



UNIVERSIDAD CATÓLICA
de Colombia
Vigilada Mineducación



MACHINE LEARNING MODEL FOR EFFECTIVENESS EVALUATION OF A DIGITAL MARKETING STRATEGY

MASTER IN INGEGNERIA E GESTIONE DELL'
INNOVAZIONE
FACULTY OF ENGINEERING
UNIVERSIDAD CATÓLICA DE COLOMBIA
UNIVERSITÀ DEGLI STUDI DI SALERNO

MARIA ALEJANDRA TRIVIÑO CAMELO
Matricola: MAT. 0670400033

Relatore:
Prof. Holman Diego Bolivar Barón Ph.D

(FIRMA ASESOR)

Direttore del Master:
Prof. Domenico Guida

ACADEMIC YEAR 2020/21



UNIVERSITÀ DEGLI STUDI
DI SALERNO



UNIVERSIDAD CATÓLICA
de Colombia

**MODELO DE MACHINE LEARNING PARA LA EVALUACIÓN DE EFECTIVIDAD
DE UNA ESTRATEGIA DE MARKETING DIGITAL**

MARIA ALEJANDRA TRIVIÑO CAMELO

**MAESTRIA EN INGENIERIA Y GESTIÓN DE LA INNOVACIÓN
FACULTAD DE INGENIERIA
UNIVERSIDAD CATÓLICA DE COLOMBIA & UNIVERSIDAD DE SALERNO
BOGOTÁ, May 21st, 2021**

**MODELO DE MACHINE LEARNING PARA LA EVALUACIÓN DE EFECTIVIDAD
DE UNA ESTRATEGIA DE MARKETING DIGITAL**

**MARIA ALEJANDRA TRIVIÑO CAMELO
MAT. 0670400033**

**Universidad Católica - Asesor: _____
Holman Diego Bolivar Barón Ph.D**

**U. Degli Studi di Salerno - Asesor: _____
Doménico Guida, Ph.D Eng.**

**MAESTRIA EN INGENIERIA Y GESTIÓN DE LA INNOVACIÓN
FACULTAD DE INGENIERIA
UNIVERSIDAD CATÓLICA DE COLOMBIA & UNIVERSIDAD DE SALERNO
BOGOTÁ, May 21st, 2021**



Atribución-NoComercial 4.0 Internacional (CC BY-NC 4.0)

This is a human-readable summary of (and not a substitute for) the [license](#). [Advertencia](#).

Usted es libre de:

Compartir — copiar y redistribuir el material en cualquier medio o formato

Adaptar — remezclar, transformar y construir a partir del material

La licenciente no puede revocar estas libertades en tanto usted siga los términos de la licencia

Bajo los siguientes términos:



Atribución — Usted debe dar [crédito de manera adecuada](#), brindar un enlace a la licencia, e [indicar si se han realizado cambios](#). Puede hacerlo en cualquier forma razonable, pero no de forma tal que sugiera que usted o su uso tienen el apoyo de la licenciente.



NoComercial — Usted no puede hacer uso del material con [propósitos comerciales](#).

No hay restricciones adicionales — No puede aplicar términos legales ni [medidas tecnológicas que restrinjan legalmente a otras a hacer cualquier uso permitido por la licencia](#).

<https://creativecommons.org/licenses/by-nc/4.0/deed.es>

NOTA DE ACEPTACIÓN

Jurado

Jurado

Holman Diego Bolivar Barón Ph.D
Asesor

Doménico Guida, Ph.D Eng.
Asesor

Bogotá, 2 de julio de 2021

TABLA DE CONTENIDO

RESUMÉN.....	11
ABSTRACT.....	12
1. INTRODUCCIÓN	13
2. PLANTEAMIENTO DEL PROBLEMA	15
3. OBJETIVOS	22
3.1. OBJETIVO GENERAL.....	22
3.2. OBJETIVOS ESPECIFICOS	22
4. MARCO CONCEPTUAL.....	23
4.1. ¿QUÉ ES MARKETING?	23
4.1.1. Etapas Del Marketing	24
4.2. MARKETING DIGITAL	26
4.3. REDES SOCIALES	28
4.4. SOCIAL MEDIA ANALYTICS	30
4.5. EFECTIVIDAD.....	31
4.6. OPERADOR DE INFORMACIÓN.....	32
5. MARCO TEORICO.....	34
5.1. MEDICIÓN DE LA EFECTIVIDAD.....	34
5.2. MACHINE LEARNING.....	37
5.2.1. Machine Learning y Marketing	40
6. ESTADO DEL ARTE	43
7. METODOLOGÍA.....	49
8. CARACTERIZACIÓN DE LOS PERFILES DE LOS CLIENTES	50
8.1. DESCRIPCIÓN DE LA CAMPAÑA	50
8.2. DESARROLLO DE LA CAMPAÑA	50
9. MODELOS DE MACHINE LEARNING (APRENDIZAJE SUPERVISADO)	54
9.1. ÁRBOL DE DECISIÓN	54
9.1.1. ¿Cómo funciona?.....	55
9.1.2. Matemática - Medidas de selección de atributos (AMS).....	56
9.2. RANDON FOREST	58

9.2.1. ¿Qué es el proceso de bagging y cómo funciona?	58
9.3. CLASIFICACIÓN DE NAÏVE BAYES	60
9.3.1. Ejemplo del algoritmo Naive Bayes.	60
9.3.2. Desventajas del algoritmo Naive Bayes.....	62
9.4. REGRESIÓN LOGÍSTICA	62
9.4.1. Matemática detrás de la Regresión Logística	62
9.4.2. Desventajas de la Regresión Logística.....	63
9.5. MÉTODOS “ENSEMBLE” (CONJUNTOS DE CLASIFICADORES)	64
9.5.1. Métodos de combinación más comunes en “ensemble”	64
9.5.2. Desventaja de los métodos “ensemble”.....	65
10. EFECTIVIDAD DEL MODELO A TRAVÉS DE SOCIAL MEDIA ANALYTICS	66
10.1. RESULTADO DATAFRAME PARA: CLASIFICACIÓN DE NAÏVE BAYES ..	67
10.2. RESULTADO DATAFRAME PARA: REGRESIÓN LOGÍSTICA.....	68
10.3. RESULTADO DATAFRAME PARA: MÉTODOS “ENSEMBLE”	69
10.4. RESULTADO DATAFRAME PARA: RAMDON FOREST (GINI)	70
10.5. RESULTADO LINKEDIN.....	72
11. RESULTADOS	74
12. CONCLUSIONES.....	75
13. FUTUROS TRABAJOS	76
14. REFERENCIAS BIBLIOGRAFICAS	77
ANEXOS	85

LISTA DE TABLAS

Tabla 1 Indicadores de acuerdo con los objetivos en LinkedIn [50].....	35
Tabla 2 Segmentación, Categoría: Sexo. Elaboración Propia.	52
Tabla 3 Segmentación, Categoría: Ciudad. Elaboración Propia.....	52
Tabla 5 Segmentación, Categoría: Máx. Nivel de Formación. Elaboración Propia.	53
Tabla 9 Segmentación, Categoría: Función Laboral. Elaboración Propia.....	54
Tabla 10 Dataframe bruto. Elaboración Propia.	66
Tabla 11 Dataframe caracterizado. Elaboración Propia.....	66
Tabla 12 Data Función Laboral. Elaboración Propia.....	71

LISTA DE FIGURAS

Figura 1 Redes sociales más usadas por organizaciones B2C [79]	16
Figura 2 Redes sociales más usadas por organizaciones B2B [79]	17
Figura 3 Beneficios de social media marketing [79].....	17
Figura 4 Planes de incrementar marketing orgánico en Facebook [79]	18
Figura 5 Planes de incrementar marketing orgánico en Instagram [79].....	18
Figura 6 Planes de incrementar marketing orgánico en LinkedIn [79]	19
Figura 7 Línea de tiempo de lanzamientos de las redes sociales más representativas. [12]	29
Figura 8 Mapa co-ocurrencias Machine learning and Social media tracking. [62]	48
Figura 9 Mapa co-ocurrencias - zoom Machine Learning. [62]	48
Figura 10 Diagrama - Metodología. Elaboración Propia.	49
Figura 11 Video Campaña LinkedIn Operador de Información [4].....	50
Figura 12 Total de seguidores obtenidos por campaña de Marketing Digital para el Operador de Información [49]	51
Figura 13 Estructura básica de árbol de decisión. Elaboración Propia a partir de [76].....	55
Figura 14 ¿Cómo funciona el algoritmo del árbol de decisión en Machine Learning?. Elaboración Propia a partir de [76].....	56
Figura 15 Ramdon Forest Básico. Elaboración Propia a partir de [56]	58
Figura 16 Ramdon Forest (Bagging - Arboles de Decisiones). Elaboración Propia a partir de [64].....	59
Figura 17 Ejemplo Modelo Naive Bayes – Probabilidades Reales. Elaboración Propia a partir de [71]	61
Figura 18 Ejemplo Modelo Naive Bayes – Gráfico formal. Elaboración Propia a partir de [71].....	61
Figura 19 Matemática detrás de la Regresión Logística. Elaboración Propia a partir de [57].....	62
Figura 20 Métodos de ensemble. Elaboración Propia a partir de [73]	64
Figura 21 Flujo data robot. Elaboración propia.	67
Figura 22 Resultado dataframe para: Clasificación de Naïve Bayes. Elaboración Data Robot.....	67
Figura 23 Resultado dataframe para: Regresión Logística. Elaboración Data Robot.	68
Figura 24 Resultado dataframe para: Métodos “Ensemble” (Conjuntos de clasificadores). Elaboración Data Robot.	69
Figura 25 Resultado dataframe para: Ramdon Forest (Gini). Elaboración Data Robot.	70
Figura 26 Gráfico de Elevación, Resultado Random Forest (Gini). Elaboración Data Robot.....	71

Figura 27 Resultado Administrador de Campañas - LinkedIn.....72

RESUMÉN

3 de cada 5 empresas que fracasan es porque no contaban con una estrategia de marketing, destacando que alrededor del 80% de las empresas que han cerrado en los últimos años no incluían en el modelo de negocio ningún presupuesto destinado a marketing. Además, existen compañías que implementan estrategias de marketing, pero no miden la eficiencia durante la campaña lo cual provoca que en caso de requerir un ajuste la reacción sea tardía.

Exitosas empresas invierten millones de dólares en estrategias de marketing con el fin de traer nuevos clientes; pero en la ejecución de estas estrategias se pueden presentar un resultado inesperado. A lo largo de la historia se han escuchado casos de grandes compañías como Burguer King, Pepsi y Coca Cola, entre muchas otras, las cuales han sufrido altas pérdidas por fracasos de campañas de marketing.

Se propone la implementación de un modelo de machine learning para evaluar la efectividad de una estrategia de marketing digital para un operador de información a través de social media analytics.

Se optó por machine learning porque es uno de los avances tecnológicos más significativos y de mayor uso a nivel de las organizaciones que tiene como objetivo generar sistemas más autónomos que permitan mejor análisis de datos y predicciones más acertadas y se integra con el análisis de redes sociales mediante el social media analytics.

Se eligió un modelo de machine learning robusto y efectivo que evaluó la estrategia de marketing expuesta mediante una red social, de esta forma se integró con los objetivos estratégicos del operador de información en la optimización de recursos por inversión en campañas que no dan ningún valor a la compañía.

Palabras Claves: Marketing digital, social media analytics, operador de información, machine learning, random forest.

ABSTRACT

3 out of 5 companies that fail is because they did not have a marketing strategy, highlighting that around 80% of the companies that have closed in recent years did not include any budget for marketing in the business model. In addition, there are companies that implement marketing strategies, but do not measure efficiency during the campaign, which means that if an adjustment is required, the reaction is delayed.

Successful companies invest millions of dollars in marketing strategies in order to bring in new clients; but in the execution of these strategies, an unexpected result may occur. Throughout history, cases have been heard of large companies such as Burger King, Pepsi and Coca Cola, among many others, which have suffered high losses due to marketing campaign failures.

The implementation of a machine learning model is proposed to evaluate the effectiveness of a digital marketing strategy for an information operator through social media analytics.

Machine learning was chosen because it is one of the most significant and widely used technological advances at the level of organizations that aims to generate more autonomous systems that allow better data analysis and more accurate predictions and is integrated with the analysis of social networks through social media analytics.

A robust and effective machine learning model was chosen that evaluated the marketing strategy exposed through a social network, in this way it was integrated with the strategic objectives of the information operator in the optimization of resources by investment in campaigns that do not give any value to the company.

Keywords: Digital marketing, social media analytics, information operator, machine learning, random forest.

1. INTRODUCCIÓN

Según Soto [77] el marketing digital es una herramienta básica para el crecimiento organizacional, especialmente cuando los consumidores se acostumbran cada vez más a satisfacer sus necesidades con un clic. En este contexto el marketing digital dispone de una serie de herramientas que permiten escuchar al cliente en tiempo real, conocer sus necesidades, hábitos, intereses, gustos y motivaciones, para así reaccionar más rápidamente o lograr predicciones más acertadas.

En este nuevo contexto digital se presentan nuevos retos, entre ellos, la gran cantidad de información que se encuentra en la web y la velocidad de propagación de esta. insideBIGDATA [36] revela cifras relacionadas con el nuevo contexto de las organizaciones, según este sitio web se estima que el 90% de la información generada en el mundo, se generó en los últimos 3 años y un ejemplo de esto son los casi 10 millones de fotos que se cargan en Facebook cada hora, los 24 petabits de información que una empresa como Google procesa cada día, o los 400 millones de tweets que se crean cada día. Para el 2020 se estimó que, alrededor de 1.7 megabytes de información será creada cada segundo por cada ser humano en el planeta, el universo de información acumulada en la web sería de 44 zettabits y se estima que un tercio de la información se transmitirá a través de la nube. Este volumen de información ha generado nuevas exigencias tecnológicas, al punto en que la información contenida en la web crece 4 veces más rápido que la economía mundial y la capacidad de procesamiento de las computadoras aumenta 9 veces más rápido.

Como se evidencia en las soluciones planteadas, la inteligencia artificial tiene infinidad de aplicaciones dependiendo de la necesidad y contexto de la empresa. Sin embargo, uno de los retos más desafiantes para la implementación de tecnologías de inteligencia artificial es la infraestructura que se requiere para la administración de la información, además, la capacitación del personal que está a cargo del entrenamiento de los modelos de machine learning y también del análisis de resultados de dichos procesos, pues es evidente que con el fin de obtener los mejores resultados se debe contar con expertos en el negocio de cada organización que cuenten con el conocimiento para ajustar los resultados y aprovechar la eficiencia que inteligencia artificial y machine learning brindan en el proceso de toma de decisiones.

El operador de información en donde se desarrolla este trabajo es una compañía que lleva más 12 años en el mercado de la Seguridad Social con la misión actual de “co-crear tecnologías de información integrales e innovadoras que aporten al bienestar y desarrollo de Colombia”. [3]

Es una compañía que provee una plataforma tecnología la cual ofrece servicios para el pago de la Planilla Integrada de Liquidación de Aportes (PILA) y Cesantías, esto

quiere decir que da el acceso a la liquidación y el pago de aportes a la Seguridad Social Integral de Salud, Pensión, Riesgos Laborales y parafiscales (Caja de Compensación Familiar, Sena e ICBF), con el fin de que la población colombiana tenga garantizado el servicio de salud y de la totalidad de los subsistemas de Seguridad Social integral como derecho fundamental establecido en el Artículo 48 de la Constitución Política de Colombia. **[15]**

Esta empresa lidera “el mercado de los Operadores de Información de PILA con un 43% de participación, y trabajan en la actualidad para convertirse en mediano plazo, en la plataforma de soluciones tecnológicas de información más consultada de Colombia”. **[3]**

Desde el área de marketing del Operador de Información se busca implementar y medir estrategias digitales encaminadas en la captación del mejor talento humano para así ofrecer una excelente experiencia para los clientes, integrando el machine learning con el social media analytics.

Siendo el machine learning uno de los avances tecnológicos más significativos y de mayor uso a nivel de las organizaciones que tiene como objetivo generar sistemas más autónomos que permitan mejor análisis de datos y predicciones más acertadas, velando por la integridad de los datos almacenados; todo esto se traduce en una mejora en la toma de decisiones.

Adicionalmente, se incorpora la estrategia de marketing digital la cual se maneja como una estrategia de marketing multicanal, en la cual cada uno de los canales en los que la organización establece contacto con el cliente deben potencializarse entre sí. Como es de esperarse, esta nueva forma de hacer marketing supone nuevos retos para la organización, Iruretagoyena **[37]** identifica los retos más relevantes al hacer uso del marketing digital, entre ellos, establecer mecanismos y metodologías que permitan cuantificar los resultados, segmentar de forma correcta los mercados y así definir los canales correctos para abordarlos, conseguir satisfacer consumidores más informados, saber aprovechar las oportunidades globales y por último establecer vínculos reales con los clientes.

Este proyecto de investigación se articula con el marco de proyectos de desarrollo tecnológico referenciados en Minciencias **[60]**, porque los resultados obtenidos originaran un modelo de machine learning el cual servirá para la evaluación de efectividad de una estrategia de marketing digital a través de social media analytics, en el mercado de los operadores de información Colombianos.

2. PLANTEAMIENTO DEL PROBLEMA

Marketing es el proceso en el cual se realizan actividades con el fin de entender las necesidades de los clientes y satisfacerlas mediante el ofrecimiento de productos o servicios. Esta definición de marketing implica que se deben desarrollar actividades de investigación de mercados, análisis de datos, desarrollo de productos o servicios, definición de precios y planes de comunicación. Para lograr esto, el marketing despliega estrategias para mantenerse en contacto constante con los clientes [43].

De la definición anterior se infiere que el campo de acción del marketing no solo es el interior de las organizaciones, por el contrario, gran parte de su desarrollo se da en el contexto que rodea la organización. Actualmente, dicho contexto es un ecosistema digital con gran variabilidad y altos volúmenes de información. Bajo estos parámetros, las interacciones entre los clientes y las organizaciones se han catapultado, adicionalmente, son los clientes quienes están dictando la naturaleza, extensión y contexto de las interacciones en el nuevo esquema de la web [28].

Este nuevo entorno digital supone nuevos retos para las empresas que procuran crear relaciones de valor con sus clientes [28]; sin embargo, la propaganda basada en web expande la habilidad del marketing para mover clientes desde el estado de alerta a estados de compromiso, consideración y fidelidad apoyándose en los diferentes puntos de interacción con sus usuarios y potencializando los esfuerzos del marketing en cada etapa [28].

La nueva dinámica en la que se generan interacciones ha creado nuevas formas en las que los consumidores buscan y valoran información, por ejemplo, durante el 2013, dos tercios de los usuarios de la web en EEUU se basaron en canales digitales para obtener información sobre viajes (eMarketer, 2013) y alrededor del 74% de los viajeros se basaron en comentarios y reseñas para planear sus viajes [18]. Si se observa el ranking mundial de los sitios web más visitados durante el 2020 se tiene que las principales redes sociales se sitúan dentro de los primeros 25 lugares de la siguiente manera: Facebook en tercer lugar del ranking, Instagram en el quinto y LinkedIn en el puesto 24 [75].

El uso de redes sociales ha sido una herramienta fundamental para dinamizar e incrementar las interacciones entre las empresas y los consumidores, sin embargo, a diferencia de la percepción común, este ambiente digital se ha venido forjando desde hace más de 20 años. De hecho, en 1997, la red social Sixdegrees permitía a sus usuarios crear un perfil, enlistar a sus amigos y agregar a amigos de sus amigos a sus listas [40]. Este ambiente de interconectividad y la posibilidad de que individuos ajenos a la organización creen, compartan y recomienden información, ha extendido el alcance del marketing. Por esta razón, varias plataformas ofrecen herramientas para generar interacciones cargadas de valor. Por consiguiente, esta evolución digital del marketing también ha afectado a los consumidores. Según

Karpinski (2005), ahora los consumidores son más inteligentes, organizados y más confiados de sus propias opiniones y las opiniones de sus pares [28].

Este comportamiento ha sido identificado por las organizaciones, haciendo que se adapten y cambien sus estrategias, lo que se evidencia en el reporte *Social Media Marketing Benchmark (2020)*, en el cual se entrevistaron a más de 5200 profesionales de marketing identificando que las 3 redes sociales más comunes para realizar social media marketing son Facebook, Instagram y LinkedIn [79]. En la siguiente gráfica se muestra la relevancia de las redes sociales para empresas B2C:

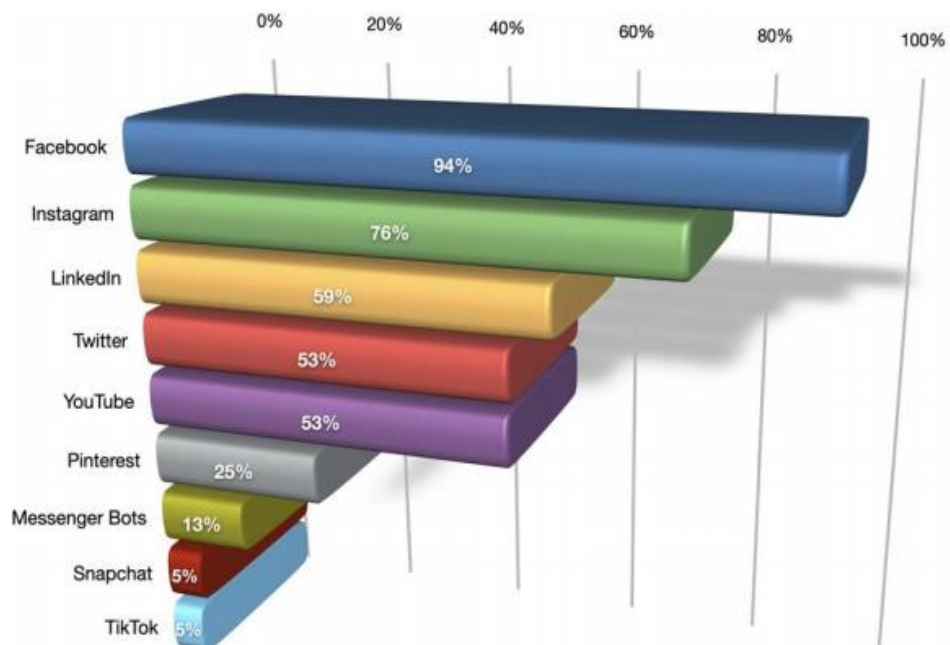


Figura 1 Redes sociales más usadas por organizaciones B2C [79]

De la misma manera, se evidencia que las mismas redes sociales se mantienen en las primeras posiciones para las empresas que desempeñan actividades B2B, sin embargo, para este tipo de organizaciones la plataforma LinkedIn cobra relevancia avanzando al segundo lugar con una importancia del 81%, como se evidencia en la siguiente gráfica:

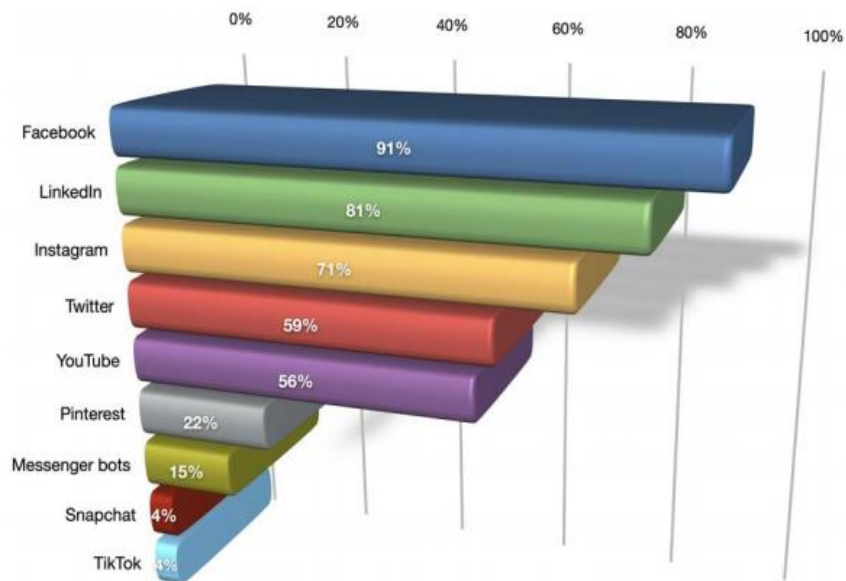


Figura 2 Redes sociales más usadas por organizaciones B2B [79]

Así mismo, cuándo se les preguntó los beneficios que perciben al hacer uso de las redes sociales, el 86% de los entrevistados coinciden con que el principal beneficio es la exposición de la marca, seguido por el incremento del tráfico en los sitios web con un 78% y la generación de leads con el 67% [79].

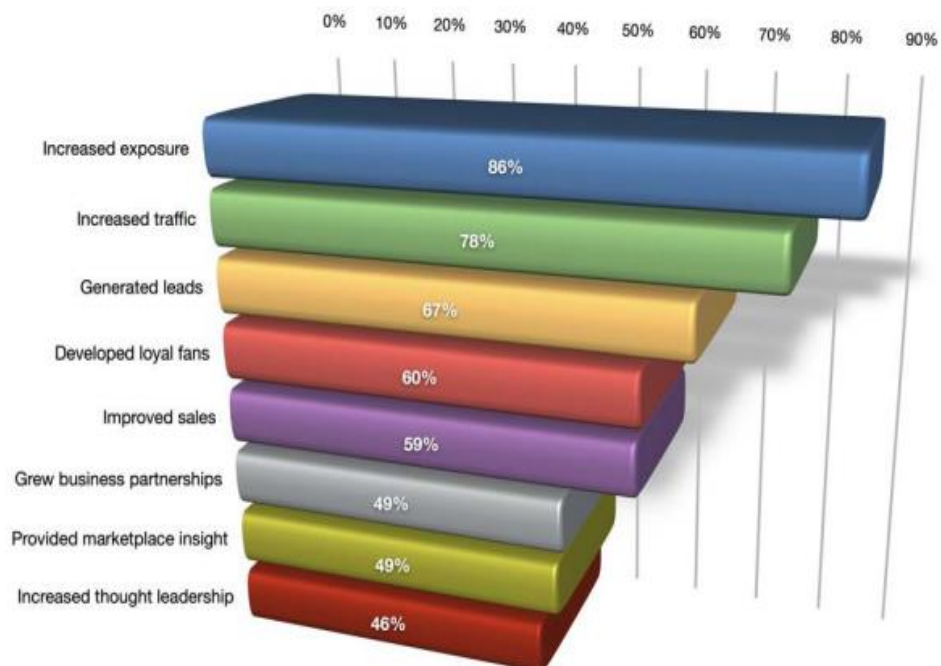


Figura 3 Beneficios de social media marketing [79]

En el Social Media Marketing Benchmarking [79], se identifica la intención de incrementar los planes de marketing a través de redes sociales, identificando que más de la mitad de los entrevistados planea incrementar los esfuerzos a través de las plataformas principales.

Para comenzar, el 50% de los encuestados planea incrementar sus esfuerzos a través de Facebook, 37% mantendrá los esfuerzos actuales, y tan solo un 13% no tiene planes de usar esta plataforma o tiene planes de disminuir sus actividades [79]:

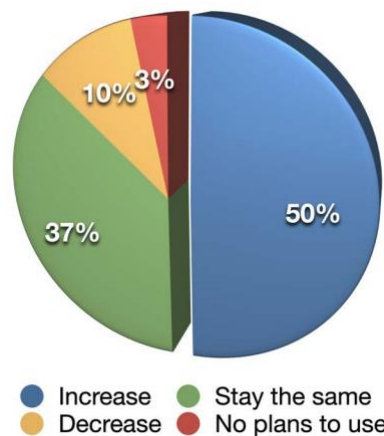


Figura 4 Planes de incrementar marketing orgánico en Facebook [79]

Por otro lado, cuando se evalúa la intención de usar Instagram, se observa que el 67% de los encuestados tiene planes de incrementar la presencia a través de este canal y solo un 14 % no tiene planes de usarla o planean decrecer su presencia. El 19% restante, planea mantenerse en el mismo nivel de uso [79]:

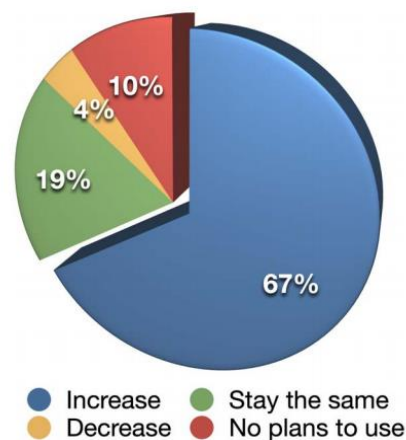


Figura 5 Planes de incrementar marketing orgánico en Instagram [79]

El 55% de los encuestados planean incrementar sus esfuerzos a través de la red social LinkedIn y para el caso particular de esta red, se tiene que el 22% de los encuestados no planea utilizar esta plataforma o planea decrecer sus esfuerzos [79]:

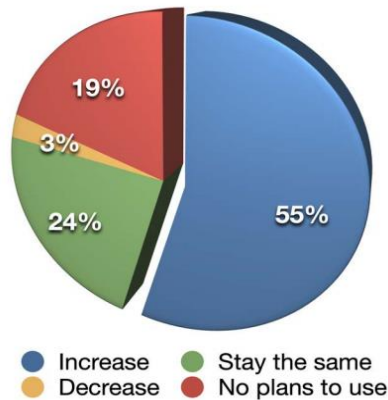


Figura 6 Planes de incrementar marketing orgánico en LinkedIn [79]

Por otra parte, la capacidad para medir la efectividad de las iniciativas de marketing en redes sociales es un enorme reto para las organizaciones, por ende, no existe una fórmula estándar que determine la mejor manera de disponer recursos en los diferentes canales disponibles para realizar social media marketing. Adicionalmente, muchas organizaciones no saben cómo medir el retorno de la inversión de las campañas dado que, las compañías que deciden establecer puntos de contacto a través de las redes sociales, deben lidiar con la dualidad de los canales propios (generados y controlados por la organización) y los canales ganados (no son generados ni controlados por la organización) [17].

Por otro lado, el concepto tradicional del retorno sobre la inversión (ROI), visto como la relación que existe entre el beneficio que se crea por invertir capital y el costo de la inversión [48], difícilmente es aplicable a una campaña de marketing digital, esto dado a que, el enfoque tradicional del retorno sobre la inversión subestima el valor de una publicación acerca de una marca o un producto. Por ende, se debe redefinir el concepto de ROI y adaptarlo a la nueva naturaleza del ecosistema digital y como complemento a una cifra estrictamente monetaria se deben tener en cuenta los comportamientos de los usuarios durante las interacciones mediante social media (likes, shares, repost, tags, menciones, comentarios) [29]. Por tanto, es evidente que la medición de la efectividad del marketing digital a través de social media se debe abordar desde dos frentes, uno cualitativo, con el que se mida la reacción de los usuarios a la presencia digital de una compañía (likes, comentarios, shares, efecto viral, followers) y otro cuantitativo, con el que se aborda el concepto tradicional de ROI. Es importante resaltar que, para poder medir una campaña de marketing digital desde un frente cuantitativo, la organización debe tener la habilidad

y la capacidad de rastrear sus ingresos y así poder determinar la cantidad que se derivan de dicha campaña [17].

Para lograr definir las características a tener en cuenta al momento de determinar la efectividad de una estrategia de marketing digital se debe conocer el método de facturación de cada una de las plataformas en las que se realizan campañas. Por instancia, en Facebook se cobra dependiendo del presupuesto que se asigne a cada una de las campañas en circulación y nunca se sobrepasa el límite establecido, sin embargo, esta plataforma se basa en un sistema de subasta para determinar qué anuncio se muestra o no y de esta forma el costo por reacción o costo por clic puede variar, lo que afecta el cálculo del ROI [23]. Cabe aclarar que, Facebook ofrece herramientas con las cuales se puede optimizar el costo de las reacciones o de los clics dependiendo del objetivo de la campaña, adicionalmente, se puede establecer un parámetro de retorno del gasto publicitario (ROAS), con lo cual se busca optimizar el rendimiento de los anuncios en esta plataforma [21].

En contraste, Instagram basa sus costos en el presupuesto y la duración de la promoción, estos parámetros se definen al momento de crear la campaña y afectan directamente el número de personas a las que se muestra dicha publicación, es decir, a mayor presupuesto se muestra la promoción a un mayor número de personas, esto repercute directamente en la medición de efectividad de la campaña [22].

Por otro lado, LinkedIn maneja un sistema de cobro basado en objetivos, el cual, en principio es similar al sistema de subasta utilizado en Facebook, sin embargo, las acciones facturables varían dependiendo del objetivo que se defina para la campaña, por ejemplo, una campaña cuyo objetivo es llevar visitantes al sitio web de la organización tiene como evento facturable los clics que redireccionan a los usuarios a dicha página web [51].

Como se hace evidente, la efectividad de las nuevas comunicaciones de marketing online es potenciada por el desarrollo continuo de publicidad en el ambiente online, las cuales pueden ser creadas usando diferentes herramientas analíticas. Por esta razón, realizar la medición de efectividad en las comunicaciones a través de un ambiente online es un proceso complejo, especialmente en un ambiente de constante cambio donde nuevos recursos de marketing son creados cada día. Sin embargo, se ha demostrado que las empresas con más alta efectividad en marketing también alcanzan mayor efectividad en el negocio en general comparado con sus competidores, para los cuales la efectividad del marketing no es una prioridad. Por esta razón, se podría decir que, el objetivo de evaluar la efectividad del marketing es analizar la relación entre las actividades de marketing y el desempeño general del negocio y por ende es muy importante que el concepto de efectividad del marketing contribuya al funcionamiento general del negocio y no sea una función aislada [46].

Con base en lo anterior, existen herramientas que ayudan a la identificación de oportunidades y medición de desempeño del marketing, actualmente, una de las más revolucionarias es Machine Learning (ML), como muestra de esto, para el 2018 cerca del 84% de las organizaciones de marketing estaban implementando o ampliando el uso de ML debido a que esto les permite realizar un mejor análisis y predicción del comportamiento de los consumidores, tanto así que según QuanticMind [8], el 97% de los ejecutivos opinan que el futuro del marketing depende de cómo los marketers utilizan el ML. De hecho, algunas de las aplicaciones de ML en marketing digital incluyen, predicción de comportamiento a través de análisis predictivo, creación de campañas de marketing predictivo basados en los comportamientos históricos de los usuarios, chatbots y marketing de contenido al personalizar las interacciones con los usuarios mediante canales digitales como las redes sociales [8]. Adicionalmente, ML es un diferenciador al momento de automatizar procesos de segmentación basados en grandes volúmenes de data que pueden ser recolectados desde las redes sociales, también, puede optimizar el análisis de los comentarios y reseñas que los usuarios expresan de forma online, de esta manera, ML cada vez tiene un rol más importante dentro de las diferentes etapas del marketing [58].

Dada la evolución del marketing en redes sociales y las características expuestas anteriormente, es evidente que una de las etapas más complejas en el desarrollo de campañas de marketing es la medición de la efectividad, ya que esta no se ha hecho de forma sistemática. Si bien una empresa conoce cifras generales como el número de seguidores o la cantidad de reacciones que se obtienen en una publicación, no se tiene un concepto definido de cómo medir la efectividad [11]. Sin embargo, con ayuda del machine learning se han desarrollado herramientas que contribuyen al alcance de los objetivos de cada campaña en específico, por ejemplo, Millares-Pechuán et al., [61]. proponen un modelo para optimizar el micro targeting durante el diseño de campañas y de esta manera, aumentar la relación entre el input requerido (tiempo de la publicidad, cantidad de personas a las que se les debe mostrar la publicación, etc.) y el resultado obtenido, con esto se lograría un impacto directo en la efectividad.

Por todo lo anterior la pregunta de investigación que sustenta este proyecto es:

¿Cuáles serían las características y componentes de un modelo de machine learning que permita evaluar la efectividad de una estrategia de marketing digital en una red social?

3. OBJETIVOS

3.1. OBJETIVO GENERAL

Implantar un modelo de machine learning para la evaluación de la efectividad de una estrategia de marketing digital a través de social media analytics.

3.2. OBJETIVOS ESPECIFICOS

- Caracterizar los perfiles de los nuevos seguidores para la aplicación de una estrategia de marketing para un operador de información.
- Identificar un modelo de machine learning para la evaluación de la efectividad de la estrategia de marketing definida.
- Validar la efectividad del modelo a través de social media analytics.

4. MARCO CONCEPTUAL

4.1. ¿QUÉ ES MARKETING?

Según la American Marketing Association [1] *“marketing son las actividades, conjunto de instrucciones y procesos de creación, comunicación, distribución e intercambio de ofertas que tienen valor para los consumidores, clientes, socios y la sociedad en general.”*

Para Kotler [45] todas las empresas exitosas tienen un foco común que es el cliente, en consecuencia, estas organizaciones están enfocadas en cómo satisfacer las necesidades de su mercado objetivo, el cual debe estar bien definido ya que por naturaleza es complejo. Es para este fin que las organizaciones hacen uso del marketing como una herramienta de primera necesidad. Así lo expresa Druker [20] al decir *“el marketing es tan fundamental, que no puede considerarse una función separada. Es todo el negocio visto desde la perspectiva de su resultado final, es decir, desde el punto de vista del cliente. Por lo tanto, en todas las áreas de la empresa debe permear el interés y la responsabilidad por el marketing.”*

En ese orden de ideas, Kotler [45] define marketing como la actividad por la cual se atraen clientes y se crean relaciones rentables con los mismos. Esta perspectiva es compartida por Stanton [78] cuando afirma que el marketing es, a groso modo, las actividades realizadas para propender el intercambio entre personas u organizaciones con el fin de satisfacer necesidades o deseos. Sin embargo, para que se pueda generar un intercambio desde el punto de vista del marketing, se deben cumplir 4 condiciones, deben intervenir dos o más personas u organizaciones las cuales tengan necesidades por satisfacer, la participación de las diferentes partes debe ser voluntaria, cada uno de los participantes debe tener algo de valor para ofrecer a sus contrapartes y finalmente las partes deben poder comunicarse y compartir información.

Por otro lado, Silva y otros [74] acotan el marketing como un proceso de planeación con el cual se busca anticipar las necesidades o deseos de un consumidor, y así asegurar que dicho consumidor se acerque a realizar un intercambio en lugar de acercarse a la competencia. Pero Silva no es el único en hacer referencia del marketing como proceso, Lambin y otros [47] exponen que el marketing es un proceso activo en el cual se desempeñan las tareas necesarias para el funcionamiento de una economía de mercado, y Armstrong y Kotler [5] escalan la visión del marketing como proceso a nivel social al afirmar que *“el marketing es un proceso social y directivo mediante el que los individuos y las organizaciones obtienen lo que necesitan y desean a través de la creación y el intercambio de valor con los demás. En un contexto de negocios más estrecho, el marketing implica la generación de relaciones de intercambio rentables y cargadas de valor con los clientes.”*. Como se observa en las diferentes definiciones, el marketing siempre

implica un intercambio entre dos o más partes para satisfacer necesidades o deseos, sin embargo, García [26] nos aclara que, a pesar que el enfoque en la satisfacción de las necesidades es relevante, en un entorno con una rápida adaptación tecnológica y el desarrollo de sistemas de gestión cada vez más robustos, una organización competidora puede llegar a ofrecer lo mismo que nosotros y al mismo mercado objetivo, por esta razón, el marketing debe ir mucho más allá y asegurarse de acoplar nuestra organización entre los clientes objetivo para que los mismos prefieran nuestra oferta de valor antes que cualquier otra opción a su alcance.

Para poder acotar un poco más el concepto de marketing, es importante aclarar el significado de sus componentes, en este caso, los componentes más representativos son valor y satisfacción, pues estos son componentes transversales en casi todas las definiciones de marketing. Baena [6] afirma que el valor es la relación entre los beneficios que se ganan al adquirir un bien y el sacrificio que implica obtener dicho bien, bajo esta perspectiva un consumidor adquirirá un producto cuando los beneficios que se ofrecen son mayores que los sacrificios o el costo de adquirirlo. A pesar de esto, no se puede asumir que cualquier oferta de valor va a asegurar la satisfacción de un consumidor, pues el concepto de satisfacción está basado en las expectativas del consumidor, en otras palabras, satisfacción es la sensación de que un producto cumple o excede las expectativas de un cliente.

4.1.1. Etapas Del Marketing. Es importante aclarar que, el enfoque del marketing no siempre fue el mismo. Toca [81] señala que el marketing ha pasado por 3 etapas diferentes para convertirse en la disciplina que es hoy en día. Inicialmente, todo estaba enfocado en producto, durante esta etapa la demanda de productos excedía la oferta y las organizaciones centraban su interés en hacer más eficientes los procesos de producción, ya que asumían que los consumidores adquirirían cualquier bien, siempre y cuando el precio fuera competitivo. Seguido a esto, el marketing evolucionó y se centró en las ventas, en este punto la competencia aumentó y se iniciaron los procesos de producción en masa, como resultado el objetivo de las organizaciones cambió y además de buscar producir con más eficiencia para lograr precios más competitivos, se centraron en vender lo que se producía sin importar la oferta del mercado y sus exigencias. Por último, el marketing se orientó al mercado y durante esta etapa las organizaciones se centran en la satisfacción de las necesidades de un mercado objetivo claramente definido. Parte de este cambio se da por que el consumidor es más educado y hay mucha competencia, como resultado, el marketing empieza a involucrarse con otros procesos de la compañía para asegurar la generación de una propuesta de valor atractiva, para este fin, el marketing se apoya en la información que se obtiene de los consumidores y así adaptarse para anticipar rápidamente las necesidades de estos.

Por otra parte, Kotler [42] describe una cuarta etapa en el marketing llamado marketing holístico, en esta perspectiva el marketing amplía el alcance y empieza a

importar la necesidad del mercado y todos los aspectos que se relacionan con la misma, Kotler define 4 componentes del marketing holístico:

- **Marketing relacional:** Hace referencia a todas las actividades encaminadas a crear relaciones duraderas y mutuamente beneficiosas con todos los actores importantes dentro del mercado objeto, y así lograr incrementar y mantener el negocio. El marketing relacional se centra en 4 grupos principales para crear estas relaciones, los clientes, los empleados, los socios de mercado y los actores financieros. Esto da como resultado lo que se denomina como una red de marketing, pues en el contexto empresarial actual, la competencia ya no se da entre organizaciones individuales, sino entre redes de marketing. Es en este contexto en el que una red de marketing sólida y definida por relaciones rentables entre sus partes se convierte en una ventaja competitiva.
- **Marketing integrado:** este componente del marketing holístico se fundamenta en las 4 P del marketing mix (producto, precio, plaza y promoción) y hace referencia a la toma de decisiones individuales que se apalanquen y potencialicen entre sí, es decir, el marketing integrado busca intervenir el proceso de toma de decisiones de tal forma que cada una de las P definidas en el marketing mix potencialice las demás. Para este fin el marketing integrado cuenta con 2 etapas. Primero, existen diversas actividades de marketing y comunicación que se pueden utilizar para generar valor. Y la segunda etapa aclara que, todas las actividades se deben pensar con el fin de maximizar los resultados de forma conjunta.
- **Marketing interno:** El principal objetivo de este componente, es garantizar que todos los funcionarios de la organización se apropien de los principios de marketing adecuados, pues se considera imposible ofrecer una propuesta de valor sin contar con la capacidad interna para generarla. Para esto, el marketing interno se desarrolla en 2 niveles, inicialmente busca que todas las actividades y decisiones de marketing estén integradas a lo largo de la compañía, y el segundo nivel está enfocado en conseguir que todas las actividades de marketing se generen desde el punto de vista del cliente. Estos dos niveles de acción garantizan que todos los funcionarios conozcan claramente el impacto de su labor en el mercado que se desea satisfacer.
- **Marketing social:** El enfoque de este componente del marketing holístico, define que las organizaciones deben satisfacer las necesidades de su público objetivo, de manera eficiente y asegurando que se mantiene o incrementa el bienestar de los consumidores y la sociedad en general, con este fin en mente se puede decir que, el marketing social son aquellas actividades con las que una organización se compromete con una causa en particular, para beneficio mutuo.

4.2. MARKETING DIGITAL

Actualmente, con la rápida adaptación de las tecnologías de información y el creciente número de usuarios interconectados, los retos y las exigencias del mercado han cambiado, por ende, el marketing también ha tenido que adaptarse. En este contexto nace el marketing digital.

Chaffey y Ellis-Chadwick [13] definen marketing digital como el cumplimiento de los objetivos de marketing haciendo uso de herramientas digitales. Sin embargo, para darle un poco más de enfoque Fischer y espejo [24] definen marketing digital como todas las estrategias de marketing que se desarrollan online.

La evolución del marketing durante esta era digital supone, según Iruretagoyena [37] nuevas características a tener en cuenta para el marketing digital como lo son, la facilidad de segmentación, la necesidad de procesamiento, análisis y almacenamiento de información, la ampliación geográfica del mercado y mayor agresividad por parte de la competencia. Actualmente, se está inmersos en estas nuevas características, Fischer [24] aclara 4 objetivos en los que se enfoca el marketing digital, inicialmente busca una comunicación directa con clientes cuidadosamente segmentados y así mismo pretende obtener una respuesta inmediata. El segundo objetivo del marketing digital es, establecer interacciones con los consumidores con el propósito de crear relaciones a largo plazo. Como tercer objetivo el marketing digital busca, apalancarse en las bases de datos para fortalecer la lealtad de los consumidores y lograr que los clientes potenciales se conviertan en clientes fieles mediante el envío de información ajustada a cada necesidad y así también reforzar la presencia de la marca. Por último, busca satisfacer las necesidades de clientes con nuevos hábitos de compra para usuarios que usualmente buscan ahorrar tiempo durante el proceso de intercambio y mayor conveniencia al tener mayor acceso a diferentes productos.

A pesar de que los objetivos del marketing continúan siendo los mismos, esta nueva etapa del marketing supone algunas diferencias con respecto a los canales tradicionales, Iruretagoyena [37] identifica 3 grandes diferencias entre el marketing tradicional y el marketing digital. Inicialmente, con el marketing digital se crean canales de comunicación de doble vía, es decir, no solo son las empresas las que disponen información para que un usuario la reciba, sino también los usuarios generan información para que las organizaciones la utilicen. Otra gran diferencia es que, con la rápida adopción de la World Wide Web ya no son las organizaciones las únicas que se encuentran en búsqueda constante de nuevos clientes, los clientes también se encuentran en búsqueda de nuevas formas de satisfacer sus necesidades y deseos. Por último, el uso de tecnologías de información permite realizar una medición más acertada del impacto que tiene una organización en su mercado objetivo y cómo este reacciona a la presencia de la organización en la web.

Es en este entorno donde la presencia de una empresa en diferentes canales debe planificarse cuidadosamente, en su libro marketing digital Chaffe y Ellis-Chadwick [13] aclaran que es necesaria una estrategia de marketing digital a fin de asegurar que esta se integre con todas las actividades de marketing incluidas en el plan organizacional, para así apoyar la consecución de los objetivos definidos por la dirección, lo que quiere decir que, la estrategia de marketing digital se concibe como una estrategia de canal, en esta se deben establecer objetivos específicos que estén alineados con la estrategia de marketing corporativo y además se debe establecer el peso de cada uno de los canales con el fin de generar valor en cada uno de los puntos de contacto con los clientes. En este sentido, la estrategia de marketing digital se maneja como una estrategia de marketing multicanal, en la cual cada uno de los canales en los que la organización establece contacto con el cliente deben potencializarse entre sí. Como es de esperarse, esta nueva forma de hacer marketing supone nuevos retos para la organización, Iruretagoyena [37] identifica los retos más relevantes al hacer uso del marketing digital, entre ellos, establecer mecanismos y metodologías que permitan cuantificar los resultados, segmentar de forma correcta los mercados y así definir los canales correctos para abordarlos, conseguir satisfacer consumidores más informados, saber aprovechar las oportunidades globales y por último establecer vínculos reales con los clientes.

Estas condiciones particulares del marketing online han creado un ambiente propicio para el crecimiento del marketing de social media, una práctica bastante atractiva para el desarrollo de estrategias de marketing digital. Estas nuevas plataformas suponen nuevas ventajas para las organizaciones, entre ellas, se pueden crear contenidos dirigidos y personalizados, las redes sociales son ambientes interactivos, es decir, se crean canales donde se recibe el feedback de los consumidores en tiempo real, también son canales de contacto inmediato y oportuno, además, en su mayoría son costo-eficientes. En contraste, los desafíos que se enfrentan en las estrategias de social media marketing son la dificultad para medir la eficiencia de la estrategia y la facilidad con la que se puede perder o tergiversar el mensaje [45].

Sin importar el enfoque desde el que se aborde el concepto de marketing, uno de los pilares fundamentales para el éxito de cualquier organización es la propuesta de valor que se ofrece a las diferentes partes. Kotler y Lane [44] afirman que la propuesta de valor hace referencia a la totalidad de beneficios que una organización puede ofrecer, en otras palabras, la propuesta de valor es aquello que los clientes pueden esperar recibir de la oferta disponible en el mercado. Con esto en mente, una oferta de valor no es suficiente por sí sola, en orden de cubrir las expectativas de los consumidores y lograr que prefieran nuestra oferta de valor, las organizaciones deben contar con un sistema de entrega de valor que cubra todos los puntos de contacto con el consumidor durante la búsqueda y adquisición del producto.

4.3. REDES SOCIALES

Para poder determinar cómo se debe desarrollar o medir una estrategia de social media marketing es necesario entender el concepto de redes sociales. Las redes sociales son plataformas virtuales que permiten a los usuarios establecer relaciones entre sí, bien sea que se conozcan o no fuera del ecosistema virtual. [65]

Una de las características principales de las redes sociales es la posibilidad de generar un intercambio de información entre sus usuarios, bien sea a través de transacciones económicas, comentarios, publicaciones, reacciones, calificaciones, links, archivos, entre otros. Basados en la descripción anterior, se entiende que una red social se puede generar en diferentes tipos de aplicativo, por ejemplo, foros, juegos en línea, blogs, perfiles o hasta buscadores de empleo. [65]

Una definición de red social un poco más acotada al contexto actual la establece Boyd y Ellison [12], quienes la definen como un sitio basado en la web en el que individuos pueden crear perfiles públicos o semipúblicos, crear listas de contactos y puedes consultar las listas de contactos hechas por otros individuos en el sistema.

Bajo esta última definición tenemos que, algunas de las características comunes de las redes sociales son, la necesidad de diligenciar información básica al momento de crear un perfil, la posibilidad de reconocer y conectar otros usuarios que tengan un perfil en la misma red social y la capacidad de enviar mensajes entre usuarios [12]. Dadas estas características, se tiene que el propósito de una red social es mantener una referencia de las personas con las cuales se tiene una relación o crear vínculos con otras personas.

Ahora, es necesario conocer el recorrido que han tenido estas plataformas, el cual inicia en 1997 con el lanzamiento de la red social SixDegrees, que es la primera red social reconocida en la cual usuarios podían crear perfiles, enlistar a sus amigos y enviar mensajes entre sí, si bien, estas funcionalidades ya existían en otras plataformas, SixDegrees fue la primera en unificarlas. Posterior al lanzamiento de SixDegrees, se resalta la aparición de LiveJournal en 1999, un sitio en el que se las personas se marcaban como amigos con el fin de poder seguir sus diarios, más adelante, en el año 2000 aparece la red social LunarStorm, seguido de Ryze en el 2001 y Friendster en el 2002. En el 2003 nacen varios sitios de redes sociales basados en el éxito que tuvo Friendster que en su mayoría se centraban en la creación de perfiles dentro de su plataforma, de este crecimiento se destaca la creación de LinkedIn centrada en perfiles y servicios profesionales, Care2 para unir activistas con causas afines y lastFM que se centraba en música. Ese mismo año, MySpace fue lanzado y su popularidad se dio debido a la decaída de Friendster, lo que causó un crecimiento dado que las personas que dejaban de usar Friendster se pasaban a MySpace. Pero no fue sino hasta 2004 que los servicios de redes sociales se convirtieron en un fenómeno mundial con la llegada de Facebook, el

cual, inicialmente ofrecía servicios para conectar a personas de Harvard y posteriormente a personas del ámbito escolar y profesional, pero en 2005 Facebook se abrió para aceptar a cualquier persona. Posterior al éxito de Facebook se ha dado un crecimiento vertiginoso de las redes sociales, donde se destaca la aparición de Youtube en 2005, Twitter en 2006 e Instagram en 2010 [12].

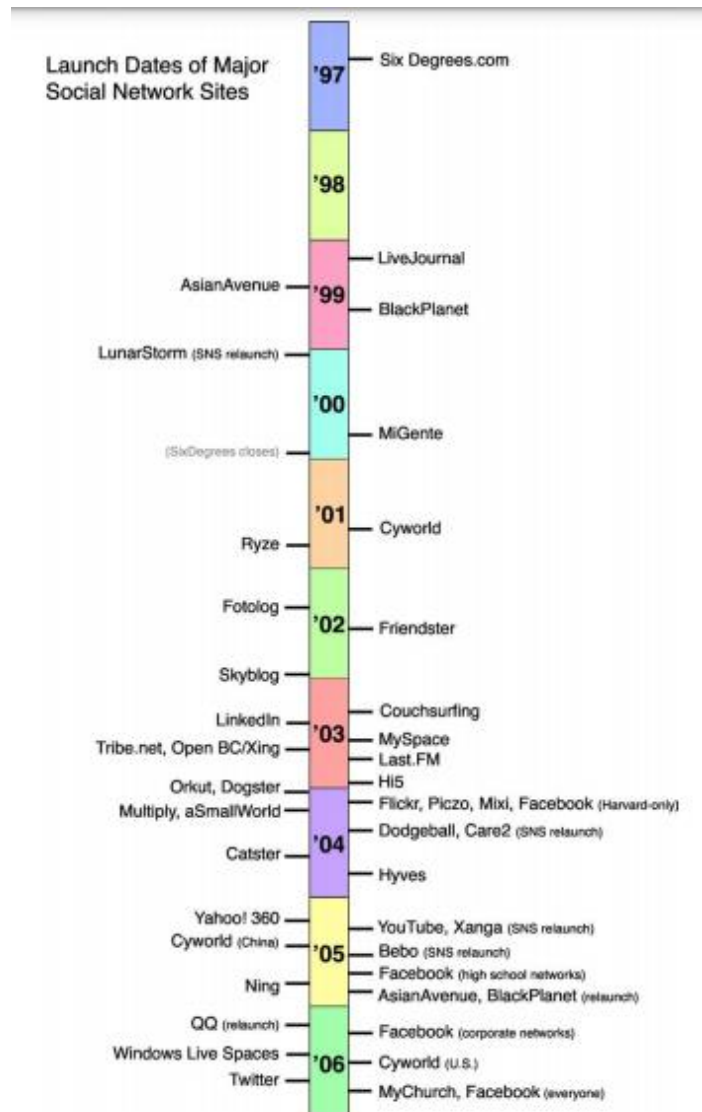


Figura 7 Línea de tiempo de lanzamientos de las redes sociales más representativas. [12]

4.4. SOCIAL MEDIA ANALYTICS

Con el fin de tomar decisiones acertadas en cuanto a las estrategias digitales y así poder medir su impacto y la efectividad, se han desarrollado bastantes herramientas, entre ellas analytics o analítica.

En la aplicación de una estrategia de social media marketing, analytics es una herramienta fundamental, Dotras [19] resalta la importancia de esta herramienta ya que mediante su uso se puede hacer seguimiento a la estrategia gracias a la recolección de datos digitales, luego permite transformar esos datos en información y por ende proporciona una brújula al momento de tomar decisiones referentes a la presencia de una organización en las redes sociales.

A diferencia de lo que se puede pensar, analytics no se enfoca en el contenido de los sitios web, en cambio, se centra en el rastro que dejan los visitantes en las páginas y así poder responder a las preguntas:

- ¿Cuántas personas se atraen?
- ¿Fuentes de tráfico, que hace referencia a cómo llegan los usuarios a la información?
- ¿Cuáles son las palabras clave que usan los usuarios para encontrar a la compañía?
- ¿Qué hacen los usuarios una vez ingresan a los sitios web?
- ¿Qué dificultades se presentan durante la navegación?
- ¿Con qué frecuencia se visita un contenido específico?
- ¿Cuánto es el tiempo de permanencia en los contenidos generados?
- ¿Cuáles son los intereses de los consumidores?
- ¿Qué dispositivos los clientes utilizan durante la navegación?

Asegurar la correcta medición de las estrategias de social media marketing genera varios beneficios, entre ellos, la oportunidad de segmentar los consumidores con mayor precisión, conocer el perfil de los usuarios que interactúan con el contenido, los puntos de contacto según el perfil del consumidor e inclusive el nivel de compromiso con la marca. Adicionalmente, con herramientas como el análisis de comentarios, las organizaciones pueden medir la imagen que los diferentes usuarios tienen acerca de la marca de la empresa; a nivel interno, analytics permite crear contenidos más relevantes para el consumidor, realizar recomendaciones más personalizadas, identificar oportunidades de mejora en los puntos de contacto con el cliente y apalancar otros canales de comunicación mediante el ofrecimiento de contenido dedicado según la segmentación de los usuarios [83].

Para asegurar una correcta medición de la estrategia de social media marketing, se debe crear un plan de medición concreto. Human Level Communications [31] afirma que, todo plan de medición debe pasar por 4 etapas que son, la fijación de un

objetivo, implementar una estrategia de medición, análisis de datos y ajuste de la estrategia. También se aclara que, al momento de crear un plan de medición no se debe confundir una métrica con uno de los indicadores de la estrategia de marketing, ya que la métrica es un dato del comportamiento de un usuario y el indicador es el que refleja el resultado de las actividades desarrolladas dentro de una estrategia con referencia a los objetivos generales de la organización. Con esto claro, se definen 4 objetivos comunes que debe manejar el plan de medición sin importar la plataforma que se mida, y estos son:

- Alcance: Hace referencia al número de usuarios que ven el contenido publicado.
- Engagement: Se mide la cantidad de interacciones de un usuario con una publicación (reacciones, menciones, comentarios, ...)
- Captación: Se puede interpretar en cantidad de ventas o en cuántos usuarios llegan al sitio web de la compañía a través de redes sociales.
- Fidelización: Este indicador refleja el valor de vida de un cliente teniendo en cuenta la información recolectada de la actividad web.

Para finalizar, Zuccherino [83] resalta algunas de las herramientas de ayuda para facilitar el proceso de analytics:

- Pirendo: Analítica de Twitter y Facebook
- Facebook Analytics: Analítica de Facebook
- YouTube Analytics: Analítica de YouTube
- Peer Index, Kred y Kloud: Analítica de reputación online.
- Google Analytics: Analítica de sitios web y apps.

4.5. EFECTIVIDAD

Uno de los componentes más importantes para las organizaciones es la efectividad de sus actividades, pero este concepto que representa un eje fundamental en el desarrollo de una empresa no ha sido fácil de definir. En algunos casos, se han listado hasta 30 criterios que se relacionan con el concepto de efectividad, sin embargo, se plantea que dichos criterios se pueden agrupar en 3 modelos fundamentales, el modelo racional, en el que se evalúa el número de unidades producidas en cierto periodo de tiempo; el modelo natural, donde se tienen en cuenta las actividades que se requieren para que el sistema se mantenga por sí solo y el modelo abierto, que tiene en cuenta actividades de creación y mantenimiento del sistema tales como adquisición de recursos y adaptabilidad. [67]

De los modelos anteriores el más aceptado es el racional y bajo esta perspectiva se puede decir que efectividad es la habilidad de producir un output con un consumo racional de recursos o inputs. Por ende, los sistemas que se desarrollan en las organizaciones tienden a minimizar recursos o maximizar salidas con el fin de lograr la efectividad deseada. Si se aplica este concepto en marketing, se tiene que la

efectividad es la relación entre los resultados de una campaña de marketing comparados con el esfuerzo y actividades que requiere alcanzar dicho resultado. **[46]**

Adicionalmente, el concepto de efectividad en marketing debe contribuir al funcionamiento general de la organización, y no debe ser una actividad aislada. Por esto, el objetivo de medir la efectividad de las actividades de marketing es analizar la relación que existe con el desempeño general del negocio, teniendo en cuenta el impacto de estas actividades en la participación del mercado, ventas, margen de ganancia e inclusive el éxito en lanzamientos de nuevos productos. **[46]**

4.6. OPERADOR DE INFORMACIÓN

El decreto 1465 de 2005 **[66]** define a los operadores de información como las entidades que realizan las siguientes acciones:

- Suministrar acceso a la Planilla Integrada de Liquidación de Aportes (PILA) de forma electrónica.
- Permitir el registro de la información detallada para el pago, su modificación antes de realizar el proceso de pago y en los casos en que haya lugar, la modificación de la información posterior al proceso de pago.
- Aplicar las validaciones establecidas por ley y presentar el informe de dicha validación antes del envío de la información, con el fin que los errores sean ajustados.
- Generar archivos de salida, reportes e informes que requieren los demás actores del sistema integrado de seguridad social.
- Almacenar el registro de identificación del aportante y la información de la planilla por un periodo mínimo de 3 meses.
- Mantener la disponibilidad y conexión con las plataformas de pago e instituciones financieras para permitir al aportante realizar el débito desde su cuenta, y a las administradoras recibir el pago correspondiente.
- Suministrar la información para realizar la distribución de pagos de forma oportuna.
- Desarrollar procesos de conciliación y contingencia durante el intercambio de información.
- Cumplir con el estándar ISO 17799 para garantizar la seguridad de la información durante los procesos recepción y envío de la misma.
- Servir de intermediario entre el aportante y las administradoras para realizar los créditos y débitos a que haya lugar.
- Socializar la información de las transacciones a los aportantes, administradoras y autoridades pertinentes.
- Aplicar las reglas de seguridad y validación vigentes en el sector financiero.
- Asegurar los procesos de conciliación y contingencia propios para las transacciones financieras. **[66]**

Teniendo en cuenta las funciones que desempeña un operador de información, el presidente de la república, mediante la ley 019 de 2012 **[16]**, en el artículo 73 establece que la vigilancia e inspección de dichas actividades son competencia de la Superintendencia Financiera de Colombia. Adicionalmente, en los artículos siguientes establece que es este organismo quien dará previa autorización a los individuos que deseen prestar servicios en calidad de operador de información con el fin de iniciar sus operaciones, también, se le confieren facultades para sancionar cualquier incumplimiento y hasta suspender las actividades de un operador de información en el caso en que haya lugar.

En resumen, la Unidad de Pensiones y Parafiscales (UGPP) **[82]** define a los operadores de información como las organizaciones que suministran a los aportantes del sistema de seguridad social, acceso a la Planilla Integrada de Liquidación de Aportes (PILA) y a su vez permiten el ingreso, modificación validación y corrección de los detalles de pago. Basados en dichas funciones, el Operador de Información **[2]** define que el objetivo de un operador de información es garantizar el correcto funcionamiento del mecanismo definido para realizar los pagos al sistema de seguridad social.

5. MARCO TEORICO

5.1. MEDICIÓN DE LA EFECTIVIDAD

Se establece que la medición de la efectividad debe proporcionar un parámetro claro que aporte visión al logro de objetivos organizacionales, por esta razón, se debe tener en cuenta la naturaleza bidireccional de cualquier interacción entre una organización y un usuario, ya que la mayoría de los esfuerzos se enfocan en un ambiente digital que es altamente interactivo. Para esto, se debe asegurar que la medición de la efectividad refleje la creación de valor para los actores del sistema en términos de salidas y procesos, y también la satisfacción de los participantes en la interacción. [80]

Para poder reflejar la creación de valor a través de estrategias de marketing digital, es necesario establecer objetivos claros para cada estrategia y de esta manera definir los indicadores de cumplimiento de dichos objetivos.

Según LinkedIn [50] los objetivos usados con mayor frecuencia en una campaña de marketing son:

- Dar a conocer la marca en su fase inicial.
- Mejorar la interacción y la percepción de la marca
- Generar leads y conversiones

Posterior a la elección de los objetivos se determinan los indicadores en función del objetivo de la campaña, en donde LinkedIn [50] se centra en los siguientes indicadores de acuerdo con los objetivos:

Objetivos	Indicadores
Dar a conocer la marca (fase inicial del proceso de marketing) Mejorar la percepción de la marca y la interacción (fase intermedia del proceso de marketing)	<i>Clics</i> : número de clics en los enlaces del anuncio.
	<i>Impresiones</i> : número de veces que la gente ha visto el anuncio.
	<i>Porcentaje de clics (CTR)</i> : número de clics dividido entre las impresiones.
	<i>Tasa de interacción media</i> : interacción total (clics de pago y gratuitos) dividida entre las impresiones. Las interacciones pueden incluir acciones sociales como recomendaciones, comentarios o veces compartido.

Generar leads y conversiones (fase final del proceso de marketing)	<i>Conversiones</i> : número de veces que alguien ha llevado a cabo una acción después de haber visto o hecho clic en el anuncio. Este indicador se aplica a campañas que usan el seguimiento de conversiones de LinkedIn.
	<i>Tasa de conversión</i> : frecuencia con la que los anuncios dan lugar a conversiones en tu sitio web.
	<i>Costo por conversión</i> : gasto del anuncio dividido entre el número de conversiones.
	<i>Leads</i> : número de leads que se han obtenido gracias a los anuncios. (número de clics sobre el número de personas que dejaron los datos)
	<i>Costo por contacto (CPL)</i> : gasto del anuncio dividido entre el número de contactos.

Tabla 1 Indicadores de acuerdo con los objetivos en LinkedIn [50]

LinkedIn [50] nos muestra algunas fórmulas de indicadores en función de los objetivos de la campaña:

Porcentaje de clics (CTR): el número de clics dividido entre las impresiones.

$$CTR = \frac{C}{I}$$

En donde:

- CTR: es el porcentaje de clics
- C: es el número de clics
- I: son las impresiones

Tasa de interacción media (TIM): interacción total (clics de pago y gratuitos) dividida entre las impresiones.

$$TIM = \frac{IT}{I}$$

En donde:

- TIM: es la tasa de interacción media
- IT: es la interacción total (clics de pago y gratuitos)
- I: son las impresiones

Costo por conversión (CV): gasto del anuncio dividido entre el número de conversiones.

$$CV = \frac{GA}{Co}$$

En donde:

- CV: es el costo por conversión
- GA: es el gasto del anuncio
- Co: son las conversiones

Leads: número de clics sobre el número de personas que empezaron a seguir al Operador de Información.

$$L = \frac{C}{S}$$

En donde:

- L: es el número de leads obtenido gracias a la publicación
- C: es el número de clics
- S: el número de personas que empezaron a seguir al Operador de Información.

Costo por contacto (CPL): gasto del anuncio dividido entre el número de contactos.

$$CPL = \frac{GA}{PL}$$

En donde:

- CPL: es el costo por contacto
- GA: es el gasto del anuncio
- Co: es el número de contactos.

Según Kietzmann et al., [40], una vez una compañía ha definido la métrica apropiada en función del objetivo de la campaña, se debe seleccionar la herramienta de evaluación apropiada. Esta se puede basar en información objetiva por ejemplo el número de visitas o seguidores, o puede usar la inteligencia colectiva de las masas, por ejemplo, sistemas de votación. Esto hace que las compañías sean capaces de monitorear cuantas veces son mencionados usando mediciones como el índice de menciones positivas contra las negativas.

Para reflejar comprensivamente la naturaleza de las interacciones de los clientes en relación con las marcas particulares, se ha propuesto el concepto de 'Compromiso de los clientes con las marcas' (CBE). CBE representa una métrica clave para el dimensionamiento del desempeño de una marca. Según los autores [30] se conceptualiza CBE como: Una influencia positiva en un cliente y su relación con una marca en niveles cognitivos, emocionales y conductuales durante o relacionados con interacciones cliente/marca. [30]

5.2. MACHINE LEARNING

El origen del machine learning generalmente es asociado con el psicólogo Frank Rosenblatt, quien se basó en ideas sobre el trabajo del sistema nervioso del ser humano y construyó junto a un equipo de trabajo una máquina para reconocer las letras del alfabeto Rosenblatt (1957, 1959, 1960). “La máquina, llamado el "perceptrón" por su creador, utiliza tanto analogía y señales discretas e incluía un elemento de umbral que convierte señales analógicas en discretas.” “Se convirtió en el prototipo de redes neuronales artificiales modernas (ANN), y el modelo de su aprendizaje estaba cerca de los modelos de animales y el aprendizaje humano desarrollado en psicología”. [25]

Según el BBVA [9], machine learning es una rama de la inteligencia artificial que permite que las máquinas aprendan sin ser programadas para ello, generando así sistemas más autónomos que permitan mejor análisis de datos y predicciones más acertadas.

Según IBM [33] dependiendo de la necesidad que quiera atender existen 4 enfoques que se pueden abarcar en el machine learning:

- Aprendizaje supervisado: este enfoque es útil cuando se tiene un conjunto establecido de datos con sus etiquetas correspondientes, con este enfoque se pueden usar imágenes con su respectiva descripción para enseñar a un ordenador a diferenciar entre millones. [33]

Recuero de los Santos, [68] afirma que los modelos o algoritmos más habituales que aplican para el aprendizaje supervisado son:

- Árboles de decisión.
 - Random Forest.
 - Clasificación de Naïve Bayes.
 - Regresión Logística.
 - Métodos “Ensemble” (Conjuntos de clasificadores).
-
- Aprendizaje no Supervisado: para cantidades de datos más grandes y sin etiquetar se usa el enfoque de aprendizaje no supervisado, este enfoque utiliza algoritmos de clasificación de datos con base en los patrones de estos y se realiza sin ningún tipo de intervención humana, es bastante útil para el análisis de datos relacionados con redes sociales y una de sus aplicaciones más comunes es la identificación de correo no deseado. [33]

Recuero de los Santos, [68] afirma que los algoritmos más habituales que aplican para el aprendizaje no supervisado son:

- Algoritmos de clustering.
 - Análisis de componentes principales.
 - Descomposición en valores singulares.
 - Análisis de componentes principales.
- Aprendizaje de refuerzo: otro de los enfoques de machine learning es el aprendizaje de refuerzo, este enfoque es conductual y se basa en el aprendizaje obtenida a través de la prueba y error, para el desarrollo de este enfoque se requiere que el usuario genere retroalimentación del análisis que realiza y como resultado el sistema refuerza el comportamiento que genere resultados más satisfactorios; a diferencia del aprendizaje supervisado, el aprendizaje de refuerzo no requiere un entrenamiento previo con datos ejemplo. **[33]**
 - Deep learning: es un método mediante el cual se emula el comportamiento del cerebro humano mediante la construcción de redes neuronales en capas sucesivas, este método se enfoca en el aprendizaje a través de iteraciones sucesivas y especialmente útil para definir patrones de datos no estructurados, mediante el deep learning una máquina puede llegar a lidiar con abstracciones y problemas mal definidos, entre las aplicaciones de este método se encuentran los sistemas de reconocimiento de imágenes, voz y los sistemas de visión por computadora. **[33]**

Hurwitz y Kirsch **[32]** explican que, para obtener los mejores resultados en el uso de machine learning, se debe garantizar que se sigue el ciclo de machine learning el cual incluye las siguientes fases:

- Identificación de los datos: En esta fase se deben identificar las fuentes de datos con un enfoque expansivo, esto con el fin de asegurar que la sostenibilidad del sistema y la mejora del análisis. **[32]**
- Preparación de los datos: Para asegurar el correcto funcionamiento del sistema, se debe asegurar que la información está calibrada con el algoritmo que se pretende usar, esto quiere decir, transformar la información para que el algoritmo pueda procesarla, si el algoritmo únicamente procesa datos numéricos no podemos tener información en palabras o imágenes, se debe asegurar que dicha información es convertida en números para poder alimentar el sistema. **[32]**
- Seleccionar el algoritmo a usar: En este paso se deben evaluar las diferentes opciones disponibles con el fin de seleccionar el algoritmo que mejor se adapte a la necesidad de la organización. **[32]**

- Entrenamiento: Se debe entrenar el algoritmo seleccionado mediante alguno de los enfoques antes visto, es decir, mediante aprendizaje supervisado, no supervisado o de refuerzo. [32]
- Evaluación: En este paso se revisan los resultados del sistema para definir las correcciones que hagan falta con el fin de mejorar los resultados del análisis. [32]
- Despliegue: Los resultados del proceso de machine learning son modelos que pueden ser aplicados a las organizaciones, en este paso, se debe integrar el resultado del sistema a la práctica. [32]
- Predecir: Después de integrar el modelo en la práctica, se deben empezar a hacer predicciones basadas en la nueva información generada. [32]
- Evaluación de los resultados: Se deben evaluar los resultados del proceso y las predicciones, la información que se genere de este análisis es la retroalimentación del sistema y debe ser una nueva entrada para mejorar la eficiencia del proceso de machine learning. [32]

Hurwitz y Kirsch [32] también explican que las organizaciones han venido utilizando la analítica de datos para entender de dónde vienen, su contexto actual y así mismo lograr predecir cómo las decisiones que se toman pueden impactar a futuro. Sin embargo, la analítica de los negocios estaba limitada a realizar predicciones basados en datos históricos, esto generaba un nivel muy bajo de adaptación a factores desconocidos que pueden impactar los resultados futuros. Por esta razón, las organizaciones han cambiado el enfoque de la analítica para centrarse en 2 tipos. La analítica descriptiva, es aquella que permite el entendimiento de las condiciones actuales del negocio y permite resolver preguntas como qué productos se están vendiendo más, que regiones tienen el mayor o menor crecimiento. La analítica predictiva, este enfoque del análisis organizacional se encarga del estudio de la información relacionada con el mercado, en orden de anticiparse a cambios y predecir los resultados dependiendo del curso de acción.

insideBIGDATA [34] aclara que la información es crucial para el éxito de un sistema de machine learning, pues la precisión de un modelo depende de la cantidad y la calidad de información con que se alimente. Bajo esta premisa es donde el concepto de big data se vuelve relevante, Barranco [7] define Big Data como la tendencia informática para referirse a enormes cantidades de datos y por ende la manipulación de forma convencional se hace muy costosa. En complemento Hurwitz y Kirsch [32] definen Big Data como una clase de fuente de información que cumple con al menos 1 de 4 características, volúmenes extremadamente grandes de información, la habilidad de mover información a una velocidad muy rápida, una variedad de fuentes de datos en constante expansión y veracidad. Una característica adicional

es que, Big Data agrupa toda clase de datos, estructurados, no estructurados y semiestructurados, provenientes de fuentes muy diversas.

Normalmente, según Barranco [7], una base de datos de Big Data contiene información proveniente de las siguientes fuentes:

- Social media: Contiene información proveniente de las redes sociales como clickstreams, contenido web, posts, etc. [7]
- Machine to Machine (M2M): Hace referencia a la información generada por dispositivos que se conectan entre sí, esta información proviene normalmente de medidores o sensores conectados a diferentes tipos de redes que transforman dicha medición en información relevante. [7]
- Datos de transacciones masivas: Esta fuente contiene información transaccional como datos de facturación o registros de llamadas. [7]
- Biométrica: Es toda la información relacionada con huellas digitales, reconocimiento facial, escaneos de retina, genética, entre otras. [7]
- Generada por los humanos: Esta información es generada por nosotros e incluye notas de voz, documentos electrónicos, estudios médicos y demás. [7]

En conclusión, Big Data puede representar una ventaja significativa si se combina con la tecnología adecuada. Hurwitz y Kirsch [32] enfatizan que los sistemas de machine learning, combinados con tecnologías para el manejo de Big Data confiere a la organización una herramienta para la toma de decisiones y la anticipación a cambios en el mercado, pues Big Data dispone de información actualizada constantemente y proveniente de diferentes fuentes confiables, lo que significa una mayor entrada de información para los sistemas de machine learning quienes los cuales generan predicciones más ajustadas al contexto real de la organización.

5.2.1. Machine Learning y Marketing. Según Miklosik et al., [59] el Machine Learning incluye mecanismos adaptativos que permiten que una computadora o máquina aprenda, basándose en la experiencia y los ejemplos, aclarando que el aprendizaje de nuevas habilidades aumentará el rendimiento del sistema con el tiempo.

En el sector del marketing, ML tiene un enorme potencial de ser utilizado en proceso de toma de decisiones, interacción con clientes y planeación estratégica. [59]

La aplicación de herramientas dirigidas por machine learning según Miklosik et al., [59] en el marketing digital, introduce nuevos y numerosos retos y oportunidades. Entre los beneficios más grandes de usar estas herramientas en marketing están:

- Desempeño óptimo: Las máquinas se desempeñan a un 100% porque ellas no pueden ser interrumpidas o distraídas. [59]
- Toma de decisiones más rápida: El tiempo que toma una máquina para tomar una decisión depende de la información disponible. Después de un cálculo rápido las máquinas pueden tomar una decisión casi de inmediato. La decisión tomada no sería afectada por factores subjetivos como sentimientos, preferencias personales, opiniones, etc. [59]
- Automatización de actividades predecibles: Machine learning puede automatizar rutinas de una forma muy efectiva. Por ejemplo, en marketing digital puede ayudar con la generación de reportes regulares de las campañas de marketing y en las de social media. [59]
- Reduce el margen de error: Elimina los errores causados normalmente por los humanos. Una máquina siempre seguirá las tareas definidas en el proceso. [59]
- Asistente digital: asistentes personales humanos son muy comunes, los sistemas manejados por machine learning pueden manejar tareas complejas y optimizar la rutina diaria. [59]
- Exploración de áreas que el ser humano es incapaz: El ser humano es incapaz de desarrollar tareas en ciertas áreas por diferentes razones. Por ejemplo, los humanos no pueden sumergirse en las áreas más profundas del océano o procesar las grandes cantidades de información que se generan cada minuto en internet. Por otro lado, las máquinas se pueden adaptar casi a cualquier condición, y pueden procesar hasta la operación estadística más compleja. [59]

Sin embargo; Miklosik et al., [59] realizan un estudio cualitativo, en el cual se pretendió identificar las barreras y riesgos de aplicar herramientas de machine learning e inteligencia artificial en marketing, así como anticipar el futuro rol de las herramientas basadas en machine learning en esta industria. El costo de implementación (62%), el tiempo requerido para aprender cómo desarrollar y usar las herramientas (47%) y las limitaciones del mercado como tamaño y escasez de oportunidades para usar ML e IA (45%) fueron las 3 barreras más comunes. Los directores de marketing tienen miedo que al usar las herramientas analíticas basadas en ML se pueda aumentar el error debido a falta de know-how (50%); tener un impacto negativo en los resultados y relaciones debido a las dimensiones éticas y morales en el comportamiento de un programa y su inteligencia (41%); o que signifique pérdida del control en el proceso de marketing (35%). Indiferente a esto, 72% de los encuestados creen que ML e IA representan el futuro del marketing digital (9% respondieron que piensan que estas herramientas se van a dejar de usar y el 19% no estaba seguro). [59]

A pesar de las cifras mostradas anteriormente en un estudio cualitativo realizado por Miklosik et al., Bayoude et al., [8] indican que 84% de las organizaciones de marketing están implementando o expandiendo el uso de machine learning.

Para Bayoude et al., [8] machine learning es una rama de estudio que brinda a computadores la habilidad de aprender y mejorarse a ellos mismos mediante el aprendizaje basado en la información de experiencias sin basarse en una programación explícita, basándose en la idea que sistemas de computadoras pueden identificar modelos y tomar decisiones con mínima intervención humana.

Según Bayoude et al., [8] utilizando modelos de inclinación y analítica predictiva, mercaderistas pueden identificar clientes potenciales, quienes son más probables de responder a sus ofertas dadas unas características correlacionadas con su comportamiento anterior.

El análisis predictivo según Bayoude et al., [8] permite a los mercaderista extraer información de la data que los usuarios deja cuando interactúan online, y la usa para predecir tendencias de compra y el modelo de comportamiento de los usuarios, en orden de desarrollar sistemas automáticos y perfiles de clientes para atacar ciertos mercados. Esta herramienta analiza grandes volúmenes de datos y ayuda a revelar los aspectos con más impacto. Esto se trabaja a través de los siguientes pasos:

- Identificar los objetivos que los datos analizados extraídos de diferentes fuentes y que requieren limpieza y preparación para determinar el patrón o modelo que mejor se ajusta nuestras necesidades.
- Creación del modelo y su validación, los modelos son validados basados en sus objetivos.
- Aplicación de los resultados del modelo para tomar ventaja de estos y guiarse para tomar decisiones de negocio correctas, también refinar el modelo continuamente para obtener mejores resultados

Como conclusión para Bayoude et al., [8] el análisis predictivo es el análisis de la información actual e histórica disponible y usando técnicas de machine learning para crear predicciones de comportamientos, preferencias y necesidades, se apunta a mayor crecimiento estratégico empresarial mediante el marketing digital.

6. ESTADO DEL ARTE

Haciendo uso del Machine Learning empresas como Spiro logran generar productos como Proactive Relationship Management, el cual es un software creado con base en inteligencia artificial y que ofrece ventajas para el personal de ventas. En primer lugar, apoyándose en inteligencia artificial posibilita una mejor toma de decisiones al sugerir el mejor curso de acción mediante la revisión constante y en tiempo real de la información disponible. Adicionalmente, está planeado como uno de los mejores asistentes personales para los representantes de ventas, ya que el software posibilita la creación de listas de tarea automáticamente, recomendaciones de acciones a tomar teniendo en cuenta la información disponible. Finalmente, Proactive Relationship Management correlaciona y analiza la información en su red para generar presentaciones a las partes interesadas y generación de informes de forma automática. [35]

Según Miklosik et al., [59] actualmente el potencial de las herramientas analíticas orientadas por machine learning se basan en la automatización de procesos periódicos que consumen el tiempo de manera poco efectiva cuando son desarrollados por personas, así mismo, aseguran que el uso más grande del machine learning es proveer mayor calidad en los datos y su uso recurrente como reportería, creación y optimización de campañas, e incluso comunicación con clientes.

En el campo de la analítica, machine learning es un método para el análisis de datos que mejora y automatiza la construcción de modelos analíticos. [8] Además de ser usado para desarrollar complejos algoritmos predictivos y modelos matemáticos más precisos que analicen mayores cantidades de información tanto actual como histórica con el fin de hacer predicciones de eventos futuros.

El Machine learning resuelve la pregunta de cómo construir computadores que mejoran automáticamente a través de la experiencia. Además, con la inteligencia artificial (IA), machine learning ha emergido como el método seleccionado para el desarrollo de programas prácticos en visión de computadoras, reconocimiento de diálogo, procesamiento de lenguaje natural, control de robots y otras aplicaciones. [53]

Jordan, and Mitchell [53] indican que conceptualmente, los algoritmos de machine learning se pueden ver como una búsqueda entre un gran espacio de programas candidatos, guiada por el entrenamiento y la experiencia, en orden de encontrar un programa que optimice la métrica de resultado. Además, que los dispositivos móviles y computación integrada permiten recolectar grandes cantidades de datos sobre individuos, y los algoritmos de machine learning pueden aprender sobre esos

datos para personalizar sus servicios según las necesidades y circunstancias de cada individuo.

Los algoritmos de aprendizaje automático de Machine Learning se pueden dividir en tres grandes categorías: aprendizaje supervisado, aprendizaje no supervisado y aprendizaje de refuerzo. En los algoritmos de aprendizaje supervisado se genera un modelo predictivo, el cual se basa en datos de entrada y salida. En donde la palabra clave “supervisado” viene de la idea de tener un conjunto de datos previamente etiquetado y clasificado, en donde el algoritmo va aprendiendo a clasificar las entradas y las compara con la etiqueta real de la muestra. **[52]**

En donde en el aprendizaje supervisado se puede encontrar algoritmos como: Árboles de decisión, Random Forest, Clasificación de Naïve Bayes, Regresión Logística y Métodos “Ensemble” (Conjuntos de clasificadores). **[68]** Estos algoritmos son útiles en los casos en que una propiedad (etiqueta) está disponible para un determinado conjunto de datos (conjunto de formación), pero debe predecirse para otras instancias. **[10]**

Por otra parte, el aprendizaje no supervisado trabaja con datos que no han sido etiquetados o clasificados y no se tiene una etiqueta que predecir. Estos algoritmos se usan principalmente en tareas donde es necesario analizar los datos para extraer nuevo conocimiento o agrupar entidades por afinidad. **[72]**

En el aprendizaje no supervisado se puede encontrar: Algoritmos de clustering, Análisis de componentes principales, descomposición en valores singulares, análisis de componentes principales. **[68]** Los algoritmos En el aprendizaje no supervisado son útiles en los casos en que la tarea consiste en descubrir relaciones implícitas en un conjunto de datos no etiquetados o no segmentados previamente. **[10]**

Y el aprendizaje de refuerzo se encuentra en alguno de estos dos extremos, en donde hay alguna forma de realimentación disponible en cada momento de predicción, pero no hay segmentación precisa o algún mensaje de error. **[10]**

El uso de las tecnologías puede incrementar el valor del marketing y entregar más experiencias a los clientes del mañana. Además, que el marketing es manejado en red a través de páginas basadas en social media, donde el compromiso de los clientes es clave para generar rentabilidad sostenida; para lo cual el machine learning y la inteligencia artificial son usados con frecuencia en el campo del marketing específicamente en la colocación de anuncios interactivos y la orientación de anuncios al contenido de video, con el fin de construir publicidad interactiva. **[58]**

La implementación de machine learning aplicada a las herramientas analíticas del marketing tiene desventajas obvias según Miklosik et al., **[59]**. Sin embargo, dependiendo del tipo de herramienta, existen diferentes preocupaciones o limitantes para su uso eficiente:

- **Creatividad:** El proceso de toma de decisiones en marketing está basado ampliamente en la creatividad. Es el resultado de mezclar varios parámetros que son naturales del ser humano, con emociones, intuición y una parte significativa de empatía. Estos 3 componentes son muy difíciles de replicar en una máquina. **[59]**
- **Principios éticos y morales:** Es la habilidad de tomar decisiones, expresar emociones, hacer sentido de los valores éticos y morales que hace a las personas humanas. Es cuestionable si las máquinas nunca tendrán esta habilidad, o si será posible desarrollarla en etapas posteriores. Hay preocupaciones justificadas sobre si las máquinas seguirán operando para el interés de la gente y dentro de límites éticos y morales cuando el progreso continúe. **[59]**
- **Preferencias del consumidor:** Los usuarios no siempre quieren interactuar con un robot o un programa. En lugar, se espera que, en el futuro, la necesidad de interacción con un humano va a permanecer como una preferencia, ya que los clientes mantendrán la confianza en ese tipo de conversaciones. **[59]**
- **Las herramientas de machine learning no pueden funcionar sin la intervención humana:** En las décadas venideras, las mentes humanas aún serán necesarias para trabajar con sistemas de machine learning para desarrollar y optimizar esta tecnología. **[59]**
- **Los algoritmos pueden estar equivocados:** Las decisiones hechas basadas en las estadísticas y cálculos matemáticos hechos por machine learning pueden resultar en acciones incorrectas, bien sea por el tipo de instrucciones o datos. Cualquier sistema de información avanzado está concebido primariamente para servirle a la gente. Dado que la gente requiere cambiar y evolucionar con el tiempo, puede resultar en que las herramientas basadas en machine learning no provean resultados óptimos con la configuración actual. Respuestas automáticas en solicitudes mediante redes sociales o descuentos generados automáticamente en caso de insatisfacción durante el proceso de compra sirven como ejemplo. **[59]**
- **Machine learning y empleo:** Hay sistemas que prácticamente pueden reemplazar la fuerza laboral en marketing digital. Por ejemplo, la carga automática de publicaciones en redes sociales, soporte de lectura de texto automática, etc. Sin embargo, es importante notar que estos sistemas nunca se comportarán correctamente sin apoyo humano. **[59]**

A pesar de las desventajas mencionadas anteriormente, la innovación tecnológica permite a las organizaciones rastrear los puntos de contacto en línea que tienen los clientes, lo que hace posible la predicción del comportamiento del cliente en un futuro. Además, indican que las nuevas tecnologías empoderan a las organizaciones a recolectar más información válida de su entorno, en donde lo importante es guiar a las personas a interactuar con esta información y así es como la tecnología ha cambiado todo, y eso permite nuevos caminos para crear experiencias de cliente, nuevos canales para contactarse con los clientes y otros participantes, también se requiere trillones de puntos de datos para entender el

comportamiento de los clientes y el impacto de los programas y las actividades de marketing. [58]

El aumento de los medios digitales interactivos ha catapultado el contacto entre empresas y consumidores del modelo tradicional Web 1.0 al mundo altamente interactivo de la Web 2.0, donde los consumidores dictan la naturaleza, el alcance y el contexto de los intercambios de marketing. [28]

El concepto de Social Media es una de las prioridades de la agenda de muchos ejecutivos de negocios en la actualidad, debido a que las redes sociales representan una nueva tendencia revolucionaria que debería ser de interés para las empresas que operan en el espacio en línea, o en cualquier espacio [39].

Kaplan and Haenlein [39] indican que los medios sociales son un grupo de aplicaciones basadas en Internet que se basan en los fundamentos ideológicos y tecnológicos de la Web 2.0 y que permiten la creación e intercambio de contenido generado por el usuario.

Según los resultados de una encuesta realizada a 51 expertos en inteligencia artificial publicados en Newstex [63] se espera que el uso de dicha disciplina crezca exponencialmente durante los primeros años de la década del 2020, ofreciendo ventajas como la generación de contenido personalizado o la segmentación eficaz de los mercados, entre muchas más. Sin embargo, existen varios retos a afrontar, entre ellos la limitante que muchas organizaciones presentan para poder administrar el flujo de información necesario para poder asegurar una implementación correcta de disciplinas como inteligencia artificial.

Para el caso de Colombia, existen empresas que han hecho grandes avances en la implementación de herramientas apalancadas en inteligencia artificial y han obtenido ventajas frente a los demás competidores en el mercado. Algunos ejemplos citados por Redacción TecnoSfera [69] son, el caso de iFood, quienes mediante la implementación de modelos de inteligencia artificial (IA) en su plataforma de domicilios no solo han conseguido generar ofertas personalizadas a sus usuarios, sino que también han logrado disminuir los tiempos de espera de sus domiciliarios. [69] Por otro lado, la empresa Cubiq go ha implementado sistemas basados en IA en la industria logística que identifican las dimensiones y el peso de las encomiendas, contribuyendo a la aceleración en el proceso de envío para empresas como TCC o 472, adicionalmente, Cubiq go ofrece servicios de autogestión de envíos a través de quioscos en donde los usuarios deben ingresar la información de la entrega, mientras el sistema analiza las características del paquete en paralelo y así el usuario solo debe hacer el pago para realizar el envío.

Según Redacción TecnoSfera [69] a esta tendencia de digital se han sumado empresas como Ecopetrol, quienes mediante el uso de metodologías de deep learning implementaron un bot llamado Lucas para resolver dudas referentes a la

normatividad del proceso de compras y proveedores, adicionalmente, con la implantación de sensores en los equipos, la IA puede analizar patrones de uso con el fin de validar el mejor momento para reemplazar piezas críticas. Por último, SAS Analítica es otra empresa liderando la adaptación de IA a diferentes campos, SAS ha desarrollado un servicio dirigido a bancos en el que por medio del análisis de datos masivos y su correlación con diferentes variables, apoya a organizaciones del sector bancario a identificar las situaciones críticas en las que se deben tomar decisiones con el fin de mejorar la experiencia del cliente.

Por otra parte, Cigüenza [14] resalta el esfuerzo del estado colombiano para la adopción de tecnologías soportadas en Machine Learning e IA como lo son el caso de la Fiscalía, quienes han implementado un sistema que les permite identificar asociaciones entre casos de delincuencia en diferentes ciudades con el fin agilizar la solución de dichos casos, el sistema utilizado por la Fiscalía puede analizar en tiempo real más de 13 millones de quejas en el sistema e identificar la información necesaria para establecer las asociaciones entre casos. Otro ejemplo [14] es la herramienta empleada por el Ministerio de Tecnologías de la Información y las Comunicaciones en conjunto con Agrosavia, con la cual logran realizar predicciones para más de 200 tipos de cultivos diferentes con el fin de generar apoyos más efectivos a los campesinos.

Otra implementación admirable de IA y machine learning según Jahuin [38], la realizó la empresa Financiera Efectiva, esta compañía quien durante un momento de expansión no contaba con la infraestructura para atender satisfactoriamente la totalidad de los requerimientos de sus clientes, decidió otra poa la implementación de tecnologías de inteligencia artificial. Es en este punto donde crearon a María, un asesor virtual que en una primera fase se encargaba de gestionar el contacto con clientes por medios digitales e iniciar el procedimiento de preaprobación de creaditos, una vez preaprobado María asignaba el trámite a un asesor. En la segunda fase María aprendió a responder las inquietudes sobre los servicios de la financiera y a gracias a su implementación la organización logró colocar créditos por más de \$20 millones.

Finalmente, para el análisis cuantitativo de los datos se incluyen las variables o key words (KW) de machine learning y social media tracking en Scopus obteniendo un archivo exportable RAIS el cual fue procesado en el software VOSviewer, como se evidencia en la siguiente figura:

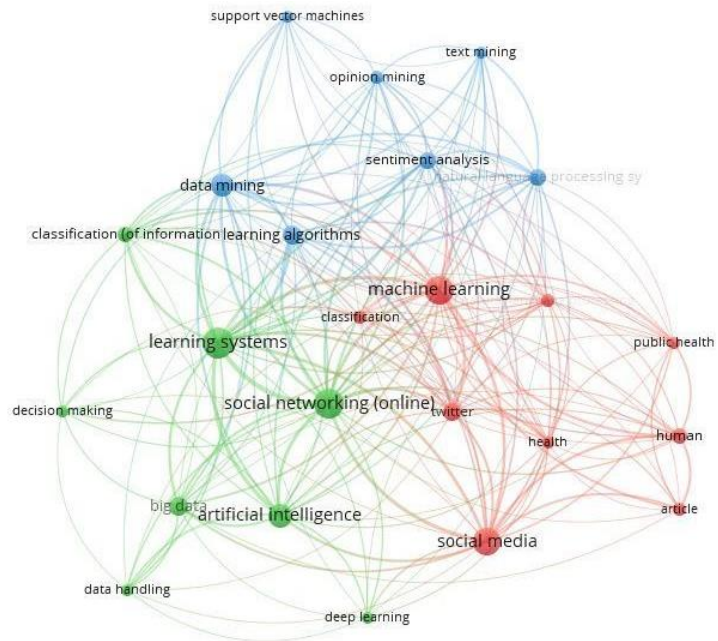


Figura 8 Mapa co-ocurrencias Machine learning and Social media tracking. [62]

Al realizarle zoom en machine learning se observa en la figura 9 en el mapa bibliométrico palabras enfáticas como inteligencia artificial, social media, data mining entre otros, lo cual nos da un contexto de evaluación y gestión de la investigación:

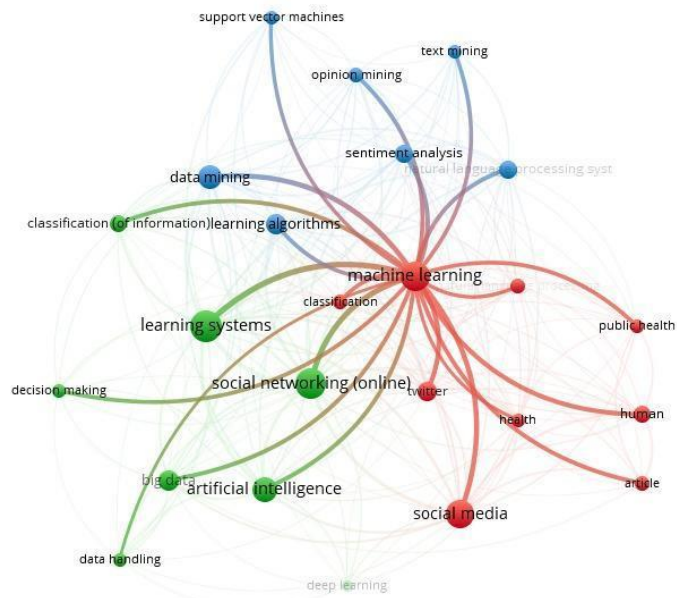


Figura 9 Mapa co-ocurrencias - zoom Machine Learning. [62]

7. METODOLOGÍA

Para el desarrollo de este proyecto se llevó a cabo el siguiente diagrama de procesos:

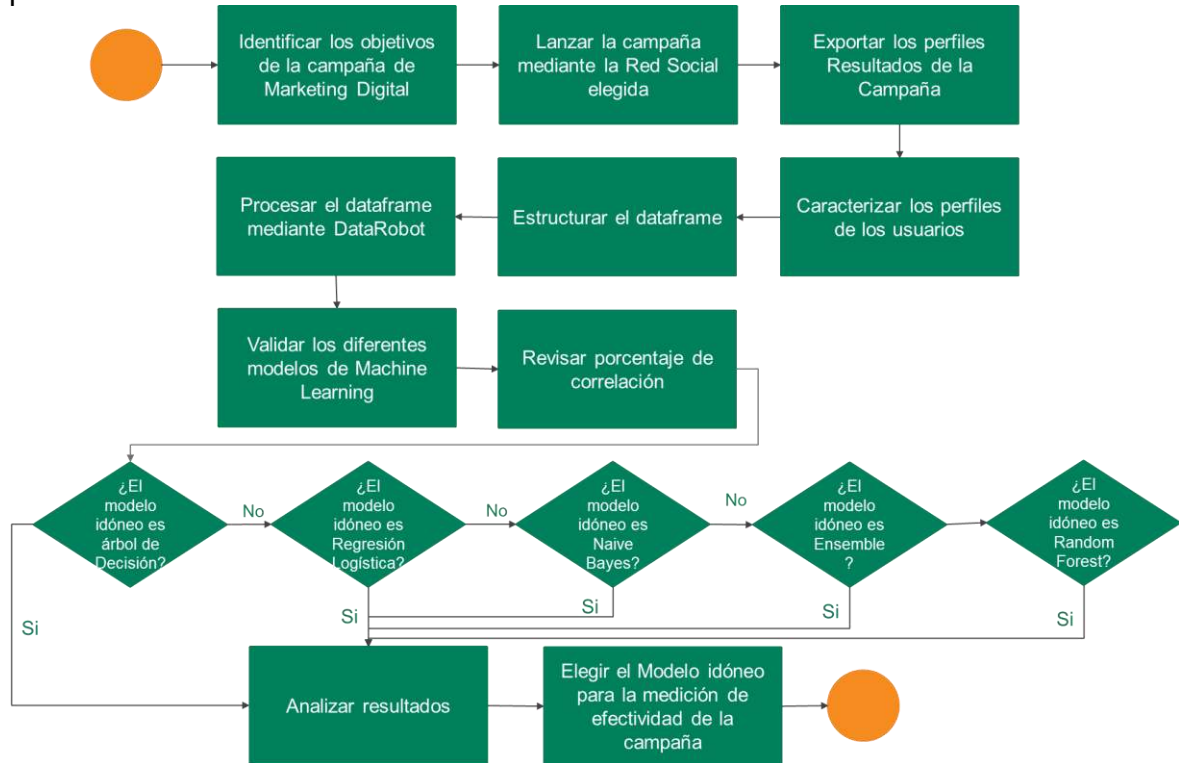


Figura 10 Diagrama - Metodología. Elaboración Propia.

Con el fin implementar el modelo de machine learning que evaluó la efectividad de una estrategia de marketing digital para el operador de información a través de social media analytics, se debió identificar los objetivos de la campaña de marketing digital, elegir la red social en donde se llevó a cabo la campaña, publicar la campaña durante el periodo de tiempo definido, exportar los perfiles de nuevos seguidores, caracterizar los perfiles de los 106 nuevos seguidores, estructurar el dataframe, procesar el dataframe mediante datarobot, validar los diferentes modelos de machine learning, revisar el porcentaje de correlación de cada modelo, elegir el modelo con mayor correlación el cual es el idóneo para la medición de la efectividad de la campaña y analizar los resultados.

8. CARACTERIZACIÓN DE LOS PERFILES DE LOS CLIENTES

8.1. DESCRIPCIÓN DE LA CAMPAÑA

La campaña fue creada por el Operador de Información con los objetivos de dar a conocer la empresa, mejorar la percepción de la empresa, aumentar los seguidores y al aumentar los seguidores se buscaba reclutar ingenieros desarrolladores.

Para la campaña se crea un video y se publica mediante la red social de LinkedIn.



Figura 11 Video Campaña LinkedIn Operador de Información [4]

El slogan de la campaña fue: “*Sé parte de un lugar y un equipo de trabajo en donde encontrarás muchas oportunidades. Únete a nuestra comunidad y conoce todas las ofertas laborales que tenemos para ti*”.

La campaña se llevó a cabo del 15 de diciembre de 2020 al 13 de enero de 2021 y el costo de inversión fue de 499,38 USD, para ingresar a la campaña se puede mediante el siguiente enlace: <https://www.linkedin.com/posts/activity-6737573074926997504-5hul/>

8.2. DESARROLLO DE LA CAMPAÑA

Se identificó que la campaña realizada obtuvo 106 nuevos seguidores en las siguientes fechas:

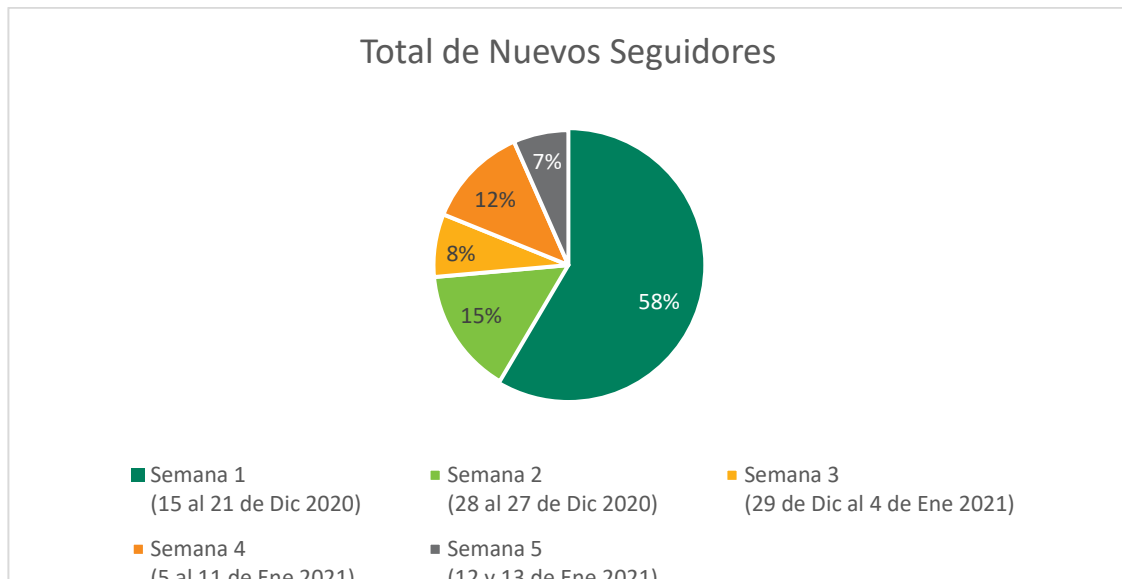


Figura 12 Total de seguidores obtenidos por campaña de Marketing Digital para el Operador de Información [49]

De esos 106 nuevos seguidores se buscan las url de los perfiles de LinkedIn, se definen los criterios que se desean obtener de cada perfil para construir la Base de Datos, los criterios definidos fueron:

- ID
- Sexo
- Ciudad
- Titulo_Profesional
- Máximo_Nivel_Formación
- Titulo_Máximo_Nivel_Formación

- Del Primer_Trabajo, se identifica:
 - Cargo
 - Empresa
 - Tiempo (meses)

- Del Penúltimo_Trabajo, se identifica:
 - Cargo
 - Empresa
 - Tiempo (meses)

- Del Último_Trabajo, se identifica:
 - Cargo
 - Empresa
 - Función Laboral
 - Tiempo (meses) Hasta2020-12

Posteriormente, de los 106 perfiles se busca manualmente los 16 criterios anteriormente definidos.

En donde se segmenta cada uno de los 8 criterios arrojando los siguientes datos:

- En la siguiente tabla se observa la segmentación de la categoría Sexo, registrándose una numeración de 0 a 2:

Sexo	
Descripción	No
Sin información	0
Masculino	1
Femenino	2

Tabla 2 Segmentación, Categoría: Sexo. Elaboración Propia.

- En la siguiente tabla se observa la segmentación de la categoría Ciudad, registrándose una numeración de 0 a 13:

Ciudad	
Descripción	No
Sin información	0
Bogotá D.C	1
Cali	2
Barranquilla	3
Guayaquil	4
Medellín	5
Ucrania	6
Atlántico	7
Cartagena	8
Soledad	9
Ipiales	10
Bucaramanga	11
Venezuela	12
Valencia	13

Tabla 3 Segmentación, Categoría: Ciudad. Elaboración Propia.

- En el Anexo A se observa la segmentación de la categoría Título Profesional, registrándose una numeración de 0 a 34.
- En la siguiente tabla se observa la segmentación de la categoría Máximo Nivel de Formación, registrándose una numeración de 0 a 7:

Máximo_Nivel_Formación	
Descripción	No

Sin información	0
Bachiller	1
Técnico	2
Tecnólogo	3
Profesional	4
Especialista	5
Máster	6
Doctor	7

Tabla 4 Segmentación, Categoría: Máx. Nivel de Formación. Elaboración Propia.

- En el Anexo B se observa la segmentación de la categoría Título Máximo Nivel de Formación, registrándose una numeración de 0 a 64.
- En el Anexo C se observa la segmentación de la categoría Cargos, registrándose una numeración de 0 a 240.
- En el Anexo D se observa la segmentación de la categoría Empresas, registrándose una numeración de 0 a 217.
- En la siguiente tabla se observa la segmentación de la categoría Función Laboral, registrándose una numeración de 0 a 25:

Función Laboral	
Descripción	No
Sin información	0
Ingeniería	1
Operaciones	2
Ventas	3
Desarrollo Empresarial	4
Tecnología de la Información	5
Educación	6
Jurídico	7
Finanzas	8
Arte y Diseño	9
Administración	10
Investigación	11
Ayuda	12
Marketing	13
Recursos Humanos	14
Gestión de Programas y Proyectos	15
Consultoría	16
Contabilidad	17

Medios de Comunicación	18
Control de Calidad	19
Servicios Sociales y Comunitarios	20
Servicios Sanitarios	21
Bienes Raíces	22
Servicios Militares y de Protección	23
Gestión de Producto	24
Liderazgo	25

Tabla 5 Segmentación, Categoría: Función Laboral. Elaboración Propia.

Con la segmentación de las 8 categorías para los 106 nuevos seguidores, se da por cumplido el objetivo de la caracterización de los perfiles.

9. MODELOS DE MACHINE LEARNING (APRENDIZAJE SUPERVISADO)

9.1. ÁRBOL DE DECISIÓN

Un árbol de decisiones [41] es la base del modelo de machine learning Random Forest. Este se compone de una serie de preguntas de sí / no, formuladas sobre los datos que eventualmente conducen a una posible solución.

En Machine Learning un árbol de decisión es similar a un diagrama de flujo en donde un nodo interno representa una característica o atributo, la rama representa una regla de decisión y cada nodo representa el resultado.

El nodo superior en un árbol de decisión se conoce como el nodo raíz. Este aprende a particionar en función del valor de la característica o atributo.

Cada nodo en el árbol actúa como un caso de prueba para alguna característica o atributo, y cada borde que desciende de ese nodo corresponde a una de las posibles soluciones. Este proceso es recursivo y se repite para cada subárbol enraizado en los nuevos nodos. [76]

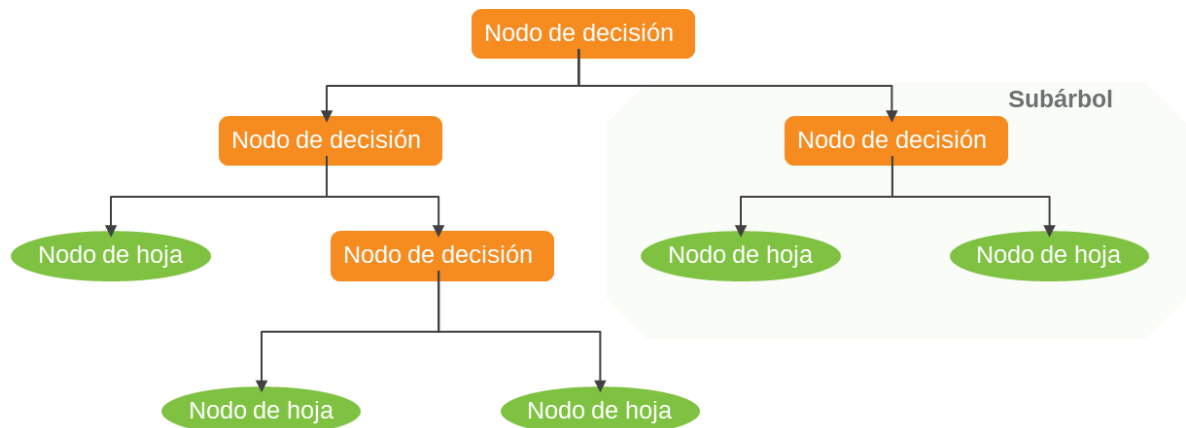


Figura 13 Estructura básica de árbol de decisión. Elaboración Propia a partir de [76]

9.1.1. ¿Cómo funciona? La idea básica detrás de cualquier algoritmo de árbol de decisión es el siguiente: [76]

- Seleccionar el mejor atributo para dividir los registros.
- Haga que ese atributo sea un nodo de decisión y divida el conjunto de datos en subconjuntos más pequeños, hasta que una de las condiciones coincida:
 - Todas las tuplas (o lista ordenada) pertenecen al mismo valor de atributo.
 - No quedan más atributos.
 - No hay más instancias.

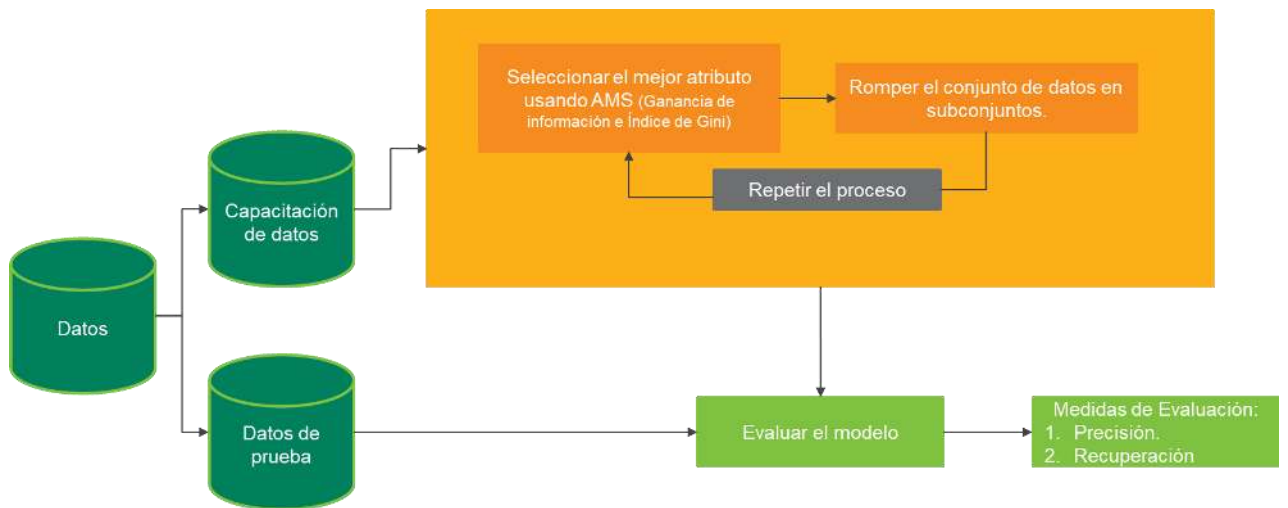


Figura 14 ¿Cómo funciona el algoritmo del árbol de decisión en Machine Learning?.
 Elaboración Propia a partir de [76]

9.1.2. Matemática - Medidas de selección de atributos (AMS). Se conoce como reglas de división porque ayuda a determinar puntos de interrupción para tuplas (o lista ordenada) en un nodo dado. [76]

Las medidas de selección de atributos proporcionan un rango para cada característica o atributo y el atributo de mejor puntuación se seleccionará como un atributo de división

Las medidas de selección más comunes son Ganancia de información e Índice de Gini.

○ **Ganancia de Información**

La ganancia de información es una propiedad estadística que mide qué tan bien un atributo dado separa los ejemplos de entrenamiento de acuerdo con su clasificación objetivo. [76]

Y para conocer la ganancia de información es necesario calcular la entropía del conjunto de datos.

La entropía mide la impureza del conjunto de entrada, además de ser la aleatoriedad o impureza en el sistema. La ganancia de información es una disminución de la entropía. [76]

Es válido decir que cuanto más “impuro” sea un conjunto de datos, mayor será la entropía y menos “impuro” un conjunto de datos, menor será la entropía.

La entropía se calcula así:

$$Entropía = -p \log_2 p - q \log_2 q$$

En donde:

p y q: es la probabilidad de éxito y fracaso respectivamente en un nodo.

Y la ganancia de información se calcula como la diferencia entre la entropía antes de la división (nodo padre) y la entropía promedio después de la división (nodo hijo) del conjunto de datos en función de los valores de los atributos de los datos, [76] así:

$$\text{Ganancia de Información} = \text{Entropía (nodo padre)} - (\text{Promedio Entropía (nodo hijo)})$$

○ Índice de Gini

La medida de selección Índice de Gini dice que, si se selecciona dos elementos de una población al azar, entonces deben ser de la misma clase y la probabilidad de esto es 1 si la población es pura. [76]

- Funciona con la variable objetivo-categorica “Éxito” o “Fracaso”.
- Realiza solo divisiones binarias.
- Cuanto mayor sea el valor de Gini, mayor será la homogeneidad.
- CART (árbol de clasificación y regresión) utiliza el método Gini para crear divisiones binarias.

El índice de Gini se calcula así:

$$Gini = (p^2 + q^2)$$

En donde:

p y q: es la probabilidad de éxito y fracaso.

Adicionalmente, el término ‘Impureza de Gini’ se determina restando el valor de Gini de 1, así: [76]

$$\text{Impureza de Gini} = 1 - Gini$$

9.2. RANDON FOREST

El modelo o algoritmo de Machine Learning Ramdon Forest, es un conjunto de árboles de decisión combinados con bagging. [56]

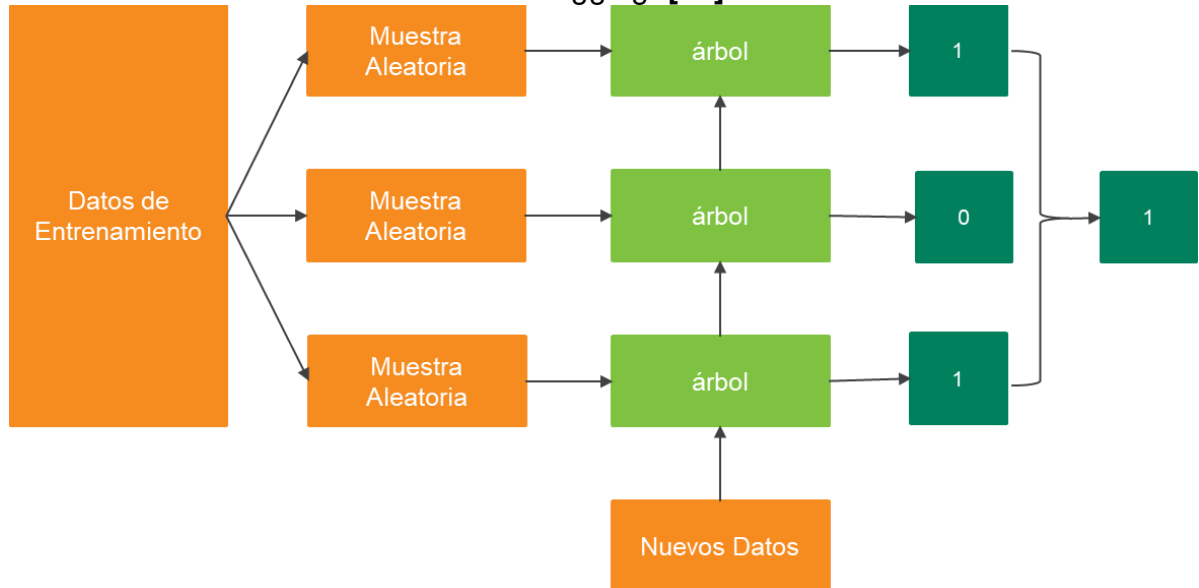


Figura 15 Ramdon Forest Básico. Elaboración Propia a partir de [56]

Todo lo relacionado a arboles de decisiones se podrá encontrar en el numeral 9.1.

9.2.1. ¿Qué es el proceso de bagging y cómo funciona? La agregación bootstrap o los métodos de bagging son métodos donde los algoritmos simples son usados en paralelo. Y lo que se busca es aprovecharse de la independecia que hay entre los algoritmos simples, ya que el error se puede reducir bastante al promediar las salidas de los modelos simples. El principal objetivo de los algoritmos de bagging es el de la reducción de la varianza. [27]

Al usar bagging, lo que en realidad está pasando, es que distintos árboles ven distintas porciones de los datos. Ningún árbol ve todos los datos de entrenamiento. Esto hace que cada árbol se entrene con distintas muestras de datos para un mismo problema. De esta forma, al combinar sus resultados, unos errores se compensan con otros y tenemos una predicción que generaliza mejor. [56]

La forma de conseguir que los errores se compensen entre sí es que cada modelo se entrena con subconjuntos del conjunto de entrenamiento. Estos subconjuntos se forman eligiendo muestras aleatoriamente (con repetición) del conjunto de entrenamiento. [55]

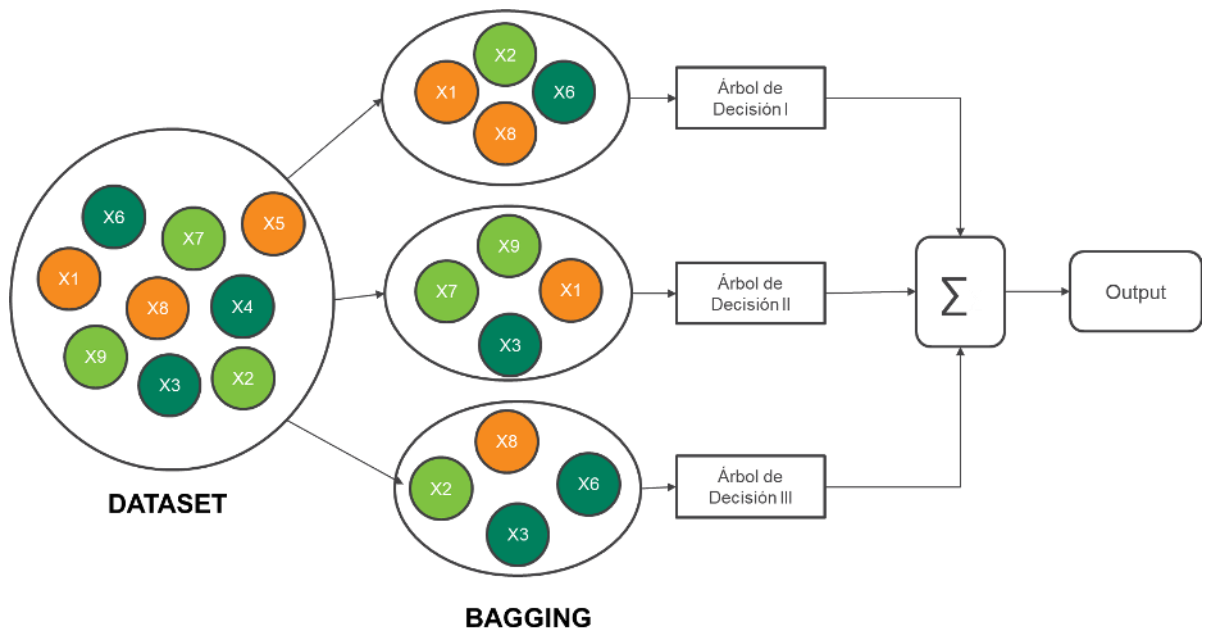


Figura 16 Random Forest (Bagging - Árboles de Decisiones). Elaboración Propia a partir de [64]

En conclusión, el Bagging sirve para: **[64]**

- Crear múltiples subconjuntos de datos
- Construir múltiples modelos
- Combinar los modelos.

En conclusión, “*Random Forest es una técnica de aprendizaje automático supervisada basada en árboles de decisión. Su principal ventaja es que obtiene un mejor rendimiento de generalización para un rendimiento durante entrenamiento similar. Esta mejora en la generalización la consigue compensando los errores de las predicciones de los distintos árboles de decisión.*” Y para asegurarse que los árboles sean distintos, lo que se hace es que cada uno se entrena con una muestra aleatoria de los datos de entrenamiento y a esta estrategia se le denomina bagging. **[56]**

9.3. CLASIFICACIÓN DE NAÏVE BAYES

Los modelos de Naive Bayes se basan en una técnica de clasificación estadística llamada “teorema de Bayes”.

Son una clase especial de algoritmos de clasificación de Machine Learning o Aprendizaje Automático, estos modelos son llamados algoritmos “Naive”, o “Inocentes” en español.

En este modelo se asume que las variables predictoras son independientes entre sí. En otras palabras, que la presencia de una cierta característica en un conjunto de datos no está en absoluto relacionada con la presencia de cualquier otra característica. [71]

La facilidad de construir estos modelos se debe por la simplicidad de este; Esta simplicidad la consiguen proporcionando una forma de calcular la probabilidad ‘posterior’ de que ocurra un cierto evento, dadas algunas probabilidades de eventos ‘anteriores’. [71]

Se calcula así [71]:

$$P(A|R) = \frac{P(R|A)P(A)}{P(R)}$$

En donde:

P(A) es la Probabilidad de A

P(R|A) es la probabilidad de que se dé R dado A

P(R) es la probabilidad de R

P(A|R) es la probabilidad posterior de que se dé A dado R

9.3.1. Ejemplo del algoritmo Naive Bayes. Dos compañeros que trabajan en la misma oficina, se sabe cómo información anterior que [71]:

- Ana viene a la oficina 3 días a la semana.
- Beto viene a la oficina 1 día a la semana.

Alguien pasa corriendo y no se logra identificar si fue Ana o Beto, de acuerdo con esta información y asumiendo que solo trabajan 4 días a la semana, las probabilidades de que la persona vista sea Ana o Beto son:

- P(Ana) = 3/4 = 0.75
- P(Beto) = 1/4 = 0.25

Cuando se vio pasar la persona, se identificó que él o ella llevaba una chaqueta roja. También se sabe lo siguiente:

- Ana viste de rojo 2 veces a la semana.
- Beto viste de rojo 3 veces a la semana.

Así que, para cada semana de trabajo, que tiene cinco días, podemos inferir lo siguiente:

- La probabilidad de que Ana vista de rojo es $\rightarrow P(\text{Rojo}|\text{Ana}) = 2/5 = 0.4$
- La probabilidad de que Beto vista de rojo $\rightarrow P(\text{Rojo}|\text{Beto}) = 3/5 = 0.6$
- Entonces, con esta información, ¿a quién se vio pasar? (en forma de probabilidad)
- Esta nueva probabilidad será la ‘posterior’.

Inicialmente conocíamos las probabilidades $P(\text{Ana})$ y $P(\text{Beto})$, y después inferíamos las probabilidades de $P(\text{rojo}|\text{Ana})$ y $P(\text{rojo}|\text{Beto})$.

De forma que las probabilidades reales son:

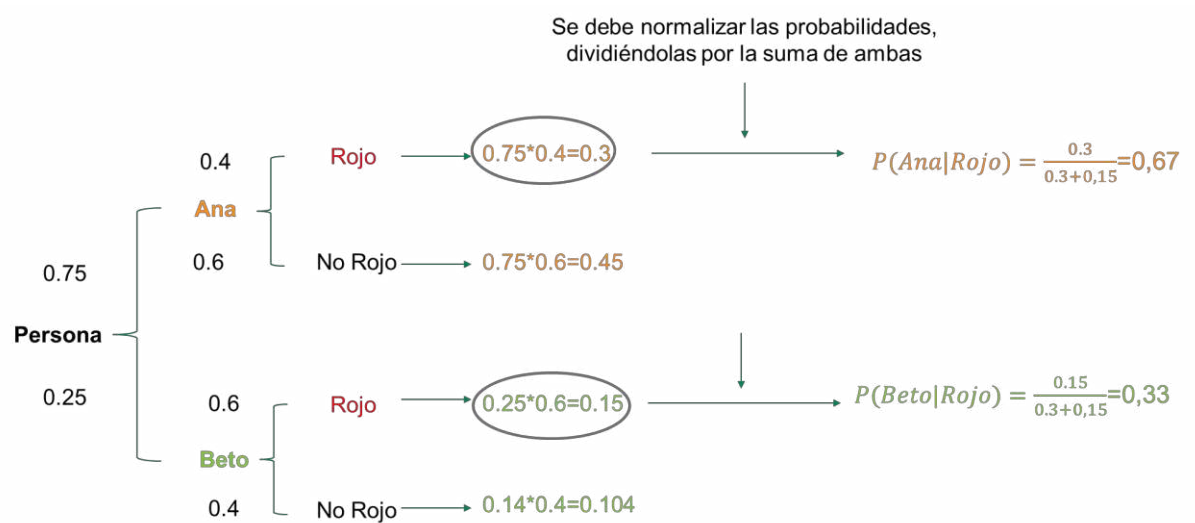


Figura 17 Ejemplo Modelo Naive Bayes – Probabilidades Reales. Elaboración Propia a partir de [71]

Y el grafico formal seria:

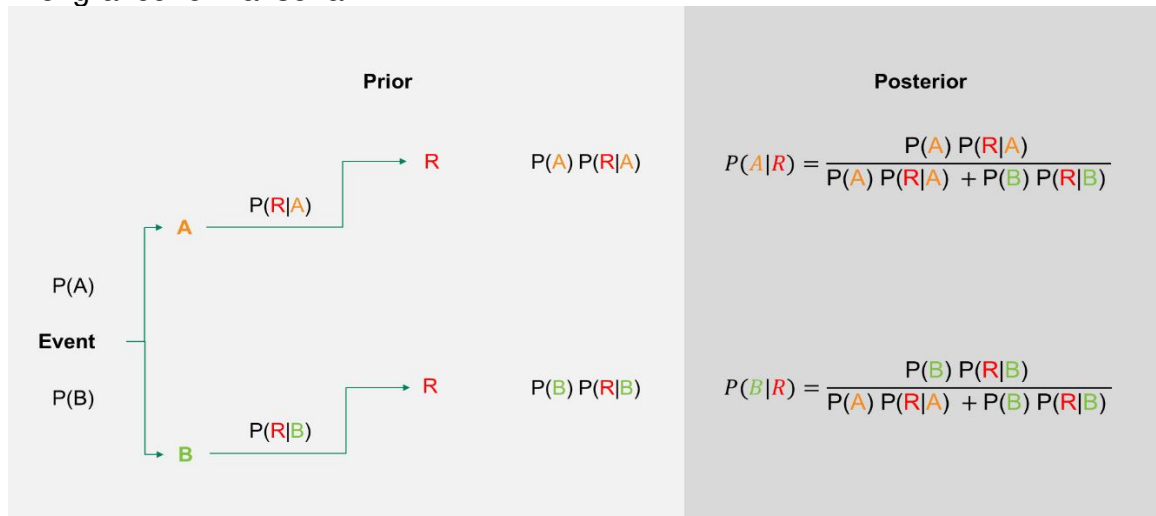


Figura 18 Ejemplo Modelo Naive Bayes – Gráfico formal. Elaboración Propia a partir de [71]

9.3.2. Desventajas del algoritmo Naive Bayes. A continuación, se relación algunas desventajas de este modelo:

- A pesar de ser clasificadores bastante buenos, los algoritmos Naive Bayes son conocidos por ser pobres estimadores. Por ello, no se toman con mucha certeza las probabilidades que se obtienen. [71]
- La presunción de independencia Naive muy probablemente no reflejará cómo son los datos en el mundo real. [71]
- Cuando el conjunto de datos de prueba tiene una característica que no ha sido observada en el conjunto de entrenamiento, el modelo le asignará una probabilidad de cero y será inútil realizar predicciones. [71]

9.4. REGRESIÓN LOGÍSTICA

La técnica de aprendizaje automatizado de Regresión Logística es una red neuronal con exactamente una neurona.

9.4.1. Matemática detrás de la Regresión Logística

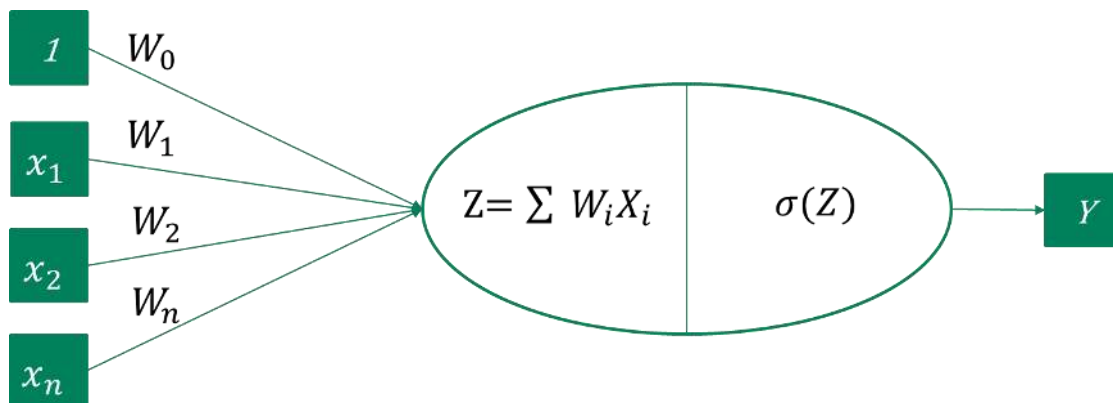


Figura 19 Matemática detrás de la Regresión Logística. Elaboración Propia a partir de [57]

En donde los valores de x corresponden a los distintos atributos del problema.

Y matemáticamente se puede formular así: [57]

$$y = \sigma(z) = \sigma(WX) = \sigma\left(\sum (w_i x_i)\right) = \sigma\left(\sum (w_0 x_0 + w_1 x_1 + \dots + w_n x_n)\right)$$

Se evidencia que la regresión logística tiene dos partes: [57]

- Una combinación lineal (a la izquierda de la neurona)
- Aplicación de la función logística (a la derecha de la neurona)

Así que todas las entradas se combinan con una línea con los coeficientes w . Y luego se aplica la función logística (también llamada sigmoide) al resultado.

- **Función logística o sigmoide**

Se expresa matemáticamente así: **[57]**

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

En donde la función logística o sigmoide: **[57]**

- Está acotada entre 0 y 1. Su valor mínimo es 0 y el máximo es 1.
- Se pueden interpretar los resultados como probabilidades.
- Para problemas de clasificación binaria, se puede suponer que los valores menores de 0.5 corresponden a la clase 0 y los superiores a 0.5 a la clase 1.

9.4.2. Desventajas de la Regresión Logística. A continuación, se relación algunas desventajas de este modelo:

- No maneja números grandes de características o variables categóricas.
- No funciona bien cuando el espacio de características es demasiado grande.
- Imposibilidad de resolver directamente problemas no lineales, debido a que la expresión que toma la decisión es lineal. Por ejemplo, en el caso de que la probabilidad de una clase se reduzca inicialmente con una característica y posteriormente suba no puede ser registrado con un modelo logístico directamente. Siendo necesario transforma esta característica previamente para que el modelo puede registrar este comportamiento no lineal. **[70]**

9.5. MÉTODOS “ENSEMBLE” (CONJUNTOS DE CLASIFICADORES)

Los métodos combinados de aprendizaje (métodos de ensemble) utiliza múltiples algoritmos de aprendizaje para obtener un rendimiento predictivo que mejore el que podría obtenerse de los algoritmos de aprendizaje individuales que lo componen. [73]

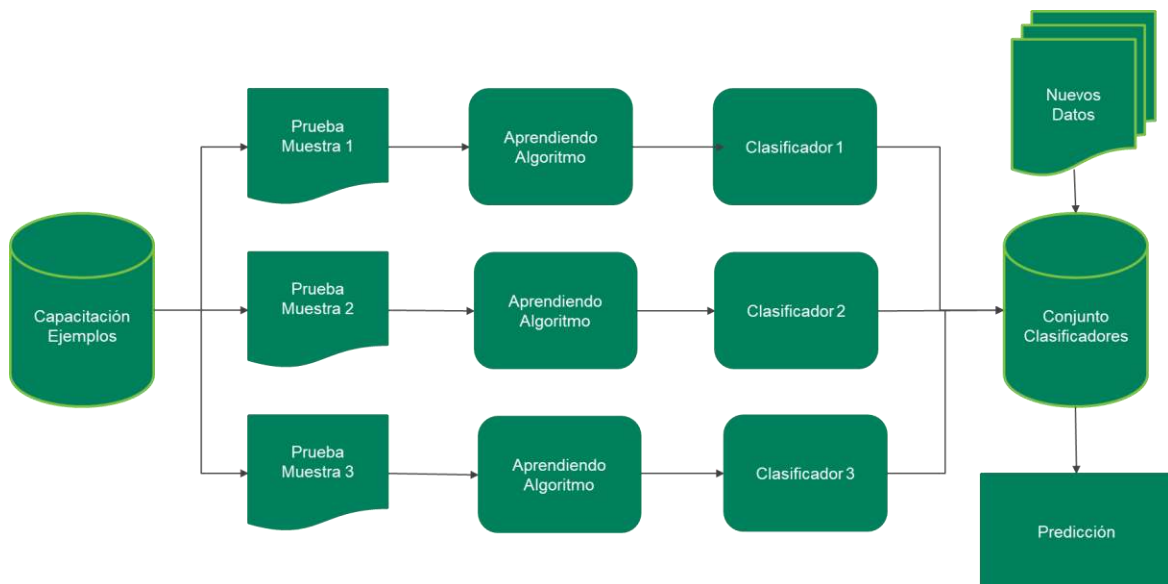


Figura 20 Métodos de ensemble. Elaboración Propia a partir de [73]

Los métodos combinados “ensemble” tienen como objetivo considerar múltiples hipótesis simultáneamente para formar una hipótesis que con suerte se comporte mejor.

9.5.1. Métodos de combinación más comunes en “ensemble”. Los métodos de combinación más comunes usados en ensemble son: Agregación Bootstrap o Bagging, Boosting y Subespacios Aleatorios. [73]

La agregación Bootstrap o Bagging se puede consultar en el numeral 9.2.1.

○ **Boosting**

En el método boosting no se crean versiones del conjunto de entrenamiento, sino que se trabaja siempre con el conjunto completo de entrada, y se manipulan los pesos de los datos para generar modelos distintos. [73] El objetivo es que en cada iteración se incremente el peso de los objetos mal clasificados por el predictor en esa iteración, por lo que en la construcción del próximo predictor estos objetos serán más importantes y será más probable clasificarlos bien.

El método de boosting más usado es el AdaBoost, en el cual se realizan los siguientes pasos: [73]

- A todos los datos del conjunto de entrenamiento se les asigna un peso idéntico: $w_i = \frac{1}{n}$ en donde n es el tamaño del conjunto de datos.
- Se entrena el modelo usando el conjunto de entrenamiento.
- Se calcula el error del modelo en el conjunto de entrenamiento, se cuentan cuántos objetos han sido mal clasificados y se identifican cuáles son.
- Se incrementan los pesos en aquellos casos de entrenamiento en los que el modelo anterior ha dado resultados erróneos.
- Se entrena un nuevo modelo usando el conjunto de pesos modificados.
- Volver al punto del cálculo del error del modelo en el conjunto de entrenamiento (y se repite el proceso hasta el número de iteraciones fijadas inicialmente).
- El modelo final se consigue por votación ponderada usando los pesos de todos los modelos. [73]

○ Subespacios Aleatorios

En el método de subespacios aleatorios cada modelo se entrena con todos los ejemplos, pero solo considera un subconjunto de los atributos. [73]

El tamaño de estos subconjuntos es el parámetro del método, y de nuevo el resultado es el promedio o votación de los resultados individuales de los modelos.

La combinación se realiza con el siguiente algoritmo: [73]

- Sea N el tamaño del conjunto de entrenamiento, y A el número de atributos de los datos.
- Sea L el número de clasificadores individuales del conjunto.
- Para cada clasificador individual, h , se toma dh con $dh < A$, el número de variables de entrada de h . Comúnmente se usa el mismo valor de dh para todos los clasificadores individuales.
- Para cada clasificador individual h , se crea un conjunto de entrenamiento seleccionado al azar dh atributos distintos de los datos de entrada, y se entrena con ellos al clasificador.
- Para clasificar un nuevo objeto, se deben combinar las salidas de los L clasificadores individuales por combinación ponderada. [73]

9.5.2. Desventaja de los métodos “ensemble”. Debido a que los métodos combinados hacen uso de varias hipótesis simultáneas, se produce una elevación en los costes computacionales. [73]

10. EFECTIVIDAD DEL MODELO A TRAVÉS DE SOCIAL MEDIA ANALYTICS

Con base en las tablas de segmentación de categorías por: Sexo, Ciudad, Título Profesional, Máximo Nivel de Formación, Título Máximo Nivel de Formación, Cargos, Empresas y Función laboral; las cuales se encuentran en el numeral 8.3. DESARROLLO DE LA CAMPAÑA, se realiza el Dataframe bruto codificándose así:

Link LinkedIn	Sexo	Ciudad	Titulo_Profesional	Máximo_Nivel_Formación	Titulo_Máximo_Nivel_Formación	Primer_Trabajo			Penúltimo_Trabajo			Último_Trabajo				
						Cargo	Empresa	Tiempo (meses)	Cargo	Empresa	Tiempo (meses)	Cargo	Empresa	Función Laboral	Tiempo (meses) Hasta2020-12	
https://www.linkedin.com/in/...	2	1	1	5	1	1	1	1	93	96	11	165	163	14	38	
https://www.linkedin.com/in/...	1	0	2	4	2	2	2	37	0	0	0	2	164	1	82	
https://www.linkedin.com/in/...	2	0	3	4	3	3	3	17	0	0	0	3	3	2	17	
https://www.linkedin.com/in/...	1	2	4	4	4	4	4	24	94	97	7	166	165	1	12	
https://www.linkedin.com/in/...	1	1	5	4	5	5	5	8	95	98	16	8	166	10	20	
https://www.linkedin.com/in/...	2	3	6	4	6	6	6	13	96	99	37	96	167	7	14	
https://www.linkedin.com/in/...	2	0	5	6	7	7	7	11	97	100	2	167	42	13	16	
https://www.linkedin.com/in/...	2	1	7	4	8	8	8	9	98	101	4	168	168	20	16	
https://www.linkedin.com/in/...	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
https://www.linkedin.com/in/...	2	1	8	5	9	9	9	7	99	102	20	169	125	1	15	
https://www.linkedin.com/in/...	2	1	5	5	10	10	10	6	100	10	24	170	10	8	1	
https://www.linkedin.com/in/...	2	1	9	4	11	11	11	1	101	11	1	171	163	4	5	
https://www.linkedin.com/in/...	2	1	10	4	12	12	12	4	102	103	2	172	170	5	25	
https://www.linkedin.com/in/...	2	3	5	4	5	13	13	102	103	104	24	173	35	3	202	
https://www.linkedin.com/in/...	1	1	11	5	13	14	14	26	104	105	37	174	105	1	1	
https://www.linkedin.com/in/...	2	2	1	4	14	15	15	16	105	106	9	175	171	4	51	
https://www.linkedin.com/in/...	2	4	12	4	15	16	16	28	106	107	27	176	19	4	99	
https://www.linkedin.com/in/...	1	1	8	4	16	17	17	35	107	102	4	177	125	1	14	
https://www.linkedin.com/in/...	1	1	12	4	15	0	0	0	0	0	0	0	0	1	0	
https://www.linkedin.com/in/...	2	5	13	2	17	18	18	49	0	0	0	0	18	18	8	49
https://www.linkedin.com/in/...	1	1	14	4	18	19	19	36	108	108	9	178	172	9	1	
https://www.linkedin.com/in/...	1	1	0	0	0	20	20	0	0	0	0	0	20	20	1	0
https://www.linkedin.com/in/...	1	1	15	4	19	21	21	13	109	21	45	179	173	1	22	
https://www.linkedin.com/in/...	1	1	16	6	20	22	22	6	110	109	10	180	174	1	1	
https://www.linkedin.com/in/...	1	1	0	0	0	23	23	0	0	0	0	23	23	2	0	
https://www.linkedin.com/in/...	1	1	0	0	0	24	24	24	0	0	0	24	24	24	24	
https://www.linkedin.com/in/...	2	1	11	4	21	0	0	0	0	0	0	0	0	8	0	

Tabla 6 Dataframe bruto. Elaboración Propia.

El dataframe bruto se convierte en dataframe caracterizado, como se evidencia a continuación:

ID	Sexo	Ciudad	Titulo_Profesional	Máximo_Nivel_Formación	Titulo_Máximo_Nivel_Formación	Primer_Trabajo			Penúltimo_Trabajo			Último_Trabajo			
						Cargo	Empresa	Tiempo (meses)	Cargo	Empresa	Tiempo (meses)	Cargo	Empresa	Función Laboral	Tiempo (meses) Hasta2020-12
0.0094	1,0000	0,0769	0,0294	0,7143	0,0156	0,0042	0,0046	0,0032	0,3875	0,4424	0,0351	0,6875	0,7512	0,5600	0,1214
0.0189	0,5000	0,0000	0,0588	0,5714	0,0313	0,0083	0,0092	0,1182	0,0000	0,0000	0,0000	0,0083	0,7558	0,0400	0,2620
0.0283	1,0000	0,0000	0,0882	0,5714	0,0469	0,0125	0,0138	0,0543	0,0000	0,0000	0,0000	0,0125	0,0138	0,0800	0,0543
0.0377	0,5000	0,1538	0,1176	0,5714	0,0625	0,0167	0,0184	0,0767	0,3917	0,4470	0,0224	0,6917	0,7604	0,0400	0,0383
0.0472	0,5000	0,0769	0,1471	0,5714	0,0781	0,0208	0,0230	0,0256	0,3958	0,4516	0,0511	0,0333	0,7650	0,4000	0,0639
0.0566	1,0000	0,2308	0,1765	0,5714	0,0938	0,0250	0,0276	0,0415	0,4000	0,4562	0,1182	0,4000	0,7696	0,2800	0,0447
0.0660	1,0000	0,0000	0,1471	0,8571	0,1094	0,0292	0,0323	0,0351	0,4042	0,4608	0,0064	0,6958	0,1935	0,5200	0,0511
0.0755	1,0000	0,0769	0,2059	0,5714	0,1250	0,0333	0,0369	0,0288	0,4083	0,4654	0,0128	0,7000	0,7742	0,8000	0,0511
0.0849	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
0.0943	1,0000	0,0769	0,2353	0,7143	0,1406	0,0375	0,0415	0,0224	0,4125	0,4700	0,0639	0,7042	0,5760	0,0400	0,0479
0.1038	1,0000	0,0769	0,1471	0,7143	0,1583	0,0417	0,0461	0,0192	0,4167	0,0461	0,0767	0,7083	0,0461	0,3200	0,0032
0.1132	1,0000	0,0769	0,2647	0,5714	0,1719	0,0458	0,0507	0,0032	0,4208	0,0507	0,0032	0,7125	0,7788	0,1600	0,0160
0.1226	1,0000	0,0769	0,2941	0,5714	0,1875	0,0500	0,0553	0,0128	0,4250	0,4747	0,0064	0,7167	0,7834	0,2000	0,0799
0.1321	1,0000	0,2308	0,1471	0,5714	0,0781	0,0542	0,0599	0,3259	0,4292	0,4793	0,0767	0,7208	0,1613	0,1200	0,6454
0.1415	0,5000	0,0769	0,3235	0,7143	0,2031	0,0583	0,0645	0,0831	0,4333	0,4839	0,1182	0,7250	0,4839	0,0400	0,0032
0.1509	1,0000	0,1538	0,0294	0,5714	0,2188	0,0625	0,0691	0,0511	0,4375	0,4885	0,0288	0,7292	0,7880	0,1600	0,1629
0.1604	1,0000	0,3077	0,3529	0,5714	0,2344	0,0667	0,0737	0,0895	0,4417	0,4931	0,0863	0,7333	0,0876	0,1600	0,3163
0.1698	0,5000	0,0769	0,2353	0,5714	0,2500	0,0708	0,0783	0,1118	0,4458	0,4700	0,0128	0,7375	0,5760	0,0400	0,0447
0.1792	0,5000	0,0769	0,3529	0,5714	0,2344	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0400	0,0000
0.1887	1,0000	0,3846	0,3824	0,2857	0,2656	0,0750	0,0829	0,1565	0,0000	0,0000	0,0000	0,0750	0,0829	0,3200	0,1565
0.1981	0,5000	0,0769	0,4118	0,5714	0,2813	0,0792	0,0876	0,1150	0,4500	0,4977	0,0288	0,7417	0,7926	0,3600	0,0032
0.2075	0,5000	0,0769	0,0000	0,0000	0,0000	0,0833	0,0922	0,0000	0,0000	0,0000	0,0000	0,0833	0,0922	0,0400	0,0000
0.2170	0,5000	0,0769	0,4412	0,5714	0,2969	0,0875	0,0968	0,0415	0,4542	0,0968	0,1438	0,7458	0,7972	0,0400	0,0703
0.2264	0,5000	0,0769	0,4706	0,8571	0,3125	0,0917	0,1014	0,0192	0,4583	0,5023	0,0319	0,7500	0,8018	0,0400	0,0032
0.2358	0,5000	0,0769	0,0000	0,0000	0,0000	0,0958	0,1060	0,0000	0,0000	0,0000	0,0000	0,0958	0,1060	0,0800	0,0000

Tabla 7 Dataframe caracterizado. Elaboración Propia.

Posteriormente, el dataframe caracterizado se pasa por data robot, con el objetivo de comparar cada uno de los algoritmos de machine learning. El data robot realiza el siguiente flujo:

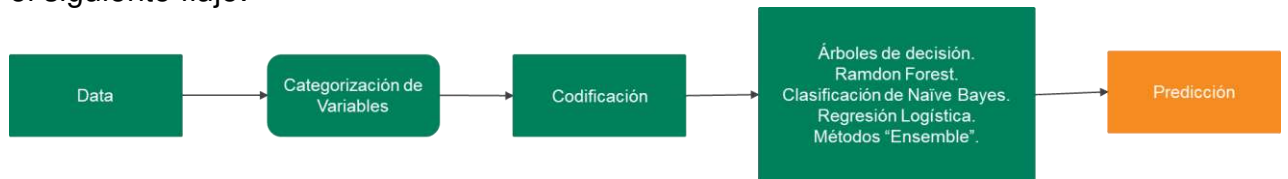


Figura 21 Flujo data robot. Elaboración propia.

Es decir, el dataframe se pasa por cada uno de los algoritmos o modelos de machine learning; y se obtienen los siguientes resultados:

10.1. RESULTADO DATAFRAME PARA: CLASIFICACIÓN DE NAÏVE BAYES

A continuación, se observa el resultado obtenido después de pasar el dataframe por data robot para el algoritmo o modelo de machine learning identificado como Clasificación de Naïve Bayes:

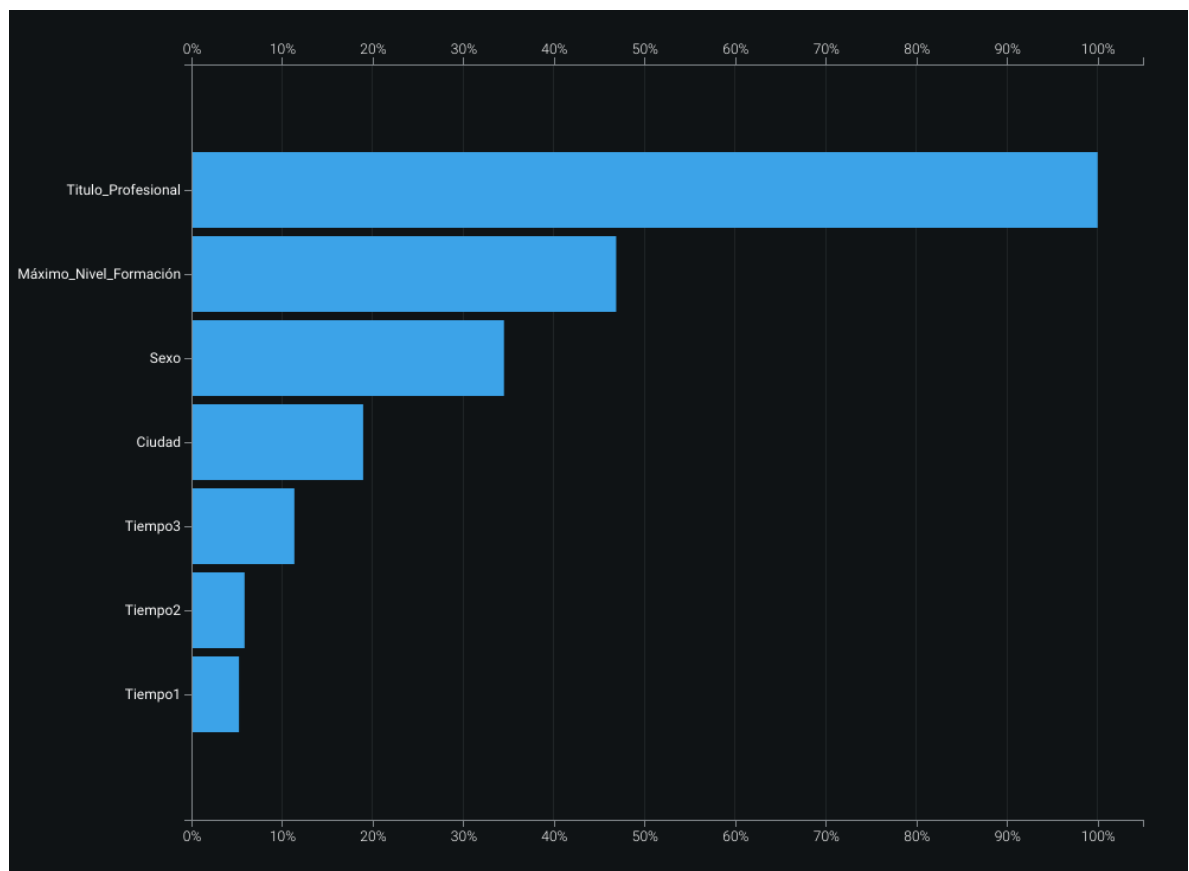


Figura 22 Resultado dataframe para: Clasificación de Naïve Bayes. Elaboración Data Robot.

Según la gráfica anterior se evidencia, que al pasar el dataframe versus la variable objetivo la cual era Función Laboral por el modelo de machine learning de Clasificación de Naïve Bayes, que: El porcentaje de correlación de la variable objetivo para 6 criterios de 7 estuvo por debajo del 80%.

10.2. RESULTADO DATAFRAME PARA: REGRESIÓN LOGÍSTICA

A continuación, se observa el resultado obtenido después de pasar el dataframe por data robot para el algoritmo o modelo de machine learning identificado como Regresión Logística:

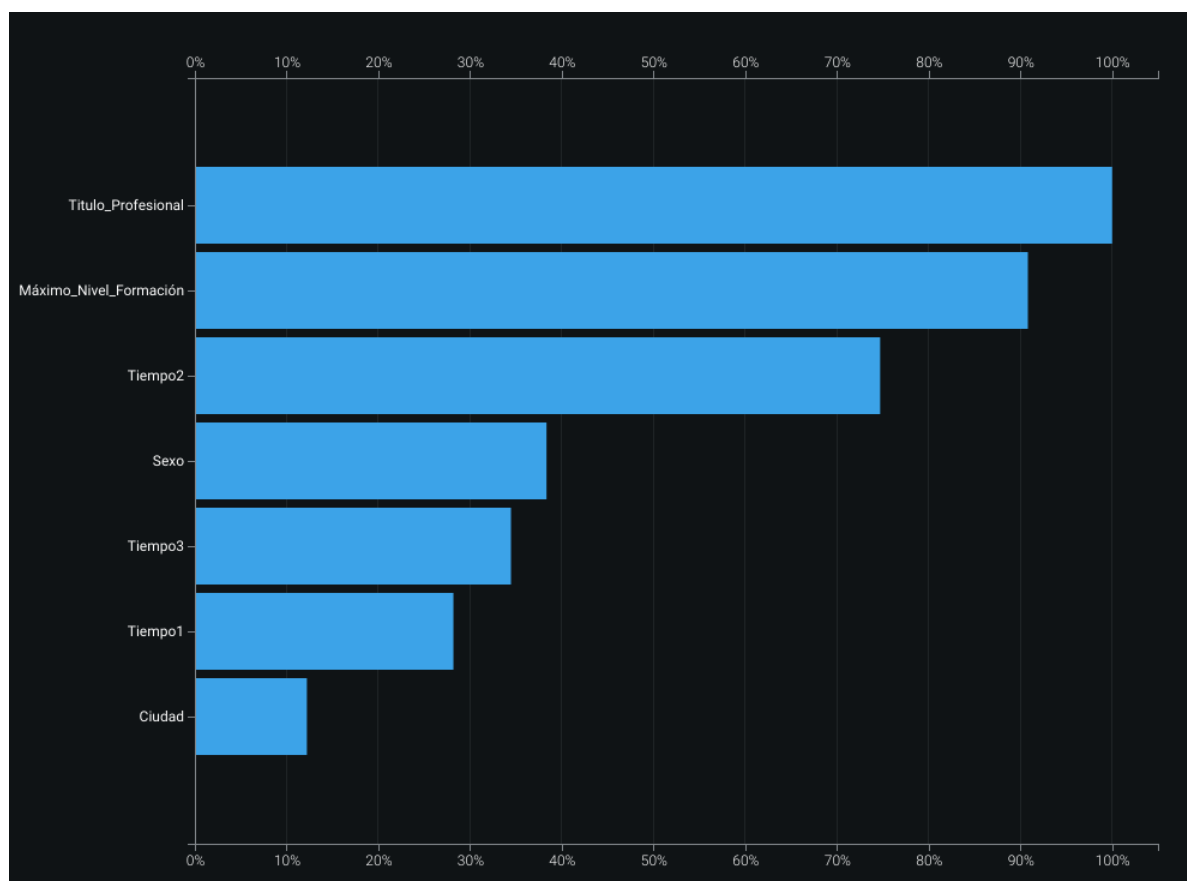
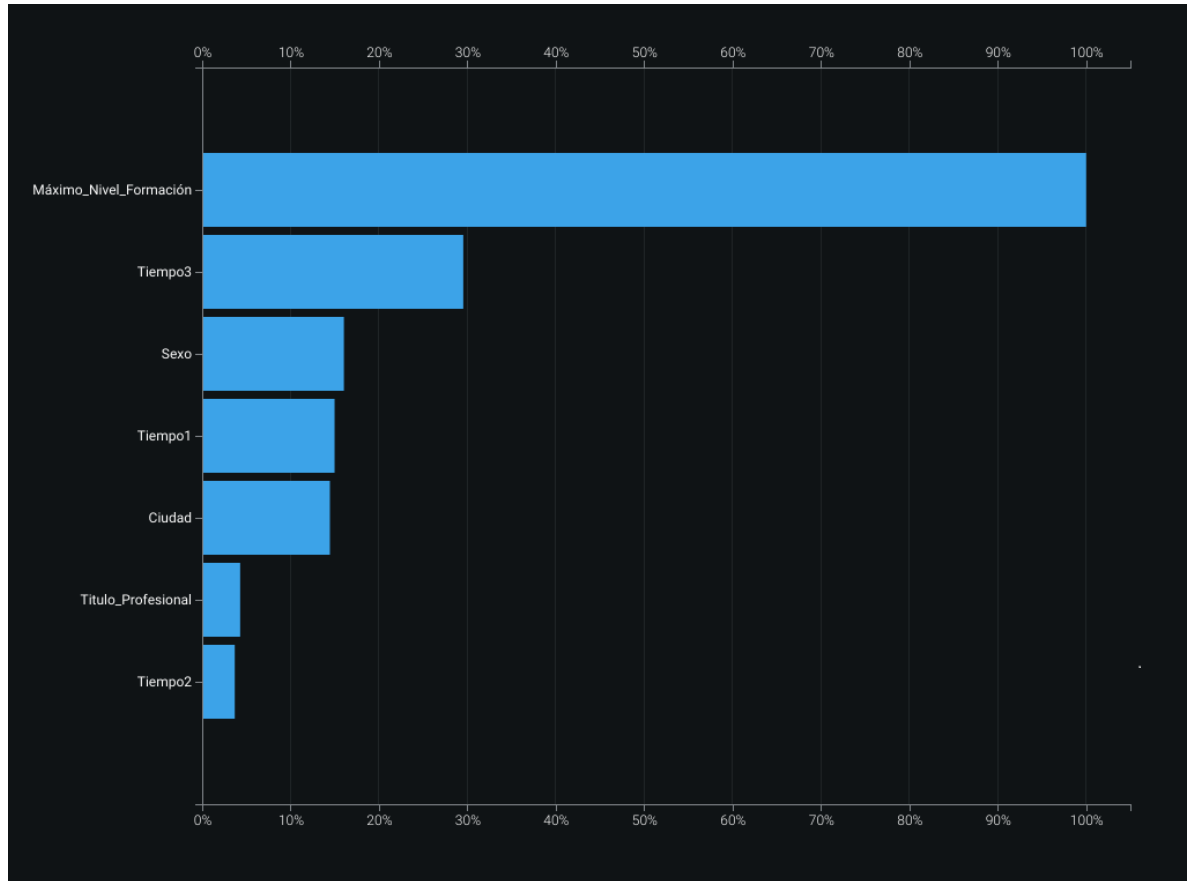


Figura 23 Resultado dataframe para: Regresión Logística. Elaboración Data Robot.

Según la gráfica anterior se evidencia, que al pasar el dataframe versus la variable objetivo la cual era Función Laboral por el modelo de machine learning de Regresión Logística, que: El porcentaje de correlación de la variable objetivo para 5 criterios de 7 estuvo por debajo del 80%.

10.3. RESULTADO DATAFRAME PARA: MÉTODOS “ENSEMBLE”

A continuación, se observa el resultado obtenido después de pasar el dataframe por data robot para el algoritmo o modelo de machine learning identificado como Métodos “Ensemble” (Conjuntos de clasificadores):



**Figura 24 Resultado dataframe para: Métodos “Ensemble” (Conjuntos de clasificadores).
Elaboración Data Robot.**

Según la gráfica anterior se evidencia, que al pasar el dataframe versus la variable objetivo la cual era Función Laboral por el modelo de machine learning de Métodos “Ensemble” (Conjuntos de clasificadores), que: El porcentaje de correlación de la variable objetivo para 5 criterios de 7 estuvo por debajo del 80%.

10.4. RESULTADO DATAFRAME PARA: RAMDON FOREST (GINI)

A continuación, se observa el resultado obtenido después de pasar el dataframe por data robot para el algoritmo o modelo de machine learning identificado como Ramdon Forest (Gini):

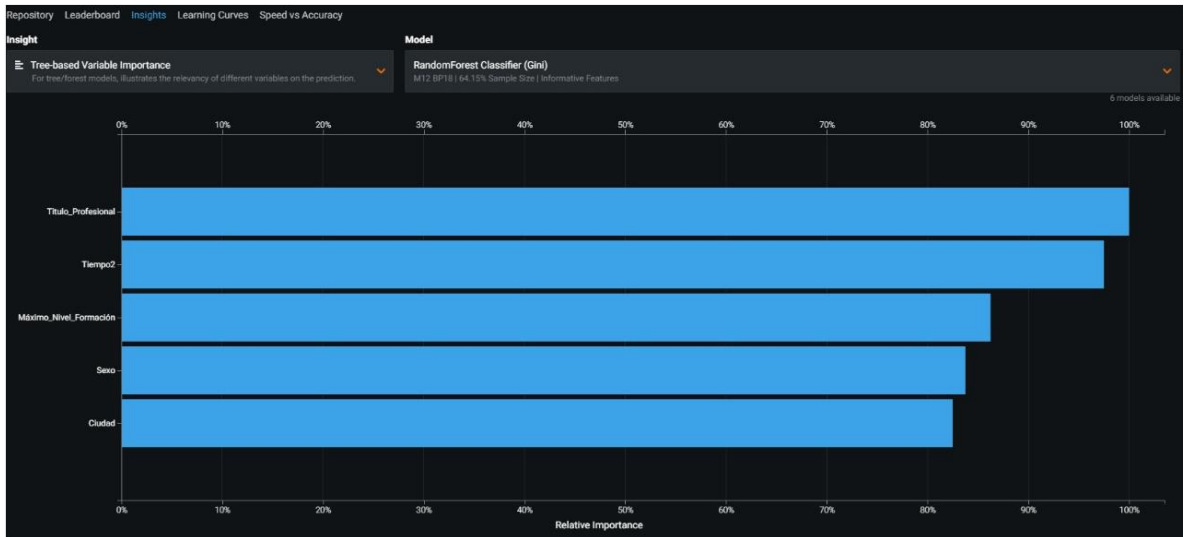


Figura 25 Resultado dataframe para: Ramdon Forest (Gini). Elaboración Data Robot.

Al aplicar el modelo de Random Forest (Gini) la correlación obtenida entre el impacto de la variable *Función Laboral* versus la información del dataframe bruto se obtiene un porcentaje mayor del 80%.

Es importante aclarar que la variable objetivo era Función Laboral, debido a que el objetivo era atraer un nuevo talento específico con perfil de ingeniero para el área de desarrollo; por ende, la predicción se hace con base a Función Laboral, en donde función laboral contenía los siguientes datos:

Función Laboral	
Descripción	No
Ingeniería	1
Operaciones	2
Ventas	3
Desarrollo Empresarial	4
Tecnología de la Información	5
Educación	6
Jurídico	7
Finanzas	8
Arte y Diseño	9

Administración	10
Investigación	11
Ayuda	12
Marketing	13
Recursos Humanos	14
Gestión de Programas y Proyectos	15
Consultoría	16
Contabilidad	17
Medios de Comunicación	18
Control de Calidad	19
Servicios Sociales y Comunitarios	20
Servicios Sanitarios	21
Bienes Raíces	22
Servicios Militares y de Protección	23
Gestión de Producto	24
Liderazgo	25

Tabla 8 Data Función Laboral. Elaboración Propia.

En la siguiente grafica en el eje x se observa el valor de objetivo promedio y en el eje y la probabilidad de ocurrencia, en el cual encuentra correlación de los datos al graficarlos todos; en donde la correlación hace referencia a que al ordenar los datos crecen o decrecen en el mismo sentido:



Figura 26 Gráfico de Elevación, Resultado Random Forest (Gini). Elaboración Data Robot.

En esta grafica se observa la línea naranja la cual hace referencia a los datos actuales y la azul a la predicción, en donde la actual son los datos que se tenían en el dataframe y la predicción la realizo uno a uno aprendiendo a partir de los datos que se tenían.

10.5. RESULTADO LINKEDIN

Se obtiene de LinkedIn un informe nombrado como Administrador de Campañas el cual muestra en Función Laboral el número y porcentaje de Impresiones y el número y porcentaje de clics:

Función Laboral	Impresiones			Clics	
	No	#	%	#	%
Ingeniería	1	1472	13,45%	24	11,76%
Operaciones	2	937	8,56%	24	11,76%
Ventas	3	812	7,42%	12	5,88%
Desarrollo Empresarial	4	686	6,27%	8	3,92%
Tecnología de la Información	5	665	6,08%	8	3,92%
Educación	6	585	5,35%	8	3,92%
Juridico	7	501	4,58%	6	2,94%
Finanzas	8	495	4,52%	4	1,96%
Arte y Diseño	9	494	4,52%	17	8,33%
Administración	10	458	4,19%	13	6,37%
Investigación	11	401	3,67%	0	0,00%
Ayuda	12	369	3,37%	9	4,41%
Marketing	13	336	3,07%	7	3,43%
Recursos Humanos	14	334	3,05%	0	0,00%
Gestión de Programas y Proyectos	15	331	3,03%	8	3,92%
Consultoría	16	328	3,00%	10	4,90%
Contabilidad	17	325	2,97%	10	4,90%
Medios de Comunicación	18	282	2,58%	9	4,41%
Control de Calidad	19	221	2,02%	0	0,00%
Servicios Sociales y Comunitarios	20	199	1,82%	0	0,00%
Servicios Sanitarios	21	170	1,55%	3	1,47%
Bienes Raices	22	117	1,07%	0	0,00%
Servicios Militares y de Protección	23	95	0,87%	0	0,00%
Gestión de Producto	24	88	0,80%	3	1,47%
Liderazgo	25	80	0,73%	0	0,00%

Figura 27 Resultado Administrador de Campañas - LinkedIn

El indicador a medir para dar a conocer la marca fue: Porcentaje de clics (CTR): el número de clics dividido entre las impresiones.

$$CTR = \frac{C}{I}$$

En donde:

- CTR: es el porcentaje de clics
- C: es el número de clics

- I: son las impresiones

El resultado del CTR fue:

$$CTR = \frac{183}{10.781} = 0,0169 = 1,6974\%$$

La media del CTR para LinkedIn es de 0.05% [54] y en la campaña se obtuvo el 1,6974% es decir que la campaña estuvo arriba de la media del CTR.

Adicionalmente, se quería reclutar nuevos talentos con perfiles de ingeniería en el campo de desarrollo o software. En donde Ingeniería fue el área con mayor impresiones y clics de la muestra. Y se logra incorporar un nuevo talento como ingeniero desarrollador durante el mes de diciembre del 2020.

11.RESULTADOS

La ventaja competitiva es que las soluciones de proyectos relacionados con machine learning y social media analytics se encuentran en las ciencias de la computación y no más del 2% en negocios y con baja relevancia en marketing.

El carácter novedoso del proyecto es la integración de un modelo de machine learning a través de social media analytics para operadores de información nacionales.

Se identifica que los países que más publican artículos de machine learning vinculados al social media tracking son Estados Unidos, India y el Reino Unido, lo cual genera apertura de la inclusión a Colombia a través de esta investigación; además, de segmentar el proyecto a operadores tecnológicos los cuales aún no han vinculado ninguna de estas tecnologías.

Entre los objetivos de esta campaña se requería reclutar nuevos talentos con perfiles de ingeniería en el campo de desarrollo o software, en donde Ingeniería fue el área con mayor impresiones y clics de la muestra. Y se logra incorporar un nuevo talento como ingeniero desarrollador durante el mes de diciembre del 2020.

De acuerdo con el procedimiento de marketing del operador de información la salida o entregable en su mayoría es una pieza impresa, un brochure o material POP; sin embargo, en estrategias no efectivas de marketing se ha evidencia el fracaso del proyecto posterior a la generación de este entregable, ocasionando que se deba destruir el material final. Teniendo en cuenta esta información el objetivo del proyecto impacto ambientalmente al eliminar la impresión de material que deberá ser desechado por campañas no efectivas de marketing.

12. CONCLUSIONES

De acuerdo con la caracterización se identifica que el impacto de la campaña decrece en el tiempo, lo cual es evidenciable al observar que en la primera semana se obtuvieron el 58% de los registros de seguidores mientras que en la quinta semana solo el 7% como se evidencia en la Figura 12.

El impacto de la caracterización tiene un comportamiento potencial decreciente con una relación aproximada de $0,45 \times 1,2$ con una intersección en el eje Y de 0,85 al graficar la relación de semana e impacto e identificar la línea de tendencia que más se ajusta.

Existían diferentes riesgos técnicos asociados al proyecto, como la pérdida de información por fallas en el almacenamiento de altos volúmenes de datos o la dependencia de las redes sociales (LinkedIn) para la captura de información. Evidenciando que para obtener mayores resultados fue necesario realizar una inversión total de 499,38 USD con el fin de que la red social mostrara en los perfiles segmentados la publicación de la campaña. Los resultados de esta inversión se evidenciaron en la caracterización de los perfiles que empezaron a seguir la página de la red social de LinkedIn del Operador de Información en donde el 30% de los seguidores su función laboral se desempeñaba en el campo de la ingeniería, siendo este el público objetivo.

Se investigaron diferentes modelos de machine learning como: Árboles de decisión, Random Forest, Clasificación de Naïve Bayes, Regresión Logística y Métodos "Ensemble" (Conjuntos de clasificadores); sin embargo, de acuerdo con características específicas el modelo ideal para la validación de efectividad de esta estrategia de marketing fue el de Random Forest con un porcentaje de correlación mayor al 80% entre el impacto de la variable Función Laboral versus la información del dataframe bruto; en donde este modelo se caracteriza por la utilidad en la resolución de problemas de clasificación.

13. FUTUROS TRABAJOS

Se recomienda continuar con el trabajo validando estrategias de marketing digital para operadores de información en las otras redes sociales (Facebook, Twitter e Instagram) teniendo en cuenta que el social media analytics ofrece el acceso a la información requerida para evaluar la efectividad de las campañas e incorporando modelos de machine learning. El mercado potencial es nacional, para este proyecto son los operadores de información autorizados por el Ministerio de la Protección Social.

14. REFERENCIAS BIBLIOGRAFICAS

- [1] American Marketing Association, Definitions of Marketing, (2017), Consultado el: 05-05-2020, Disponible en: <https://www.ama.org/the-definition-of-marketing-what-is-marketing/>.
- [2] Aportes en Línea S.A, ¿Qué es un Operador de Información?, Consultado el: 12-05-2020, Disponible en: http://aportesenlinea.custhelp.com/app/answers/detail/a_id/69/~/%C2%BFqu%C3%A9-es-un-operador-de-informaci%C3%B3n%3F.
- [3] Aportes en Línea S.A, ¿Quiénes somos?, Consultado el: 04-06-2020, Disponible en: <https://www.aportesenlinea.com/Home/QuienesSomos.aspx>.
- [4] Aportes en Línea, Sé parte de un lugar y un equipo de trabajo en donde encontrarás muchas oportunidades. Únete a nuestra comunidad y conoce todas las ofertas laborales que tenemos para ti, LinkedIn, Consultado el: 10-05-2021, Disponible en: <https://www.linkedin.com/posts/activity-6737573074926997504-5hul/>.
- [5] Armstrong. Gary, Kotler. Philip, Fundamentos de marketing, Pearson Educación, (2013), ISBN: 9786073217224, Disponible en Biblioteca digital de Bogotá: <https://www.bibliotecadigitaldebogota.gov.co/resources/2215444/>.
- [6] Baena. Verónica, Fundamentos del marketing: entorno, consumidor, estrategia e investigación comercial, Universitat Oberta de Catalunya, (2011). ISBN: 9788497884648, Disponible en: https://books.google.com.co/books/about/Fundamentos_de_marketing_entorno_consumi.html?id=lf9sfHpscC&redir_esc=y.
- [7] Barranco Fragoso. Ricardo, ¿Qué es Big Data?, IBM, (2012). Consultado el: 01-06-2020, Disponible en: <https://www.ibm.com/developerworks/ssa/local/im/que-es-big-data/index.html>.
- [8] Bayoude. KENZA, Ouassit. Youssef, Ardchir. Soufiane, Azouazi. Mohamed, How Machine Learning Potentials are transforming the Practice of Digital Marketing: State of the Art, Periodicals of Engineering and Natural Sciences, Pg. 1-7, (2018), ISSN 2303-4521, <http://pen.ius.edu.ba>.
- [9] BBVA - Communications, 'Machine learning': what is it and how does it work?, Banco Bilbao Vizcaya Argentaria, S.A, (2019), Consultado el: 01-06-2020, Disponible en: <https://www.bbva.com/es/machine-learning-que-es-y-como-funciona/>.
- [10] Bermejo. Eduard, Martinez. Alex, Los 10 Algoritmos esenciales en Machine Learning, Raona, Consultado el: 10-05-2020, Disponible en: https://www.raona.com/los-10-algoritmos-esenciales-machine-learning/?cli_action=1620689600.574.
- [11] Boon-Long. Supond, Wongsurawat. Winai, Social media marketing evaluation using social network comments as an indicator for identifying consumer purchasing decision effectiveness, Journal of International Business Studies, Pg. 1-20, (2015), doi: 10.1057/dddmp.2015.51, https://www.researchgate.net/publication/283902399_Social_media_marketing_ev

aluation using social network comments as an indicator for identifying consumer purchasing decision effectiveness.

[12] Boyd. Danah, Ellison. Nicole, Social Network Sites: Definition, History, and Scholarship, Journal of Computer-Mediated Communication, Pg. 1-17, (2007), ISSN: 10836101, doi: 10.1111/j.1083-6101.2007.00393.x, <https://academic.oup.com/jcmc/article/13/1/210/4583062>.

[13] Chaffey. Dave, Ellis-Chadwick. Fiona, Marketing digital, Pearson Educación, (2014) ISBN: 9786073227476, Disponible en: <https://www.marcialpons.es/libros/marketing-digital/9786073227476/>.

[14] Cigüenza Riaño. Noelia, What are local companies using artificial intelligence today?, LA REPÚBLICA, (2019), Consultado el: 03-06-2020, Disponible en: <https://www.larepublica.co/internet-economy/en-que-estan-usando-las-empresas-locales-la-inteligencia-artificial-en-la-actualidad-2930773>.

[15] Constitución Política de la República de Colombia, Constitución Política, Derechos de autor reservados - Prohibida su reproducción, Consultado el: 04-06-2020, Disponible en: http://www.secretariassenado.gov.co/senado/basedoc/constitucion_politica_1991_p_r015.html.

[16] CORTE CONSTITUCIONAL, REPÚBLICA DE COLOMBIA, Ley 019 de 2012, Consultado el: 11-05-2020, Disponible en: <https://www.corteconstitucional.gov.co/relatoria/2013/C-097-13.htm>.

[17] Costa. Susana, Oliveira. Paulo Alexandre, Resende. Sara, How companies evaluate the ROI of social media marketing programmes: insights from B2B and B2C, Journal of Business & Industrial Marketing, Pg. 1-14, (2020), ISSN: 0885-8624, <https://www.emerald.com/insight/content/doi/10.1108/JBIM-06-2019-0291/full/html>.

[18] De Pelsmacker. Patrick, van Tilburgb. Sophie, Holthofb. Christian, Digital marketing strategies, online reviews and hotel performance, International Journal of Hospitality Management, Pg. 1-9, (2018), doi: 10.1016/j.ijhm.2018.01.003, ISSN: 0278-4319, <https://www.sciencedirect.com/science/article/pii/S0278431917305303>.

[19] Dotras. Alberto, Social Media: herramientas y estrategias empresariales, Ediciones de la U, (2018), ISBN: 9789587627848, Disponible en: <https://edicionesdelau.com/producto/social-media-herramientas-y-estrategias-empresariales/>.

[20] Drucker. Peter, The Practice of Management, HarperCollins, (1954), ISBN: 9780062005441, Disponible en: https://www.scribd.com/book/163567585/The-Practice-of-Management?&language_settings_changed=English.

[21] Facebook, FACEBOOK for Business, Business Help Center, About Minimum ROAS, Consultado el: 14-02-2021, Disponible en: <https://www.facebook.com/business/help/1113453135474912?id=2196356200683573>.

[22] Facebook, FACEBOOK for Business, Business Help Center, Setting a Budget for Instagram Ads, Consultado el: 14-02-2021, Disponible en: <https://www.facebook.com/business/help/1514927061975521?id=332010350818053>.

- [23] Facebook, FACEBOOK for Business, Help service for companies, How ad billing works on Facebook, Consultado el: 13-02-2021, Disponible en: https://es-la.facebook.com/business/help/716180208457684?id=1792465934137726&recommended_by=105373712886516.
- [24] Fischer. Laura, Espejo. Jorge, Mercadotecnia, McGraw-Hill Interamericana S.A, (2017), ISBN: 9786071513922, Disponible en: <https://librerianacional.com/producto/356603>.
- [25] Fradkov. Alexander L, Early History of Machine Learning, IFAC-PapersOnLine, Pg. 1-6, (2020), ISSN 2405-8963, Disponible en: <https://www.sciencedirect.com/science/article/pii/S2405896320325027>.
- [26] García. Juan, Marketing logístico, Universidad Complutense de Madrid, (2004), ISSN: 1131-6985, Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=1301195>.
- [27] Gonzalo. Alvaro, ¿Cuál es la diferencia entre los métodos de bagging y los de boosting?, machinelearningparatodos.com, Consultado el: 11-05-2021, Disponible en: <https://machinelearningparatodos.com/cual-es-la-diferencia-entre-los-metodos-de-bagging-y-los-de-boosting/>.
- [28] Hanna. Richard, Rohm. Andrew, Crittenden. Victoria L, We're all connected: The power of the social media ecosystem, Business Horizons, Pg. 1-10, (2011), doi: 10.1016/j.bushor.2011.01.007, ISSN: 0007-6813, <https://www.sciencedirect.com/science/article/pii/S0007681311000243>.
- [29] Hoffman. Donna L, Fodor. Marek, Can You Measure the ROI of Your Social Media Marketing?, MIT Sloan Management Review, Pg. 1-11, (2010), ISSN: 15329194, <https://n9.cl/oa24o>.
- [30] Hollebeek. Linda D, Glynn. Mark S, Brodie. Roderick, Consumer Brand Engagement in Social Media: Conceptualization, Scale Development and Validation, Journal of Interactive Marketing, Pg. 1-17, (2014), ISSN 1094-9968, Disponible en: <https://www.sciencedirect.com/science/article/pii/S1094996813000649>.
- [31] Human Level Communications, Maciá. Fernando, Santonja. Maria, Marketing en redes sociales, Difusora Larousse - Anaya Multimedia, (2017), EISBN: 9788441539013, Disponible en: <https://elibro-net.ucatolica.basesdedatosezproxy.com/es/ereader/ucatolica/122932?page=17>.
- [32] Hurwitz. Judith and Kirsch. Daniel, Machine Learning For Dummies, IBM Limited Edition Published by John Wiley & Sons, (2018), Consultado el: 01-06-2020, Disponible en: <https://www.ibm.com/downloads/cas/GB8ZMQZ3>.
- [33] IBM, ¿Qué es Machine Learning?, Consultado el: 01-06-2020, Disponible en: <https://www.ibm.com/co-es/analytics/machine-learning>.
- [34] insideBIGDATA, From Data to Metadata for Machine Learning Platforms, Equipo Editorial, (2020), Consultado el: 01-06-2020, Disponible en: <https://insidebigdata.com/2020/05/15/from-data-to-metadata-for-machine-learning-platforms/>.
- [35] insideBIGDATA, How Artificial Intelligence is Revolutionizing How We Sell, Equipo Editorial, (2019), Consultado el: 31-05-2020, Disponible en: <https://insidebigdata.com/2019/12/21/how-artificial-intelligence-is-revolutionizing-how-we-sell/>.

- [36] insideBIGDATA, Infografía: La escasez de científicos de datos, Equipo Editorial, (2018), Consultado el: 31-05-2020, Disponible en: <https://insidebigdata.com/2018/08/19/infographic-data-scientist-shortage/>.
- [37] Iruretagoyena. Santiago, Comercio electrónico, McGraw-Hill Interamericana S.A, (2015), ISBN: 9788448613235, Disponible en: <https://editorialmultilibros.com/producto/comercio-electronico/>.
- [38] Jahuin Gamarra. Gaby, 5 Success stories of companies in Peru that implemented Artificial Intelligence, LinkedIn, (2017), Consultado el: 03-06-2020, Disponible en: <https://www.linkedin.com/pulse/5-casos-de-%C3%A9xito-empresas-en-el-per%C3%BA-que-inteligencia-jahuin-gamarra>.
- [39] Kaplan. Andreas, and Haenlein. Michael, Users of the world, unite! The challenges and opportunities of Social Media, Business Horizons, Pg. 1-10, (2010), DOI: 10.1016/j.bushor.2009.09.003, Disponible en: https://www.researchgate.net/publication/222403703_Users_of_the_World_Unite_The_Challenges_and_Opportunities_of_Social_Media.
- [40] Kietzmann. Jan H, Hermkens. Kristopher, McCarthy. Ian P, Silvestre. Bruno S, Social media? Get serious! Understanding the functional building blocks of social media, Business Horizons, Pg. 1-11, (2011), ISSN: 0007-6813, <https://www.sciencedirect.com/science/article/pii/S0007681311000061>.
- [41] Koehrsen. Will, An Implementation and Explanation of the Random Forest in Python, Towards Data Science, (2018), Consultado el: 05-05-2021, Disponible en: <https://towardsdatascience.com/an-implementation-and-explanation-of-the-random-forest-in-python-77bf308a9b76>.
- [42] Kotler. Philip, Dirección del marketing, Pearson Educación, (2006), ISBN: 9789702607632, Disponible en Biblioteca digital de Bogotá: <https://www.bibliotecadigitaldebogota.gov.co/resources/2215331/>.
- [43] Kotler. Philip, J Levy. Sidney, Broadening the Concept of Marketing, Journal of Marketing, Pg. 1-10, (1969), doi: 10.2307/1248740, <http://www.jstor.org/stable/1248740>.
- [44] Kotler. Philip, Lane. Kevin, Dirección de marketing, Pearson Educación, (2006), ISBN: 9786073212458, Disponible en Biblioteca digital de Bogotá: <https://www.bibliotecadigitaldebogota.gov.co/resources/2215332/>.
- [45] Kotler. Philip, Marketing, Pearson Educación, (2017), ISBN: 9786073238601, Disponible en Biblioteca digital de Bogotá: <https://www.bibliotecadigitaldebogota.gov.co/resources/2215633/>.
- [46] Krizanova. Anna, Lăzăroi. George, Gajanova. Lubica, Kliestikova. Jana, Nadanyiova. Margareta, and Moravcikova. Dominika, The Effectiveness of Marketing Communication and Importance of Its Evaluation in an Online Environment, Sustainability, Pg. 1-19, (2019), doi:10.3390/su11247016, https://www.researchgate.net/publication/337842627_The_Effectiveness_of_Marketing_Communication_and_Importance_of_Its_Evaluation_in_an_Online_Environment.
- [47] Lambin. Jean Jacques, Gallucci. Carlo, Sicurello. Carlos, Dirección de marketing: gestión estratégica y operativa del mercado, McGRAW-HILL, (2009). ISBN: 978-970-10-6710-9, Disponible en:

https://www.academia.edu/11747236/Direcci%C3%B3n_de_marketing_Gesti%C3%B3n_estrat%C3%A9gica_y_operativa_del_mercado.

[48] Lappe. Marc, Spang. Konrad, Investments in project management are profitable: A case study-based analysis of the relationship between the costs and benefits of project management, International Journal of Project Management, Pg. 1-10, (2013), doi: 10.1016/j.ijproman.2013.10.005, ISSN: 2637863, <https://www.sciencedirect.com/science/article/abs/pii/S0263786313001361>.

[49] LinkedIn, Export your LinkedIn Page analytics report, LinkedIn in Help, Consultado el: 05-05-2021, Disponible en: <https://www.linkedin.com/help/linkedin/answer/96320/exportar-el-informe-de-analisis-de-tu-pagina-de-linkedin?lang=es>.

[50] LinkedIn, Marketing Solutions, Analyze the performance of your campaign, Consultado el: 06-04-2021, Disponible en: <https://business.linkedin.com/es-es/marketing-solutions/success/best-practices/analyze-your-performance>.

[51] LinkedIn, Marketing Solutions, How much does it cost to advertise on LinkedIn?, Consultado el: 14-02-2021, Disponible en: <https://business.linkedin.com/es-es/marketing-solutions/ads/pricing>.

[52] Luna. Javier, Tipos de aprendizaje automático, SoldAI - Medium, Consultado el: 10-05-2020, Disponible en: <https://medium.com/soldai/tipos-de-aprendizaje-autom%C3%A1tico-6413e3c615e2>.

[53] M. I. Jordan, and T. M. Mitchell, Machine learning: Trends, perspectives, and prospects, American Association for the Advancement of Science, Pg. 1-7, (2015), DOI: 10.1126/science.aaa8415, Disponible en: <https://science.sciencemag.org/content/349/6245/255>.

[54] Martín. Sara, Cómo hacer anuncios en LinkedIn, Metricool, Consultado el: 18-05-2021, Disponible en: <https://metricool.com/es/anuncios-en-linkedin/>.

[55] Martinez. Jose, Ensembles: voting, bagging, boosting, stacking, IArtificial.net, Consultado el: 11-05-2021, Disponible en: <https://www.iartificial.net/ensembles-voting-bagging-boosting-stacking/>.

[56] Martinez. Jose, Random Forest (Bosque Aleatorio): combinando árboles, IArtificial.net, Consultado el: 11-05-2021, Disponible en: <https://www.iartificial.net/random-forest-bosque-aleatorio/>.

[57] Martinez. Jose, Regresión Logística para Clasificación, IArtificial.net, Consultado el: 11-05-2021, Disponible en: <https://www.iartificial.net/regresion-logistica-para-clasificacion/>.

[58] Miklosik. Andrej, Evans. Nina, Impact of big data and machine learning on digital transformation in marketing: A literature review, Institute of Electrical and Electronics Engineers Inc, Pg. 1-9, (2017), doi: 10.1109/ACCESS.2020.2998754, https://www.researchgate.net/publication/341762521_Impact_of_Big_Data_and_Machine_Learning_on_Digital_Transformation_in_Marketing_A_Literature_Review.

[59] Miklosik. Andrej, Kuchta. Martin, Evans. Nina and Zak Stefan, Towards the Adoption of Machine Learning-Based Analytical Tools in Digital Marketing, Institute of Electrical and Electronics Engineers Inc, Pg. 1-14, (2019), ISSN: 2169-3536, DOI: 10.1109/ACCESS.2019.2924425, Disponible en: <https://ieeexplore.ieee.org/abstract/document/8746184>.

- [60] Minciencias, Tipología de proyectos calificados como de carácter científico, tecnológico e innovación, Minciencias, Versión 5, Consultado el: 04-06-2020, Disponible en: [http://www.colciencias.gov.co/sites/default/files/upload/paginas/documento de tipologias - version 5.pdf](http://www.colciencias.gov.co/sites/default/files/upload/paginas/documento_de_tipologias_-_version_5.pdf).
- [61] Miralles-Pechuán. Luis, Ponce. Hiram, Martínez-Villaseñor. Lourdes, A novel methodology for optimizing display advertising campaigns using genetic algorithms, Pg. 1-13, (2018), doi: 10.1016/j.elerap.2017.11.004, ISSN: 1567-4223 [https://www.researchgate.net/publication/321259861 A Novel Methodology for Optimizing Display Advertising Campaigns Using Genetic Algorithms](https://www.researchgate.net/publication/321259861_A_Novel_Methodology_for_Optimizing_Display_Advertising_Campaigns_Using_Genetic_Algorithms).
- [62] Nees Jan van Eck, and Ludo Waltman, Vosviewer CWTS, Centro de Estudios de Ciencia y Tecnología, Universidad de Leiden, Países Bajos, Consultado el: 03-06-2020, Disponible en: <https://www.vosviewer.com/download>.
- [63] Newstex, MarketingProfs: Marketing Opinions - MarketingProfs Daily Fix: The Marketing Impact of AI and Machine-Learning: 3 Predictions by 51 ML Marketing Executives, ProQuest, (2017), Consultado el: 03-06-2020, Disponible en: <https://search-proquest-com.ucatolica.basesdedatosezproxy.com/docview/2370771037?accountid=45660>.
- [64] Orellana. Johanna, Arboles de decision y Random Forest, bookdown.org, Consultado el: 11-05-2021, Disponible en: <https://bookdown.org/content/2031/>.
- [65] Prato. Laura, Villoria. Liliana, Aplicaciones Web 2.0 - Redes sociales, Eduvim, (2010), ISBN: 9789871727131, Disponible en: <https://www.eduvim.com.ar/libro/9789871727131-aplicaciones-web-20-redes-sociales>.
- [66] PRESIDENTE DE LA REPÚBLICA DE COLOMBIA, El servicio público es de todos - Función Pública, DECRETO 1465 DE 2005, Consultado el: 11-05-2020, Disponible en: <https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=16499>.
- [67] Quinn. Robert, Rohrbaugh. John, A spatial model of effectiveness criteria: towards a competing values approach to organizational analysis, Management Science, Pg. 1-16, (1983), ISSN: 251909, DOI: 10.1287/mnsc.29.3.363, Disponible en: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1070.319&rep=rep1&type=pdf>.
- [68] Recuero de los Santos. Paloma, Tipos de aprendizaje en Machine Learning: supervisado y no supervisado, Telefónica - Think Big / Empresas, Consultado el: 10-05-2020, Disponible en: <https://empresas.blogthinkbig.com/que-algoritmo-elegir-en-ml-aprendizaje/#:~:text=Un%20algoritmo%20no%20es%20m%C3%A1s,es%20capaz%20de%20hacer%20predicciones>.
- [69] Redacción Tecnósfera, Colombian cases where artificial intelligence makes a difference, EL TIEMPO, (2020), Consultado el: 03-06-2020, Disponible en: <https://www.eltiempo.com/tecnosfera/novedades-tecnologia/casos-de-exito-de-inteligencia-artificial-en-colombia-488062>.

- [70] Rodriguez. Daniel, La regresión logística, Analytics Lane, Consultado el: 18-05-2021, Disponible en: <https://www.analyticslane.com/2018/07/23/la-regresion-logistica/>.
- [71] Roman. Victor, Algoritmos Naive Bayes: Fundamentos e Implementación, Consultado el: 12-05-2021, Disponible en: <https://medium.com/datos-y-ciencia/algoritmos-naive-bayes-fudamentos-e-implementaci%C3%B3n-4bcb24b307f>.
- [72] Sanchez. José, ¿Cómo aprenden las máquinas? Machine Learning y sus diferentes tipos, datos.gob.es, Consultado el: 10-05-2020, Disponible en: <https://datos.gob.es/es/blog/como-aprenden-las-maquinas-machine-learning-y-sus-diferentes-tipos>.
- [73] Sancho Caparrini. Fernando, Métodos combinados de aprendizaje, Disqus, Consultado el: 18-05-2021, Disponible en: <http://www.cs.us.es/~fsancho/?e=106>.
- [74] Silva. Harold, Juliao. David, Ortiz. Mauricio, Martinez. Domingo, González. Jaime, Giraldo. Mario, Marketing: conceptos y aplicaciones, Universidad del Norte, (2014), ISBN: 9789587414943, Disponible en: https://books.google.com.co/books/about/Marketing.html?id=P76QCgAAQBAJ&re_dir_esc=y.
- [75] SimilarWeb, Top Websites Ranking, Consultado el: 06-02-2021, Disponible en: <https://www.similarweb.com/top-websites/>.
- [76] Sitio big data.com, Árbol de decisión en Machine Learning (Parte 1), Sitio big data.com, Consultado el: 11-05-2021, Disponible en: <https://sitiobigdata.com/2019/12/14/arbore-de-decision-en-machine-learning-parte-1/#>.
- [77] Soto. Ximena, Describe el poder del CLIC, Forbes Centro América, Consultado el: 21-02-2020, Disponible en: file:///C:/Users/gyc02178/Downloads/Forbes%20CA%20Agosto2020_new.pdf.
- [78] Stanton. William, Fundamentos de Marketing, McGraw-Hill Interamericana, (2007), ISBN: 9789701062012, Disponible en Biblioteca digital de Bogotá: <https://www.bibliotecadigitaldebogota.gov.co/resources/2215443/>.
- [79] Stelzner. Michael A, Social media marketing industry report - How marketers are using social media to grow their businesses, Social Media Examiner, Pg. 1-48, (2020), Consultado el: 13-02-2021, Disponible en: <https://www.socialmediaexaminer.com/report2020/>.
- [80] Stewart. David, Pavlou. Paul, From Consumer Response to Active Consumer: Measuring the Effectiveness of Interactive Media, Pg. 1-21, (2002), ISSN: 920703, DOI: 10.1177/009207002236912, Disponible en: <https://link.springer.com/article/10.1177/009207002236912>.
- [81] Toca. Claudia, Fundamentos del marketing: guía para su estudio y comprensión, Universidad del Rosario, (2009), ISBN:9789587380576, Disponible en: <https://editorial.urosario.edu.co/fundamentos-del-marketing-guia-para-su-estudio-y-comprension.html>.
- [82] UGPP. la unidad DE PENSIONES Y PARAFISCALES, Operadores y bancos Operadores de información, Consultado el: 11-05-2020, Disponible en:

<https://www.ugpp.gov.co/parafiscales/nuestros-procesos/integracion/operadores-y-bancos>.

[83] Zuccherino. Santiago, Social Media Marketing: LA Revolución de los Negocios y la Comunicación Digital, Temas Grupo Editorial, (2016), ISBN: 9789873887291, Disponible en: <https://www.buscalibre.com.co/libro-social-media-marketing/9789873887291/p/47522429>.

ANEXOS

ANEXO A – Segmentación de la categoría Título Profesional

Titulo_Profesional	
Descripción	No
Sin información	0
Profesional en Salud Ocupacional	1
Ingeniero Electrónico	2
Ingeniero Informático	3
Ingeniero Químico	4
Administrador de Empresas	5
Abogado	6
Psicólogo	7
Ingeniero de Sistemas	8
Arquitecto	9
Ingeniero Civil	10
Contador Público	11
Economista	12
Técnico en procesos contables	13
Mercadeo y Publicidad	14
Ingeniero Eléctrico	15
Ingeniero Electricista	16
Ingeniero Industrial	17
Técnico Disciplina académica asesor comercial y operaciones en entidades financieras	18
Comunicador Social	19
Diseñador	20
Ingeniero Financiero	21
Administrador Turístico y Hotelero	22
Ingeniero Mecánico	23
Ingeniero de Software	24
Administrador en Finanzas y Negocios Internacionales	25
Tecnólogo en Contabilidad	26
Técnico en contabilidad sistematizada	27
Arquitecto Técnico	28
Ingeniero Electrónico y Telecomunicaciones	29
Contador	30
Internacionalista	31
Tecnólogo en Mantenimiento de Equipos de Cómputo, Diseño e Instalación de Cableado Estructurado	32
Química Farmacéutica	33
Profesor de Educación Media	34

ANEXO B – Segmentación de la categoría Título Máximo Nivel Formación

Título_Máximo_Nivel_Formación	
Descripción	No
Sin información	0
Especialista en Ingeniería Ambiental	1
Ingeniero Electrónico	2
Ingeniero Informático	3
Ingeniero Químico	4
Administrador de Empresas	5
Abogado	6
Master en Dirección de Comunicación de Empresas	7
Psicólogo	8
Especialista en Seguridad de la Información e Informática	9
Especialista en Administración Financiera	10
Arquitecto	11
Ingeniero Civil	12
Especialista en Gestión de Riesgos Financieros	13
Profesional en Salud Ocupacional	14
Economista	15
Ingeniero de Sistemas	16
Técnico en procesos contables	17
Mercadeo y Publicidad	18
Ingeniero Eléctrico	19
Maestría en ciencias en ingeniero eléctrico	20
Contador Público	21
Gestión financiera	22
Técnico Disciplina académica asesor comercial y operaciones en entidades financieras	23
Especialista en Derecho Público Disciplina académica	24
Especialista en Auditoría Internacional y Aseguramiento de Información	25
Técnico profesional en procesos comunicativos empresariales	26
Diseñador	27
Ingeniero Financiero	28
Especialista en Gerencia de Mercadeo	29
Maestría en Administración de Empresas	30
Especialista en Gestión Empresarial	31
Maestría en Finanzas	32
Especialista en Finanzas	33
Ingeniero de Software	34
Especialización en Ingeniería de Software	35

Ingeniero Industrial	36
Administrador en Finanzas y Negocios Internacionales	37
Tecnólogo en Contabilidad	38
Especialización en gerencia Financiera	39
Maestría en Planeación Urbana y Rural	40
Master of Science in Business Analytics	41
Maestría en Analítico de Datos	42
Especialista en Alta Gerencia	43
Técnico en contabilidad sistematizada	44
Arquitecto Técnico	45
Especialización en Gerencia de proyectos	46
Especialista en Análisis y Administración Financiera	47
Master en Gerencia de Tecnología de Información	48
Comunicador Social	49
Maestría en Administración - MBA	50
Master of Business Administración	51
Especialista en riesgos laborales, seguridad y salud en el trabajo	52
Contador	53
Master en Administración y Dirección de Empresas	54
Internacionalista	55
Especialización Gerencia en Finanzas	56
Especialización en Publicidad y Mercadeo	57
Tecnólogo en Mantenimiento de Equipos de Cómputo, Diseño e Instalación de Cableado Estructurado	58
Especialización en Administración	59
Especialización Derecho Comercial	60
Ingeniero Electrónico y Telecomunicaciones	61
Química Farmacéutica	62
Doctorado en Ciencias Humanas	63
Ingeniero Mecánico	64

ANEXO C – Segmentación de la categoría Cargos

Cargos	
Descripción	No
Sin información	0
Profesional en Seguridad y Salud en el Trabajo	1
Ingeniero Electrónico	2
Analista de requerimiento y QA	3
Monitor Administrativo	4
Asistente Administrativo	5
Bilingual Customer Service Representative	6
Comunicador Social	7
Auxiliar administrativo	8
Practicante en Seguridad de la información	9
Oficial I -Gestión de Documentos Derivados	10
Asistente Backstage Bcapital	11
Encargada almacén	12
Gerente de Marca Sr.	13
Analista de Calidad	14
Analista salud y seguridad social	15
Analista de Talento Humano	16
Trainee IT	17
Dogger / Alimentos DG S.A.S	18
Auxiliar de recreación	19
Productor	20
Coordinador de proyecto	21
I+D+I (Forschung und Entwicklung)	22
Customer Service Agent	23
Auxiliar de cocina	24
Analista de créditos	25
Especialista, Experto	26
Asesor comercial	27
Operador de tecnología	28
Asistente departamento sistemas	29
Agente de Servicio de Atención al Cliente	30
Asistente jurídico	31
Practicante ingeniería de sistemas	32
Auxiliar de servicio al asociado y administrativa	33
Auditora Interna	34
Comunicador social practicante	35
Auxiliar de Ventas	36

Psicólogo	37
Diseñador	38
Analista de Valoración y Control de Fondos	39
Líder de Investigación Semillero Finanzas Corporativas y Emprendimiento	40
Ejecutivo de Cuenta	41
Analista de cuentas por pagar	42
Ejecutiva de Cuentas Corporativas y Reservas	43
Auxiliar Oficina	44
Digitadora	45
Analista profesional Desarrollo	46
Auxiliar VIII	47
Software Engineer	48
Auxiliar de ingeniería	49
Trabajo de campo	50
Auxiliar en Facturación	51
Aprendiz	52
Auxiliar contable	53
Construction Inspector	54
Credit Card Coordinator	55
Asistente de Cartera	56
Technical Support Specialist	57
TECNICO EN SISTEMAS	58
Administrador de Empresas	59
Secretaria Auxiliar Contable	60
Auxiliar de Arquitectura	61
Analista 1 reporting	62
Analista Herramientas de Gestión	63
Security Engineer	64
Contratista independiente	65
Líder de Planeación Financiera	66
Técnico de laboratorio	67
Cajero bancario	68
Auxiliar de Nuevos Medios	69
Pasante Modalidad Duale	70
Asistente de Gestión Humana y Nomina - Practica Profesional	71
Compras	72
Auxiliar de nóminas de recursos humanos	73
Monitor académico	74
Contable	75
Auxiliar de sistemas (Pasantía)	76

Monitora Académica de Periodismo Internacional	77
Auxiliar Operativo	78
Consultor SAP ABAP HCM	79
Analista de Mercadeo Corporativo	80
Técnico de sistemas	81
Auxiliar de Operaciones Apoyo Comercial y Operativo- Of. Parque Washington	82
Senior IT Analyst	83
Juridico Interno	84
Practicante Universitaria	85
Helpdesk Analyst 1	86
Empleado en prácticas de proyecto	87
Ing. Electrónico	88
Gerente	89
Auxiliar de Despachos	90
Jefa Dpto Práctica Docente del Programa Educación	91
Solution Owner for logistics Process	92
Coordinadora Contrato Interadministrativo	93
Practicante Ingeniero Químico	94
Auxiliar de puestos de pago	95
Profesional de Calidad	96
Investigador en Comunicación DCEI	97
Recepcionista	98
Líder de seguridad Informática	99
Oficial De Operaciones I - Sebra Inclusión SEN	100
Asistente de casting y vestuario Bcapital	101
Agente de servicios al cliente	102
GERENTE DE MERCADEO	103
Líder de Riesgos	104
auxiliar sst y calidad	105
Subgerente de Talento Humano	106
Analista I, Ingeniero Desarrollador de Software	107
Director de marketing	108
Líder Mentor de Ingeniería	109
e-Mobility Business Development	110
Asesor comercial crédito	111
Asistente comercial	112
Asistente técnico testing	113
Gestor y Prevención de Riesgos en Originación	114
Profesional Jurídico Junior	115
Software Developer	116

Auxiliar administrativa y de cartera	117
Analista de Auditoria	118
Analista bienestar y comunicaciones	119
Pasante Delegatura de Riesgo de Mercado y Liquidez	120
Director Creativo	121
Key Account Manager	122
Gerente de finanzas	123
Asesor externo	124
Ingeniera Junior (Analista de terreno)	125
Auxiliar III Transformación del Servicio	126
Bizagi Consultant	127
Asistente contable	128
Mid Senior Developer	129
Coordinadora de surtido	130
Auxiliar de producción	131
Técnico de mantenimiento	132
Coordinadora Administrativa Financiera	133
Ingeniero Residente de Estructuras	134
Sr Manager Research & Data Analytics	135
Asesor educativo	136
Analista de infraestructura TI	137
INGENIERO DE SISTEMAS	138
Analista de producto	139
Consultant	140
Auxiliar de facturación	141
Gerente Financiera, Administrativa y de Talento Humano	142
Ingeniero de preventa - Gerente de Proyectos	143
Auxiliar líder administrativo	144
Profesional de Comunicaciones	145
Director Proyecto de Innovación	146
Analista de Gestión Humana	147
Coordinadora Técnica	148
Generalista de RR. HH.	149
Estudiante tutor	150
Account Manager	151
Gestora de proyecto	152
Subgerente	153
Líder Funcional SAP HCM	154
Organización de Eventos y Ventas	155
Logistics Operations Specialist	156

Gerente Supernumeraria Banca Oficial (Gobierno)	157
Territory Account Manager	158
Analista de Reclutamiento y Selección	159
Eventos	160
Líder de Calidad	161
PSICOLOGA	162
Gerente Comercial	163
Asesora para el Re-diseño Curricular	164
Profesional en SST	165
Ingeniero de investigación	166
Analista de Comunicaciones Internas y Contenido Digital	167
Psicóloga	168
Coordinador - Líder de Seguridad Informática y Ciberseguridad	169
Analista de mercado de capital	170
Luxury Division Marketing Assistant	171
Asesora Servicios	172
Gerente Región Caribe	173
Gerente de Riesgos y Procesos	174
Auxiliar gestión humana	175
Gerencia/Jefe/ Especialista en Talento Humano	176
Analista I. Ingeniero Desarrollador de Software	177
Analista SEM	178
Director Administrativo	179
Senior Consultan	180
Analista III	181
Director	182
Représentante service clients	183
Coodinador de ventas Online	184
Business Analyst Quality Assurance	185
Asesora Comercial	186
Abogada Grupo de Defensa del Consumidor	187
Técnico de cartera contributivo	188
Contralor	189
Locutor y productor	190
Analista de Operaciones	191
Analista sénior	192
Ingeniero graduado	193
UX/UI DESIGNER	194
Profesional Comercial Sector Privado	195
Controlador financiero	196

Analista de servicios administrativos	197
QC Inspector	198
Técnico en Bonos y Cuotas Partes	199
Solutions Sales	200
Senior Software Developer	201
Desarrollador de software	202
LIDER DE ETIQUETADO	203
Técnico de mantenimiento SAC	204
Administración de proyecto	205
Analista de contabilidad	206
Profesional Especializado	207
Líder de Analítica y Marketing Daviplata	208
Asesor Negocios Institucionales Corporativo	209
Technical Support Streaming	210
Scrum Master, lider tecnico, Focal Point, Arquitecto de Datos	211
Auxiliar de Arquitectura y Construcción	212
Líder de Servicio al Cliente y Sena	213
Product Owner	214
Cloud Engineer	215
Coordinador administrativo y de operaciones	216
Director Financiero y Administrativo	217
Profesional Reglas de Negocio	218
Agente servicio al pasajero	219
Coordinador de Marketing Digital	220
Docente / Gestor de Proyectos e Innovación	221
Jefe de Talento Humano	222
Student	223
Profesional Área Académica Posgrados de Profundización de la Escuela de Ingenierías	224
Coordinadora Gestión humana y SST	225
Abogado	226
Gerente Comercial Colombia - Data Quality	227
Gestor de proyectos de nuevos negocios	228
Gerente de zona	229
SAP HCM CONSULTANT	230
Sales Developer	231
Operations Key Account Manager	232
Gerente Supernumeraria	233
Account Executive	234
Profesional Social	235

ADMINISTRADORA	236
Gerente de Proyectos y Calidad	237
Psicoterapeuta en consulta privada	238
GERENTE COMERCIAL Y DE LOGISTICA	239
Inmigrante en España. Comunitaria Italiana	240

ANEXO D – Segmentación de la categoría Empresas

Empresas	
Descripción	No
Sin información	0
AMP INGENIEROS	1
SEDIC S.A.	2
Agencia Nacional Digital	3
Universidad del Valle (CO)	4
Distribuciones Bionatural`s SAS (Actualmente H&N)	5
BILATERAL	6
Codensa SA ESP	7
notaria 64	8
Nexa BPO	9
Banco de Bogotá	10
3cero2	11
Panorama Gold	12
Quala S.A.	13
Soluciones Viales Ballen-Ingeniería	14
Redes SAS	15
Cervecería Nacional	16
GroupM	17
Auxiliar contable	18
Autónomo	19
Oscar Rodriguez producciones	20
Cointelco S.A.	21
Belectric GmbH	22
Amazon	23
Tempotrabajamos	24
Banco Mundo Mujer	25
UkrSibbank	26
Porvenir Pensiones y Cesantías	27
Inversiones dentales Santander	28
Falabella Retail S.A	29
Manufacturas eliot	30
INSTITUTO NACIONAL PENITENCIARIO Y CARCELARIO INPEC	31
Componente Serviex	32
Febor - Fondo empleados Banco de la Republica	33
Corporación para la Recreación Popular	34
FOCUN ORGANIZACIÓN	35
Servi Norte Ltda.	36

APUESTAS CUCUTA 75 PRESENTE EN LA FIRMA DEL PACTO	37
Independiente	38
Skandia Colombia	39
Universidad Autónoma de Bucaramanga	40
Hello Innovation	41
Avianca	42
Toyota Tsusho America	43
Villeta Resort Hotel	44
MONTAJES JM S.A.	45
Finleco Ltda	46
Banco de Occidente	47
ETB	48
Gofive	49
Monomeros Colombo Venezolanos S.A.	50
Creaciones Nurys	51
E.S.E Hospital San Francisco	52
Colpatria Multibanca del Grupo Scotiabank	53
js contadores sas	54
Grupo Réditos	55
Consortio ArconCatapila	56
Davivienda	57
ITAL LENT	58
Teleperformance	59
IDIPRON	60
AUTOPARTES 76	61
Quality Vali S.A.S.	62
Ventas y Servicios S.A.	63
Fluid by Fluidsignal	64
Autoservicio Surtihogar	65
Clínica Sonríe	66
Andina Trim S.A	67
UPS Supply Chain Solutions	68
BBVA	69
Universidad del Norte	70
CienTech - Centro de Transferencia de Conocimiento e Innovación	71
GRICOL S.A.C.	72
Constructora Colpatria	73
Papeleria Mundipapeles	74
CopetranLtda	75
Brigard Urrutia	76

Universidad del Rosario	77
INDUSTRIAS FUNDEQS SAS	78
Wilcos S.A	79
Microcrédito S.A.C.	80
BS Consulting SAS	81
Celeritech	82
Laboratorio Clínico Santa Lucía	83
Banco Caja Social	84
Fundacion Mundo Mujer	85
Caja Promotora de Vivienda Militar y de Policia	86
GRUPO EDITORIAL EDUCAR	87
Getronics	88
Movilidad bogota	89
DTS AUDIO	90
Muebles y trasteos de Colombia	91
Laboratorios Maus SAS	92
PROENFAR S.A	93
UNERMB	94
SabMiller latam	95
ESE HOSPITAL HABACUC CALDERON	96
PROPAL S.A.	97
Grupo Éxito	98
Millenium BPO	99
Grupo de Investigacion en Dirección de Comunicación Empresarial e Institucional	100
Centro Comercial Santafé	101
Mareigua	102
MILLENIUM BPO S.A.	103
PICADITAS	104
Alliance Enterprise SAS	105
Servicios temporales global service	106
Consortio Puerto Limpio	107
Macerada Autónomo	108
Enel X	109
Tuya S.A	110
SMA INGENIERIA LTDA	111
VASS	112
Banco Falabella Colombia	113
Agencia Nacional de Tierras	114
Profesional independiente	115

Cooperativa de ahorro y crédito Crearcoop	116
Quality Brands Company SA QBCo	117
Tys temservice	118
Superintendencia Financiera de Colombia	119
Granados y condecoraciones	120
Simoniz S.A.	121
Tecnologías Teravision	122
Asesor de implementación normas turísticas. Hotel Meson Hondano	123
SK Rental S.A.	124
Aportes en Línea	125
Bizagi	126
JEN S.A	127
MAS Global Consulting	128
Laboratorios Cofarma	129
Droguerías Juliao S.A.	130
CHINOIN Productos Farmacéuticos S.A.	131
AJV Ingenieros SAS	132
GO Kart S.A.s	133
Construcciones obycon sas	134
Jerónimo Martins	135
Grupo Formarte	136
Cencosud Colombia S.A	137
Oxy	138
Impulsa Colombia	139
Vision Software	140
Transporte Seguro y Especializado S.A.S	141
CARGEX	142
Computec Outsourcing S.A.S.	143
Singular Comunicaciones S.A.	144
Uniandinos - Oficial	145
MULTIDIOMAS LTDA	146
GRICOL	147
Universidad Industrial de Santander	148
AUTOTANQUES DE COLOMBIA S.A.S.	149
NeoSecure	150
ACOPI ATLANTIC	151
Bancompartir	152
Grupo Inditex	153
Ajdulce Producciones Autónomo	154
Rayo	155

Forcepoint	156
Banco de las Microfinanzas - Bancamia S.A.	157
EVENTOS PARAISO	158
Bienestar Familiar	159
TESTEC S.A	160
TRANSPORTES JC LOGISTIC	161
Universidad Católica Cecilio Acosta	162
Alcaldía Municipal de Sutatausa	163
Integral Ingenieros Consultores	164
CIAT - Corporación BIOTEC	165
Centro de Educación para el Desarrollo CED	166
Indra Colombia Ltda	167
Fundacion hogares claret	168
Distribebauté	169
Teleperformance Colombia	170
CODESA	171
Imaginatio Agencia de Marketing Digital	172
Esinco	173
HINICIO	174
MEDCO-AVANTAGE L.L.C.	175
Pijamas Alismor	176
Nuvu	177
Superintendencia de Industria y Comercio	178
SaludVida EPS - Oficial	179
UNICATÓLICA CALI	180
La Rfm más música (emisora virtual de USA)	181
Publicar - Publicidad y Soluciones Multimedia S.A	182
Servicios Postales Nacionales	183
CredibanCo	184
FONCEP	185
EcoMi	186
Hexacta	187
Extreme Technologies S.A.	188
Redes Humanas	189
Funsarep	190
Cadena S.A.	191
Superintendencia de Notariado y Registro	192
Banco Davivienda	193
PROTECCIÓN S.A.	194
Ticketcode	195

SERVICIO INTEGRAL MEDICO ASISTENCIAL SAS	196
Globant	197
Dromos Project	198
Nueva EPS	199
LATAM	200
Viva tu Crédito	201
Universidad de la Costa CUC	202
Apoyo y Servicios Temporales S.A.S	203
Holberton School	204
INVERSIONES NEVADA LTDA	205
Soler&Mejía	206
DEYDE Colombia Calidad	207
COMERCIALIZADORA M&M	208
Multipagas	209
EPI-USE	210
Interfell	211
Dynatrace	212
Cooperativa Multiactiva Rapiaseos de Colombia	213
Alinatech SAS	214
Psicólogo Independiente	215
TRANSOLCAR S.A.S	216
Homologando Título de Doctor	217