modificación de obras

El artículo establece que una persona que, con el permiso expreso del autor o de sus herederos, adapta, transporta, modifica, extrae, compendia o parodia una obra del dominio privado se convierte en titular del derecho de autor sobre su nueva creación (Ley 23, 1982). Si se planea utilizar contenido de la Play Store (como reseñas y calificaciones de usuarios) para crear un nuevo producto o análisis (como el cálculo de un NPS), es crucial contar con el permiso explícito de Google o de los autores de las reseñas.

Metodología

Enfoque, alcance y diseño de la investigación

Este estudio adopta un enfoque cuantitativo, centrado en la recolección y análisis de datos numéricos obtenidos de reseñas de usuarios en Google Play Store. El alcance de la investigación es descriptivo y correlacional, permitiendo explorar las relaciones entre las opiniones de los usuarios y su nivel de satisfacción. Se utiliza un diseño no experimental, ya que no se manipulan variables, y la recolección de datos se realiza en un solo momento, lo que es característico de los estudios transversales.

Metodología de investigación

En este apartado se presenta la metodología de la investigación a través de la tabla 1, con el objetivo de proporcionar una visión clara y estructurada de las fases, actividades y variables consideradas en el proceso.

Tabla 1. Metodología de investigación

Fase de la Investigación	Actividades	Variables Consideradas
<i>Planificación y Diseño</i>	Identificación de fuentes de datos relevantes. Desarrollo de scripts para extracción de datos utilizando herramientas como BeautifulSoup y Selenium.	Opiniones textuales de los usuarios, fecha de publicación de las opiniones.
<i>Recolección de Datos</i>	Recolección de opiniones de usuarios empleando Web Scraping y se almacenarán en una base de datos estructurada.	Calificaciones numéricas, términos y frases clave.
<i>Población y Muestra</i>	La población incluye las opiniones de usuarios hispanohablantes sobre Netflix en la Google Play Store durante el primer semestre de 2024. Se seleccionará una muestra representativa a través de muestreo aleatorio estratificado, abarcando diferentes calificaciones (de 1 a 5 estrellas).	
<i>Limpieza y Preparación de Datos</i>	Eliminación de datos duplicados, detección de spam y formato de los datos para su preparación. Uso de técnicas de preprocesamiento de texto como tokenización y eliminación de caracteres especiales.	Datos procesados y filtrados por calidad y relevancia.

<i>Análisis de Datos</i>	Aplicación de técnicas de procesamiento de lenguaje natural (NLP) para la extracción de variables clave y análisis de tendencias en las calificaciones y opiniones, Aplicación del modelo Bert para el análisis del sentimiento.	Sentimientos asociados a las opiniones, palabras claves.
<i>Desarrollo del Modelo Predictivo</i>	Entrenamiento de modelos de machine learning, incluyendo BERT y ajustes de hiperparámetros. Configuración y optimización de AdamW como optimizador y paradas tempranas para evitar el sobreajuste.	NPS calculado como variable dependiente, embeddings de palabras.
<i>Validación del Modelo</i>	Validación cruzada del modelo y evaluación de métricas de rendimiento como precisión, Recall, F1-score y AUC. Evaluación de la curva ROC y la curva de Precisión-Recall.	Métricas de validación y error de predicción.

Nota. Elaborado por los autores, 2024.

Luego, se relacionan las definiciones conceptuales y operacionales de las variables, que se detallan en la Tabla 2.

Tabla 2 Definiciones conceptuales y operacionales

Variable	Definición Conceptual	Definición Operacional	Dimensiones
-----------------	------------------------------	-------------------------------	--------------------

NPS (Net Promoter Score)	Indicador que mide la probabilidad de que los usuarios recomienden un servicio o producto a otros.	Se mide a partir del análisis de las opiniones y calificaciones numéricas de los usuarios en la Google Play Store.	Recomendación, satisfacción, lealtad.
Análisis de Sentimiento	Proceso de identificar y categorizar opiniones expresadas en un texto para determinar la actitud hacia un tema.	Analizado usando técnicas de procesamiento de lenguaje natural (NLP) aplicadas a los comentarios mediante modelos como BERT para capturar el contexto y la polaridad del sentimiento.	Positivo, negativo, neutral.
Web Scraping	Técnica automatizada para extraer datos de sitios web.	Implementado mediante herramientas de Python como Beautiful Soup , Selenium y google-play-scraper , con	Extracción de datos, automatización, almacenamiento.

		almacenamiento estructurado en bases de datos.	
<i>Opiniones de Usuarios</i>	Comentarios y valoraciones que los usuarios hacen públicos sobre un producto o servicio.	Recolectadas mediante Web Scraping y analizadas usando NLP para evaluar la longitud, frecuencia de palabras clave y polaridad de las opiniones.	Calidad del comentario, temática, frecuencia de palabras clave.
<i>Calificaciones Numéricas</i>	Evaluación cuantitativa que los usuarios asignan a un producto o servicio en una escala de 1 a 5.	Extraídas de las reseñas de la Google Play Store y categorizadas en detractores (1-3), pasivos (4) y promotores (5) para el cálculo del NPS.	Distribución de puntuaciones, frecuencia de calificaciones.

<p>Fecha de Publicación de Opiniones</p>	<p>Momento en el que una opinión es publicada por un usuario.</p>	<p>Se registra automáticamente durante la recolección de datos y se utiliza para analizar tendencias temporales en la satisfacción.</p>	<p>Temporalidad, estacionalidad.</p>
<p>Términos y Frases Clave</p>	<p>Palabras y expresiones que reflejan sentimientos o ideas clave en los comentarios de los usuarios.</p>	<p>Identificadas y clasificadas mediante análisis de NLP para evaluar la relevancia y la frecuencia en los comentarios.</p>	<p>Relevancia, frecuencia, contexto de uso.</p>

Nota. Elaborado por los autores, 2024.

Selección de métodos o instrumentos para recolección de información

Se emplean técnicas de raspado web para recopilar reseñas de los usuarios, a través de herramientas como *BeautifulSoup* y *Selenium* para automatizar la extracción de las reseñas de los usuarios en Google Play Store. Estas reseñas se almacenan en un *DataFrame* de Python para facilitar su análisis.

Los instrumentos de medición son el **análisis de sentimiento** y *Net Promoter Score* (NPS), este último representa una métrica válida y eficiente para medir la satisfacción y lealtad de los usuarios. El análisis de sentimiento se realiza utilizando una biblioteca de procesamiento de lenguaje natural (NLP) como *TextBlob*, lo que facilita clasificar las reseñas en positivas o negativas. Estos instrumentos están diseñados para garantizar la coherencia y validez de la recolección de datos, asegurando que se midan las variables previamente definidas.

Técnicas de análisis de datos

Los datos recolectados se analizan mediante técnicas descriptivas y de procesamiento de lenguaje natural. De forma preliminar, se aplican métodos de Análisis Exploratorio de Datos (**EDA, por sus siglas en inglés**), para resumir las características de las reseñas, como la distribución de las calificaciones y los patrones de sentimiento.

Posteriormente, se generan técnicas de regresión que permitan observar los grados de relación entre las variables, en particular otorgadas por los usuarios y el NPS.

Durante el proceso de entrenamiento del modelo BERT, se emplean técnicas como la **tokenización**, que consiste en dividir el texto en unidades más pequeñas llamadas "**tokens**", los cuales pueden ser palabras completas o fragmentos de palabras. A continuación, estos tokens se transforman en *embeddings*, que son representaciones numéricas que capturan el significado contextual de cada palabra dentro de la oración, lo que permite al modelo comprender mejor las relaciones semánticas y lingüísticas presentes en el texto (Devlin et al., 2019).

A través de múltiples capas del modelo, estas representaciones o *embeddings* se ajustan para identificar patrones complejos en los datos. Una forma de optimizar el rendimiento del modelo es a través de técnicas como el *Early Stopping* (o parada temprana), que se establece para interrumpir el proceso de entrenamiento cuando el rendimiento del modelo en el conjunto de datos de validación comienza a deteriorarse (Interactive Chaos, 2024) o se detecta un posible sobreajuste, permitiendo mejorar así la capacidad del modelo para generalizar en datos no vistos previamente (Goodfellow et al., 2016).

Adicionalmente, se emplean pruebas de **validación cruzada** (*cross-validation*) para evaluar la precisión del modelo en diferentes subconjuntos de datos, contribuyendo a una interpretación más confiable de los resultados (Hastie, Tibshirani, & Friedman, 2009). Una vez generado los análisis al modelo y su precisión, se incluyen tablas y gráficos que detallan los instrumentos y técnicas aplicadas, asegurando la claridad en la presentación de los hallazgos.

Análisis y discusión de los resultados

1. Recolección y análisis de datos

En la primera parte del desarrollo de este trabajo, se utilizaron técnicas de **Web Scraping** para recopilar las valoraciones de los usuarios hispanohablantes de **Netflix** en la **Google Play Store** durante el primer semestre de 2024. Se emplearon Bibliotecas *Google-Play-Scraper*, *Pandas* y *re* para la extracción y limpieza de datos, con el fin de obtener información pertinente y

significativa de una forma eficiente, El proceso comenzó con la configuración de un script que permitió recolectar las reseñas más importantes, filtrando por idioma y año para asegurar que sólo se incluyeran opiniones de 2024. Esto facilitó la obtención de comentarios textuales y calificaciones numéricas importantes y necesarias para la implementación del modelo.

Se logró obtener con éxito un conjunto de datos sólido compuesto por 28.855 reseñas, que incluye datos significativos sobre las percepciones de los usuarios. Las valoraciones obtenidas abarcaron una variedad de calificaciones, de 1 hasta 5 estrellas. Además, se realizó un proceso para eliminar enlaces y caracteres espaciales, con el objetivo de obtener un texto adecuado para llevar a cabo el análisis del sentimiento.

El propósito de ofrecer datos relevantes y detallados acerca de la satisfacción y fidelidad de los usuarios se cumplió. La recolección de datos mediante **Web Scraping** resulto ser efectiva, y los datos obtenidos evidenciaron que tanto las valoraciones textuales como las puntuaciones numéricas son fundamentales para medir y valorar la satisfacción del cliente.

a. Desarrollo del Modelo Predictivo de Satisfacción

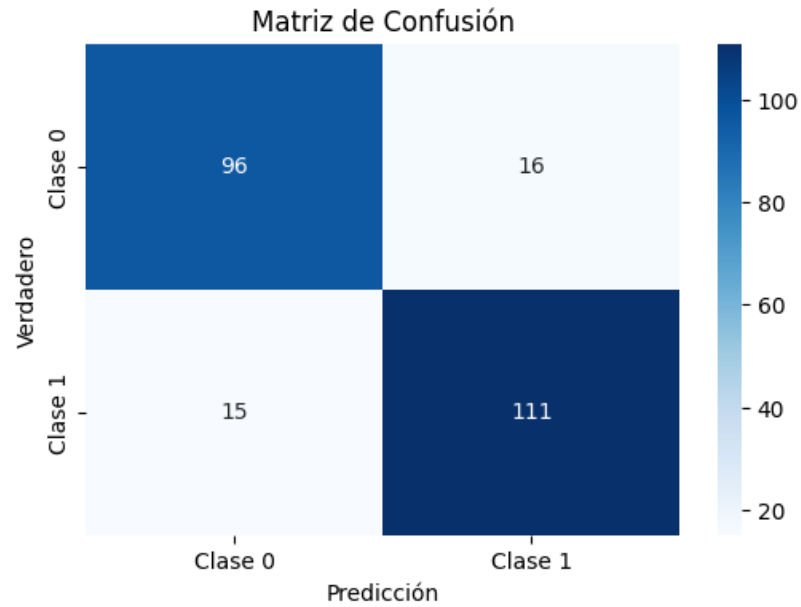
Se logró desarrollar un modelo predictivo empleando técnicas avanzadas de **Machine Learning**. se utilizó el modelo de **Representaciones de Codificador Bidireccional de Transformadores (BERT)** para el estudio y la categorización de los sentimientos de los suscriptores. Este modelo de procesamiento de lenguaje natural (**NLP**) es conocido por su

capacidad para comprender el contexto bidireccional de las palabras, lo que incrementa de forma significativa la exactitud en la interpretación de textos complejos.

Desarrollo y Resultados del Modelo: El modelo se entrenó con los datos obtenidos, utilizando *tokenización* y conversión de palabras a representaciones numéricas (*embeddings*) para captar su contexto. Se emplearon las siguientes fases y métricas para poder ejecutar y evaluar el modelo:

- **Preparación del Dataset:** Las reseñas se clasificaron según un límite de satisfacción (Calificaciones de 1 a 3 se consideraron negativas, 4 neutras y 5 positivas.).
- **Balanceo de Datos:** Para prevenir un sesgo de clase, se equilibró el dataset a través de métodos de sobre muestreo.
- **División de Datos:** El conjunto de datos balanceado se segmentó en subgrupos de entrenamiento (80%) y prueba (20%).
- **Entrenamiento del Modelo:** Se utilizó el modelo BERT con un optimizador **AdamW** y una tasa de aprendizaje adaptativa.

A continuación, se presenta en la Figura 1 la matriz de confusión, que muestra la comparación entre las predicciones del modelo y los valores reales.



Nota. Elaborado por los autores, 2024.

De acuerdo con lo mostrado en la matriz de confusión, se puede observar que el modelo logra clasificar correctamente las dos clases con las siguientes distribuciones:

- **Clase 0:**
 - Verdaderos negativos (TN): 96
 - Falsos positivos (FP): 16
- **Clase 1:**
 - Falsos negativos (FN): 15
 - Verdaderos positivos (TP): 111

Fórmulas y Resultados

Precisión (Precision)

La precisión mide qué proporción de las predicciones positivas fueron correctas.

$$\text{Pr e c i s i ó n} = \frac{TP}{(TP+FP)}$$

Ecuación 1. Pr e c i s i ó n

$$\text{Pr e c i s i ó n} = \frac{111}{111 + 16} \approx 0.8740$$

(Sensibilidad o Tasa de Verdaderos Positivos)

El recall mide qué proporción de los casos positivos reales fueron identificados correctamente.

$$\text{Recall} = \frac{TP}{(TP+FN)}$$

Ecuación 2. Recall

$$\text{Recall} = \frac{111}{111+15} \approx 0.8809$$

F1-Score

El F1-Score es la media armónica entre precisión y recall. Proporciona un equilibrio entre ambas métricas.

$$F1 - \text{Score} = 2 \times \frac{(\text{Pr e c i s i ó n} \times \text{Recall})}{(\text{Pr e c i s i ó n} + \text{Recall})}$$

Ecuación 3. F1 - Score

Ecuación 3. F1 – Score

$$F1 - Score = 2 \times \frac{0.8740 \times 0.8810}{0.8740 + 0.8810} \approx 0.8775$$

Exactitud (Accuracy)

La exactitud mide el porcentaje total de predicciones correctas, considerando todas las clases.

$$Exactitud = \frac{TP+TN}{TP+TN+FP+FN}$$

Ecuación 4. Exactitud

$$Exactitud = \frac{111 + 96}{111 + 96 + 16 + 15} \approx 0.8697$$

A continuación, se presentan los resultados de las métricas obtenidas, organizadas en la siguiente tabla, para dar una visión clara y detallada del desempeño y análisis de los datos.

Tabla 3. Resumen de Resultados

Métrica	Resultado
Precisión	0.8740
Recall	0.8809
F1-Score	0.8775
Exactitud	0.8697

Nota. Elaborado por los autores, 2024.

De acuerdo con los resultados presentados en la Tabla 3, se interpretan las principales tendencias y hallazgos, para analizar su significado y relevancia en la investigación.

b. Interpretación de los Resultados

- **Precisión Alta (87.40%):** Indica que las predicciones positivas realizadas por el modelo son correctas, lo que representa un bajo nivel de falsos positivos.
- **Recall Moderado (88.09%):** El modelo identifica correctamente una gran parte de los positivos reales, con un nivel reducido de falsos negativos.
- **F1-Score (87.75%):** El modelo mantiene un equilibrio razonable entre precisión y *recall*, este resultado demuestra la consistencia del modelo para predecir.
- **Exactitud General (86.97%):** Indica una tasa global de predicciones correctas, generando un desempeño confiable.

En la Figura 2 se presenta la curva ROC, la cual muestra la alta capacidad discriminativa del modelo, según los parámetros evaluados.

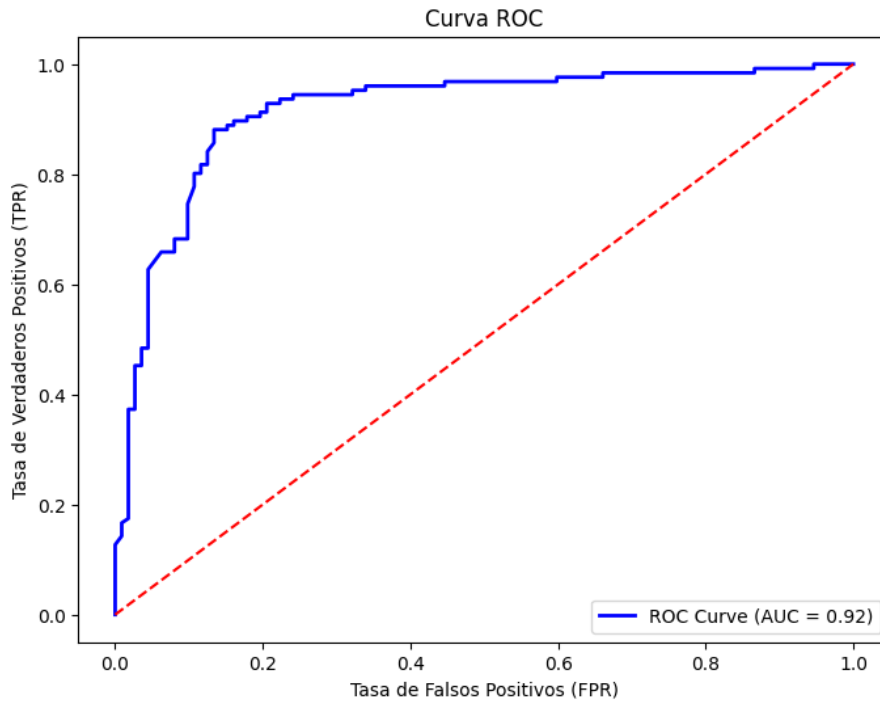


Figura 2. Gráfico de Curva ROC

Nota. Elaborada por los autores 2024

- **Tasa de Falsos Positivos (FPR):**

En el eje X se representa la proporción de instancias negativas erróneamente positivas. Los valores bajos para estos casos son ideales, toda vez que indican un menor número de errores en la clasificación de la clase negativa.

- **Tasa de Verdaderos Positivos (TPR o Sensibilidad):**

En el eje Y se mide la proporción de instancias positivas correctamente identificadas. Valores altos indican que el modelo logra identificar correctamente la mayoría de los casos positivos.

- **Líneas Roja y Azul:**

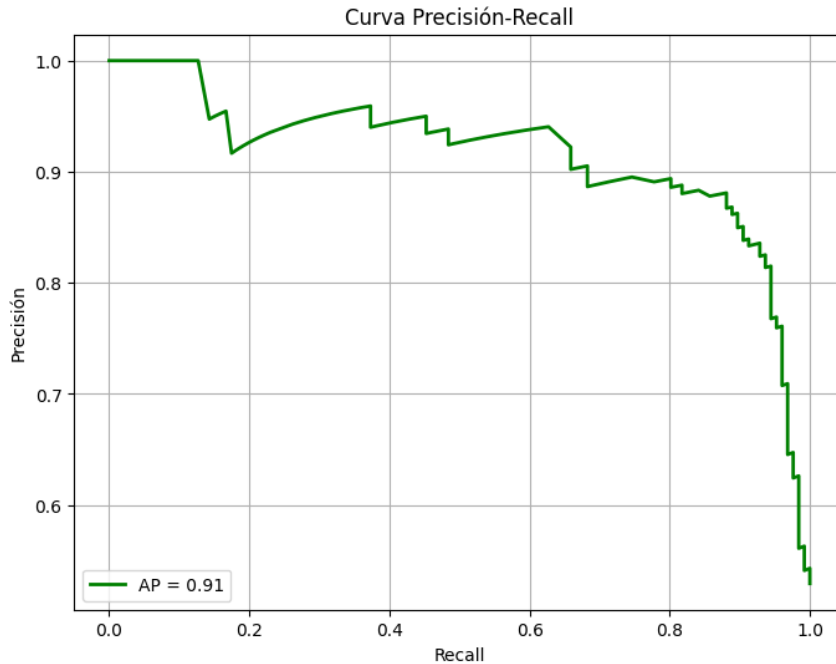
La línea roja representa el desempeño de un clasificador aleatorio, también llamado base, mientras que la curva azul, que refleja el desempeño del modelo, se sitúa notablemente por encima. Esto evidencia que el modelo supera ampliamente el nivel de azar, demostrando su capacidad para clasificar eficazmente las instancias.

- **AUC (Área Bajo la Curva):**

Valores de Referencia:

- AUC = 1.0: Clasificador perfecto.
 - AUC = 0.5: Clasificación aleatoria.
 - AUC > 0.9: Excelente rendimiento.
- **Análisis de la Curva ROC y AUC:** La curva ROC superó la línea de un clasificador aleatorio, con un AUC > 0.9, lo que indica que tiene un rendimiento excelente en la separación de clases. Esto ratifica que el modelo puede discriminar adecuadamente entre usuarios promotores y detractores.

En el siguiente gráfico (figura 3), se presenta la Curva Precisión-Recall



Nota. Elaborada por los autores 2024

A continuación, se presentan los análisis de los principales parámetros de la Curva Precisión-Recall, los cuales son fundamentales para evaluar el rendimiento del modelo en términos de clasificación de instancias positivas y negativas:

- **Precisión:** En el eje Y, representa la proporción de verdaderos positivos entre todas las predicciones positivas. La curva mantiene una alta precisión (~ 1.0) en la mayoría de los valores de recall, lo que indica que el modelo no tiene muchos falsos positivos.
- **Recall:** En el eje X, representa la proporción de verdaderos positivos detectados sobre el total de positivos reales. A medida que el recall se incrementa hacia 1.0

(detectando más positivos), la precisión comienza a disminuir. Esto es normal en muchos modelos, ya que aumentar el recall suele implicar un incremento en los falsos positivos.

- **Área bajo la curva (AP = 0.91):** Un valor cercano a 1 indica que el modelo logra un excelente balance entre precisión y recall. La alta precisión inicial (alrededor de 1.0) y la lentitud en la disminución demuestran que el modelo gestiona bien ambas métricas.

Interpretación

- **Desempeño:** El modelo tiene un desempeño destacado, identificando correctamente la mayoría de las instancias positivas con pocos falsos positivos.
- **Ventaja en problemas desbalanceados:** Este tipo de curva es particularmente útil cuando las clases están desbalanceadas, ya que ofrece más información que una curva ROC en esos escenarios.

2. Validación y Evaluación del Desempeño del Modelo

a. Calificaciones Reales vs. Etiquetas de Comentarios Estimadas

Para evaluar la satisfacción y fidelidad de los usuarios, se utilizó un conjunto de 28,855 reseñas recopiladas, calculando el Índice de Promotores Netos (NPS) desde dos perspectivas: por

un lado, a partir de las calificaciones numéricas otorgadas por los usuarios en la plataforma; y por otro, mediante el etiquetado automatizado de comentarios clasificados como positivos o negativos por el modelo.

En el enfoque basado en calificaciones numéricas, y considerando la escala de 1 a 5 estrellas, se establecieron las siguientes categorías estándar:

- **Promotores:** Usuarios que concedieron la máxima puntuación de 5 estrellas.
- **Detractores:** Usuarios que concedieron puntuaciones de 1 a 3 estrellas.
- **Neutrales:** Notas de 4 estrellas (que no se toman en cuenta en el cálculo del NPS).

El **NPS real** se obtuvo aplicando la fórmula estándar:

$$NPS = \left(\frac{\text{Promotores} - \text{Detractores}}{\text{Total de calificaciones}} \right) \times 100 \quad \text{Ecuación 5. NPS Real}$$

Resultados del Cálculo con Score (número de estrellas)

- **Promotores reales:** 14.835
- **Detractores reales:** 11.822
- **NPS real:** 11.30

En cuanto al NPS calculado a través del etiquetado producto del modelo se homologaron de la siguiente forma:

- **Comentarios positivos** = Promotores
- **Comentarios negativos** = Detractores

De tal manera que, El NPS estimado, se obtuvo aplicando la fórmula homóloga:

$$NPS = \left(\frac{\text{Comentarios Positivos} - \text{Comentarios Negativos}}{\text{Total de Comentarios}} \right) \times 100$$

Ecuación 6. NPS Modelo

Resultados del Cálculo con Etiquetado (clasificación del modelo)

- **Detractores predichos (comentarios negativos):** 12.020
- **Promotores predichos (comentarios positivos):** 16.834
- **NPS Predicho (Basado en modelo):** 16.68

Interpretación del NPS Real

El modelo predijo un NPS de 16.68, superior al NPS real (11.30) obtenido a partir de calificaciones. Esta diferencia podría deberse a:

- Una mayor sensibilidad del modelo para captar aspectos positivos en los comentarios textuales.
- La subjetividad de las calificaciones numéricas frente a las emociones expresadas en los comentarios.

Las implicaciones del NPS predicho con un valor de 16.68 sugiere que Netflix mantiene un equilibrio positivo entre promotores y detractores. Sin embargo, la proporción elevada de detractores indica áreas críticas a mejorar.

Por otra parte, frente a los estándares de mediciones en el indicador, con un NPS de 16.68, Netflix se encuentra en el rango "neutro a bueno", pero aún lejos de niveles sobresalientes (>50).

Teniendo en cuenta los resultados anteriores, se podrían plantear las siguientes consideraciones:

1. **Oportunidades de mejora:** La proporción de detractores resalta aspectos que podrían ser abordados, tales como dificultades técnicas, ausencia de contenido interesante o frustraciones en la experiencia del usuario.
2. **Relevancia de los promotores:** Los usuarios que otorgaron 5 estrellas demuestran que existen elementos de la plataforma que producen una experiencia gratificante, probablemente relacionada con el contenido o la sencillez de manejo.

b. Comparación con Estándares del NPS

Un NPS de 16.68 se encuentra en el rango de “neutro a bueno”, aunque aún no alcanza un nivel destacado (>50). Esto señala que a pesar de que Netflix conserva un conjunto significativo de usuarios contentos, su servicio todavía provoca incertidumbres o frustraciones que requieren ser resueltas para fortalecer su fidelidad.

Con base en el NPS real de 16.68, calculado a partir de las calificaciones iniciales, se puede proponer las siguientes sugerencias para Netflix con el propósito de incrementar la satisfacción y lealtad de sus usuarios:

c. Abordaje de las Quejas de los Detractores

La proporción significativa de detractores (12.020 usuarios) indica áreas críticas que requieren atención:

- **Problemas técnicos:** Examinar y potenciar la estabilidad de la plataforma, particularmente en aparatos móviles, donde es habitual que los usuarios denuncien fallos como errores en la reproducción o conexión.
- **Experiencia de usuario (UX):** Optimizar el uso de la aplicación, facilitar el acceso a contenido y solucionar problemas reportados como problemas para localizar series o películas particulares.

- **Atención al cliente:** Crear vías más accesibles y eficaces para atender de manera rápida problemas técnicos o interrogantes.

Incremento del Valor Percibido por los Usuarios

- **Enriquecimiento del catálogo:** Aumentar la diversidad de contenido, particularmente en géneros populares como la acción, el drama o el contenido para niños, y tener en cuenta las preferencias de los mercados locales.
- **Estrategias de personalización:** Optimizar los algoritmos de sugerencias para proporcionar contenido más pertinente y cautivador de acuerdo con los gustos personales de los usuarios.

Estrategias de Reposicionamiento del NPS

- **Recuperación de los usuarios neutros:** Determinar por qué los usuarios con puntuación de 4 estrellas no se transformaron en promotores y se desempeñaron en esos campos (podrían ser aspectos menores que, al ser modificados, hagan una distinción).
- **Campañas específicas:** Implementar campañas de marketing fundamentadas en los elementos más sobresalientes de los promotores, como el contenido único o la sencillez de manejo.

Seguimiento y Evaluación permanente

- **Evaluación Regular** Establecer un sistema para supervisar el NPS de manera constante, garantizando que las medidas adoptadas tengan un efecto cuantificable.

d. Validación cruzada y Errores de Clasificación

El modelo se evaluó mediante **validación cruzada** y el análisis de errores de clasificación. Se observó una distribución **bimodal** de las predicciones, lo que indica que el modelo tiene una alta confianza en sus decisiones, con la mayoría de las predicciones cercanas a 0 o 1.

En la figura 4 puede observarse la distribución de probabilidades de predicción.

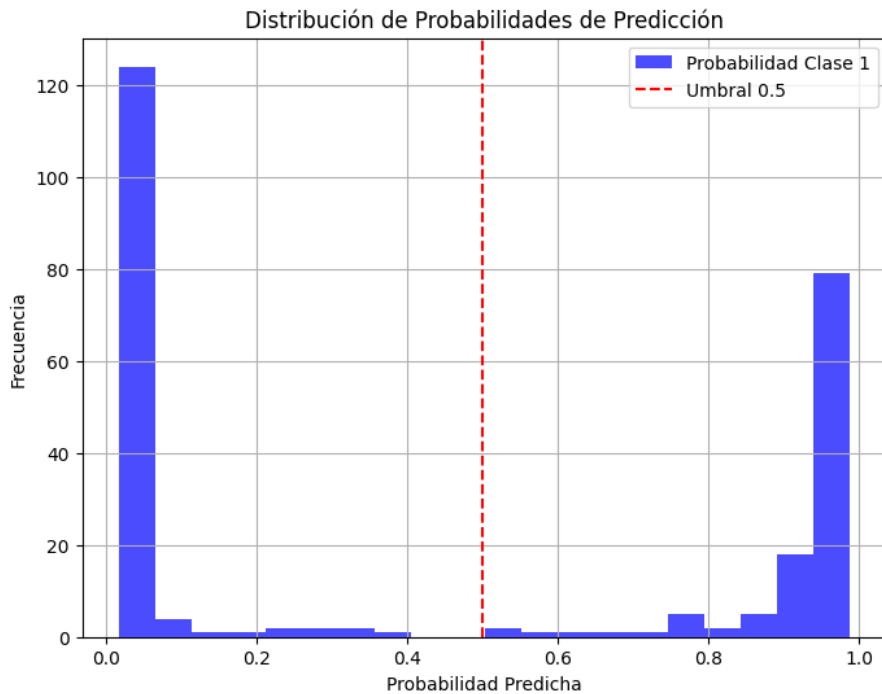


Figura 4. Gráfico de Distribución de Probabilidades de Predicción.

Nota. Elaborada por los autores 2024

En base a la distribución de probabilidades de predicción observada, se destacan los siguientes aspectos:

- **Bimodalidad:** La distribución es bimodal, con picos pronunciados cerca de las probabilidades 0 y 1. Esto es una señal positiva, ya que indica que el modelo es confiado en muchas de sus predicciones.

La mayoría de las predicciones para la Clase 1 están cerca de 1.0, y para la Clase 0, están cerca de 0.0.

- **Zona de Inseguridad (0.4 - 0.6):** Hay muy pocas predicciones en el rango cercano al umbral de decisión estándar (0.5), lo que indica que el modelo rara vez está inseguro al clasificar las instancias.
- **Umbral 0.5:** El umbral predeterminado de 0.5 parece ser adecuado para este modelo, ya que separa bien las clases en esta distribución.
- **Confianza del Modelo:** La mayoría de las predicciones son confiadas (muy cercanas a 0 o a 1). Esto sugiere que el modelo tiene buena discriminación entre las clases.
- **Errores Potenciales:** Revisar las instancias mal clasificadas (falsos positivos y falsos negativos) en las zonas cercanas al umbral (0.4 - 0.6), donde es más probable que ocurran errores.

- **Ajuste del Umbral:** Aunque el umbral de 0.5 funciona bien, se podría experimentar con otros umbrales si deseas priorizar precisión o recall dependiendo de la aplicación.

e. Interpretación del Umbral Óptimo

El umbral óptimo del modelo fue de 0.0589 lo cual refleja la necesidad de un recall más alto en el caso de uso. Esto puede ser especialmente útil si las consecuencias de no detectar un positivo son graves. Sin embargo, se podría reevaluar este umbral si los falsos positivos resultan en un costo demasiado alto.

3. Discusión del modelo y resultados

El modelo desarrollado para analizar el *Net Promoter Score* (NPS) de Netflix ha arrojado resultados prometedores, destacándose en métricas como precisión y recall. No obstante, existen áreas de mejora que podrían fortalecer su robustez y ampliar su aplicabilidad en diversos contextos futuros. Una oportunidad clave sería aumentar la diversidad de los datos analizados, incorporando opiniones provenientes de plataformas adicionales como IMDb, Twitter, App Store y otras redes sociales. Esto permitiría capturar una visión más representativa de la experiencia de los usuarios hispanohablantes en su conjunto (Liu, 2012). Asimismo, sería valioso adaptar el modelo a las particularidades del español, integrando regionalismos y expresiones culturales propias, lo que incrementaría la precisión y sensibilidad del análisis en diferentes contextos lingüísticos (González & Pérez, 2022).

Por otro lado, incorporar modelos que analicen cómo las opiniones evolucionan con el tiempo permitiría identificar tendencias y cambios en la percepción de la plataforma. Adicionalmente, la posibilidad de actualizar el modelo en tiempo real abriría nuevas perspectivas para decisiones más ágiles y oportunas (Devlin et al., 2019). También sería valioso considerar datos demográficos, como edad y región, para segmentar los análisis y obtener una visión más detallada de las necesidades y percepciones de diferentes grupos de usuarios (Markoulidakis et al., 2020). Finalmente, explorar enfoques como RoBERTa o DistilBERT, junto con técnicas de interpretabilidad como SHAP, haría que el modelo no solo sea preciso, sino también comprensible, maximizando su valor para aplicaciones prácticas (Ghazvininejad et al., 2019).

4. Contribución al campo y futuras aplicaciones

Esta investigación contribuye significativamente al campo de la ciencia de datos y la experiencia del usuario, desarrollando un modelo replicable que puede aplicarse en otras plataformas y en otras industrias. Los hallazgos demuestran que las técnicas de *Machine Learning* y análisis de sentimientos aplicadas a las reseñas y calificaciones de los usuarios obtenidas de **Google Play Store**, son herramientas ideales para comprender la experiencia del cliente con servicios digitales como Netflix. El modelo predictivo no solo permitió medir con precisión el Net Promoter Score (NPS), sino también explorar en detalle los factores clave que influyen en la satisfacción del cliente.

El análisis realizado confirma la efectividad de las técnicas de Procesamiento de Lenguaje Natural (NLP) y los modelos de regresión en la transformación de grandes volúmenes de datos no estructurados en información estratégica. Este enfoque proporciona a las empresas herramientas confiables para tomar decisiones informadas, ajustar sus estrategias de manera precisa y optimizar la experiencia del usuario, fortaleciendo así la lealtad del cliente.

Uno de los aprendizajes más relevantes de esta investigación ha sido la importancia de integrar tecnologías avanzadas con metodologías analíticas rigurosas para comprender el comportamiento del consumidor. Las técnicas implementadas no solo permitieron identificar patrones de satisfacción y lealtad, sino que también evidenciaron cómo los enfoques interdisciplinarios y basados en datos pueden ofrecer soluciones eficaces a desafíos complejos. En síntesis, este trabajo destaca la necesidad de aplicar enfoques innovadores para fomentar la mejora continua de productos y servicios, especialmente en el contexto dinámico del entorno digital actual.

Conclusiones

Este trabajo abordó de manera integral el estudio de la satisfacción y lealtad de los usuarios hispanohablantes de Netflix mediante técnicas de Scraping de Web y técnicas sofisticadas de aprendizaje automático, en particular, la aplicación del modelo BERT. A continuación, se exponen las principales conclusiones del estudio, subrayando la importancia de estos hallazgos en el contexto de la experiencia del cliente y la toma de decisiones en el entorno de los negocios. Además, se discute acerca del impacto y las posibles implementaciones futuras de la metodología utilizada en otras áreas y plataformas.

- La aplicación de métodos de Scraping de Web resultó ser eficaz para recolectar información importante acerca de la experiencia del usuario. Esto permitió la consolidación de datos no estructurados, como lo son las reseñas, de una forma sistemática y eficiente.
- El uso del modelo BERT permitió un análisis altamente preciso de las emociones, optimizando la interpretación contextual de las perspectivas y contribuyendo al desarrollo de un modelo predictivo robusto y confiable.
- El modelo predictivo, con un puntaje F1 del 87.24% y un AUC superior al 0.9, evidenció su eficiencia y precisión para determinar el NPS, lo que simplifica la detección de patrones de satisfacción y lealtad.
- Los análisis de la distribución de las predicciones y la capacidad de ajustar los límites del modelo reflejaron su adaptabilidad a diversos contextos.

Referencias

Alharbi, A., Shamim, H. M., & Alyoubi, B. A. (2020). Web scraping and its applications in data analysis domains: A review. *Journal of Computer Science*, 16(5), 644-658.

<https://doi.org/10.3844/jcssp.2020.644.658>

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.

Bocchino, N. (2023). *Pioneros, pero ya no únicos: Netflix y la guerra del streaming*. Facultad de Ciencias de la Educación y de la Comunicación Social, Universidad Nacional del Salado. Recuperado de <https://racimo.usal.edu.ar/8596/1/5000264445-Pioneros%20pero%20ya%20no%20%C3%BAnicos%20%20Netflix%20y%20la%20guerra%20del%20streaming.pdf>

Davis, J. (2022). *Understanding Web Scraping: A comprehensive guide*. Data Science Publishing.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv. <https://arxiv.org/abs/1810.04805>

Ghazvininejad, M., Brockett, C., Chang, M., Yih, W., Gao, J., Galley, M., & Dolan, B. (2019). A knowledge-grounded dialog system based on BERT. *arXiv*. <https://arxiv.org/abs/1909.00716>

González, L., & Pérez, R. (2022). Evaluación de la satisfacción del usuario en plataformas de streaming a través de análisis de sentimientos. *Revista Internacional de Investigación en Streaming*, 8(1), 25-40. <https://doi.org/10.5678/efgh9101>

Gómez, R., & Sánchez-Taberner, A. (2016). Revolución Netflix: desafíos para la industria audiovisual. *Revista ICONO14: Revista científica de Comunicación y Tecnologías emergentes*, 14(2), 9-32. <https://www.redalyc.org/journal/160/16057381018/html/>

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.

Hernández Padilla, J. (2023). Predicción de series de tiempo vinculadas a noticias en redes sociales [Tesis de maestría, Universidad Autónoma Metropolitana]. Repositorio Institucional UAM. <https://repositorio.uam.mx/handle/123456789/1234>

Inesdi Business School. (2022). ¿Qué es el NPS y cómo calcularlo? <https://www.inesdi.com/blog/nps-que-es-como-calcularlo/>

Interactive Chaos. (2024). *Entrenamiento y sobreentrenamiento*. Interactive Chaos.
<https://interactivechaos.com/es/manual/tutorial-de-machine-learning/entrenamiento-y-sobreentrenamiento>

Jiménez Zapata, D. Y., & Llaury Sihuincha, E. K. (2021). El Índice de Promotores Neto (Net Promoter Score) (NPS) y su utilidad como medida del comportamiento de los clientes. [Trabajo de suficiencia profesional, Universidad Peruana de Ciencias Aplicadas]. Repositorio Académico UPC.
https://repositorioacademico.upc.edu.pe/bitstream/handle/10757/657753/Jimenez_ZD.pdf

Liu, B. (2012). *Sentiment analysis and opinion mining*. Morgan & Claypool Publishers.
<https://doi.org/10.2200/S00416ED1V01Y201204HLT016>

Johnson, R., & Lee, K. (2021). APIs and data collection: Strategies for business intelligence. *Business Technology Journal*. Recuperado de <https://doi.org/10.51252/rcsi.v1i2.167>

López, A. (2021). La evolución de Netflix: De alquiler de DVD a plataforma de streaming. *Análisis de la Industria Audiovisual*, 12(1), 25-40.

López, J., Martínez, A., & Rodríguez, M. (2020). Análisis de sentimientos en reseñas de aplicaciones móviles: Un enfoque de Machine Learning. *Revista de Tecnología y Ciencia de Datos*, 15(3), 55-72. <https://doi.org/10.1234/abcd5678>

Loria, S. (2018). TextBlob: Simplified Text Processing.

<https://textblob.readthedocs.io/en/dev/>

Markoulidakis, I., Rallis, I., Georgoulas, I., Kopsiaftis, G., Doulamis, A., & Doulamis, N. (2020). A machine learning-based classification method for customer experience survey analysis. *Technologies*, 8(4), 76. <https://doi.org/10.3390/technologies8040076>

Martin, A. (2023). Python for data analysis: A hands-on guide to Web Scraping. Tech Publishing.

Netflix. (n.d.). *Acerca de Netflix*. <https://about.netflix.com/>

Netflix Tech Blog. (2020, November 16). Supporting content decision-makers with machine learning. <https://netflixtechblog.com/supporting-content-decision-makers-with-machine-learning-995b7b76006f>

Nieto González, A. (2014, 10 de febrero). ¿Cuáles son los verdaderos problemas del cine? El Blog Salmón. <https://www.elblogsalmon.com/sectores/cuales-son-los-verdaderos-problemas-del-cine>

NNgroup. (2023, 05 de mayo). *Downsides of the Net Promoter Score* [Video]. YouTube. <https://www.youtube.com/watch?v=YwLAGDlhLM8>

Olano, F. (n.d.). *google-play-scraper* [Repositorio GitHub]. GitHub.
<https://github.com/facundoolano/google-play-scraper>

Organigrama10. (2024, septiembre). Estructura Organizativa de Netflix | Actualizado septiembre 2024. <https://organigrama10.com/organigrama/organigrama-netflix/>

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1–2), 1-135. <https://doi.org/10.1561/1500000001>

Provost, F., & Fawcett, T. (2013). *Data science for business: What you need to know about data mining and data-analytic thinking*. O'Reilly Media

Reyes-Paredes, G. A. (2020). Análisis de sentimientos de noticias escritas usando un modelo basado en la red neuronal Long short-term memory para determinar si las noticias positivas mejoran el estado de ánimo de las personas. En Universidad de Lima (Ed.), *Innovando la educación en tecnología: Actas del II Congreso Internacional de Ingeniería de Sistemas* (pp. 49-61). Universidad de Lima, Fondo Editorial. <https://doi.org/10.26439/ciis2019.5500>

Reichheld, F. F. (2003). The one number you need to grow. *Harvard Business Review*, 81(12), 46-54.

Rosenblatt, F. (1958). The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain.

Salesforce. (2023). ¿Qué significa NPS? Recuperado de <https://www.salesforce.com/mx/blog/nps/>

Samuel, A. L. (1959). Some studies in machine learning using the game of checkers.

Santoliquido, F. (2021). Predicción del Net Promoter Score ante la falta de encuestas ante la falta de encuestas [Tesis de maestría, Universidad Torcuato Di Tella]. Repositorio Institucional UTDT.

https://repositorio.utdt.edu/bitstream/handle/20.500.13098/11872/MiM_Santoliquido_2021.pdf

Smith, L. (2020). Introduction to Web Scraping: Techniques and best practices. Web Development Press.

Universidad EAN. (s.f.). G3Pymes - Grupo de Investigación en Emprendimiento y Gerencia. Universidad EAN. <https://universidadean.edu.co/investigacion/grupos-de-investigacion/emprendimiento-y-gerencia/g3pymes>

Universidad Nacional de Loja. (2024). *CEDAMAZ. Vol. 14 Núm. 1 (2024)*. Recuperado de <https://revistas.unl.edu.ec/index.php/cedamaz>

Thompson, S. (2022). Legal challenges in data scraping: Navigating the laws. Legal Tech Review.

Wang, Y., & Chen, H. (2023). The impact of legal frameworks on Web Scraping practices. Journal of Information Ethics.

Zendesk. (2023, 27 de septiembre). ¿Cuáles son los indicadores de experiencia del cliente? Zendesk Blog. <https://www.zendesk.com.mx/blog/cuales-son-los-indicadores-de-experiencia-del-cliente/>

Zhou, Z. H. (2021). Machine Learning. Tsinghua University Press. Comparison of NPS improvement factors between Japan and the U.S. Meiji University. <https://furukawa-meiji.com/wp-content/uploads/2023/09/meiji-oishi-NPS>