



EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA



Instituto Tecnológico de Cuautla

Tecnológico Nacional de México
Instituto Tecnológico de Cuautla
Coordinación de Posgrado

“Machine learning aplicado a la mercadotecnia digital”

Alan Galván Espinoza

Tesis para Obtener el grado de
Maestro en Ingeniería Administrativa

Comité Tutorial

M.C. Silvestre Guillermo Puebla Serrano

DRA. Armida González Lorence

Directores de Tesis:

M.C. Julio Pérez Machorro

M.A. Leonor Angeles Hernández



Cuautla, Morelos

Diciembre. 2023



EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA



INSTITUTO TECNOLÓGICO
NACIONAL DE MÉXICO

Instituto Tecnológico de Cautla
Subdirección Académica
Coordinación de Posgrado

H.H. Cautla, Morelos, **05/diciembre/2023**
OFICIO: ITC/SA/CPI/169/2023

ASUNTO: Autorización de impresión

**C. ALAN GALVÁN ESPINOZA
PRESENTE**

Me es grato comunicarle que después de la revisión realizada por el Jurado designado para su examen de grado de Maestría en Ingeniería Administrativa, se acordó autorizar la impresión de su trabajo de Tesis titulado:

"MACHINE LEARNING APLICADO A LA MERCADOTECNIA DIGITAL"

Es muy satisfactorio para la Coordinación de Posgrado compartir con usted el logro de esta meta.

Espero que continúe con éxito su desarrollo profesional y dedique su experiencia e inteligencia en beneficio de nuestro país.

ATENTAMENTE

"En su ideal de Servicio reside la grandeza del hombre"

**M.C. GISELA VEGA TORRES
SUBDIRECTORA ACADÉMICA**

Ccp: archivo
GVT/MLAM/ebt*



Libramiento Cautla - Oaxaca, S/N, Col. Juan Morales, C. P. 62826, Yecapixtla, Morelos,
Tels. (735) 12-2-22-42 y 35-3-64-96 email: posgrado@cautla.tecnm.mx
tecnm.mx | www.cautla.tecnm.mx

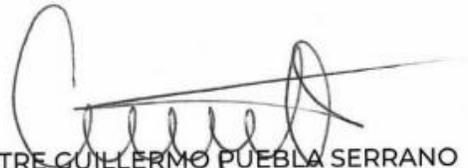


H.H. Cuautla, Morelos, **23/noviembre/2023**
OFICIO: ITC-SA-CPI/159/2023

El presente trabajo de Tesis **"MACHINE LEARNING APLICADO A LA MERCADOTECNIA DIGITAL"**, presentado por el **ING. ALAN GALVÁN ESPINOZA**, con **número de control M20680010**, ha sido aprobado por su comité Tutorial como requisito parcial para obtener el grado de:

MAESTRO EN INGENIERÍA ADMINISTRATIVA

Director de Tesis



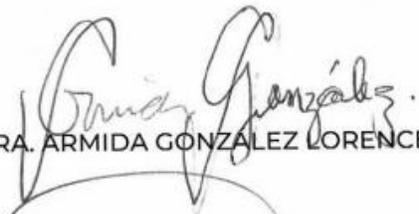
M.C. SILVESTRE GUILLERMO PUEBLA SERRANO

Secretario



M.C. JULIO PÉREZ MACHORRO

Vocal



DRA. ARMIDA GONZÁLEZ LORENCE

Vocal Suplente



M.A. LEONOR ANGELES HERNÁNDEZ



LICENCIA DE USO OTORGADA POR Alan Galván Espinoza, de nacionalidad Mexicana mayor de edad, con domicilio ubicado en Calle Engrane N°202 Ext. 2 Col. La rueda, San Juan del Rio Querétaro, en mi calidad de titular de los derechos patrimoniales y morales y autor de la tesis denominada "Machine learning aplicado a la mercadotecnia digital" en adelante "**LA OBRA**" quien para todos los fines del presente documento se denominará "**EL AUTOR Y/O TITULAR**", a favor del Instituto Tecnológico de Cuautla del Tecnológico Nacional de México, la cual se registrará por las cláusulas siguientes:

PRIMERA – OBJETO: "**EL AUTOR Y/O TITULAR**", mediante el presente documento otorga al Instituto Tecnológico de Cuautla del Tecnológico Nacional de México, licencia de uso gratuita e indefinida respecto de "**LA OBRA**", para almacenar, preservar, publicar, reproducir y/o divulgar la misma, con fines académicos, por cualquier medio en forma física y a través del repositorio institucional y del repositorio nacional, éste último consultable en la página: (<https://www.repositorionacionalcti.mx/>).

SEGUNDA - TERRITORIO: La presente licencia se otorga, de manera no exclusiva, sin limitación geográfica o territorial alguna, de manera gratuita e indefinida.

TERCERA -ALCANCE: La presente licencia contempla la autorización para formato uso de "**LA OBRA**" en cualquier formato o soporte material y se extiende a la utilización, de manera enunciativa más no limitativa a los siguientes medios: óptico, magnético, electrónico, virtual (red), mensaje de datos o similar conocido por conocerse.

CUARTA – EXCLUSIVIDAD: La presente licencia no implica exclusividad en favor del Instituto Tecnológico de Cuautla; por lo tanto, "**EL AUTOR Y/O TITULAR**" conserva los derechos patrimoniales y morales de "**LA OBRA**", objeto del presente documento.

QUINTA – CRÉDITOS: El Instituto Tecnológico de Cuautla y/o el Tecnológico Nacional de México reconoce que el "**AUTOR Y/O TITULAR**" es el único, primigenio y perpetuo titular de los derechos morales sobre "**LA OBRA**"; por lo tanto, siempre deberá otorgarle los créditos correspondientes por la autoría de la misma.

SEXTA – AUTORÍA: "**EL AUTOR Y/O TITULAR**" manifiesta ser el único titular de los derechos de autor que derivan de "**LA OBRA**" y declara que el material objeto del presente fue realizado por él, sin violentar o usurpar derechos de propiedad intelectual de terceros; por lo tanto, en caso de controversia sobre los mismos, se obliga a ser el único responsable. Dado en la Ciudad de San Juan del Rio Oro., a los 24 días del mes de noviembre de 2023.



Ing. Alan Galván Espinoza

"EL AUTOR Y/O TITULAR"

DEDICATORIA

Esta tesis está dedicada a mi familiar más cercano, a mi padre, quien siempre ha sido mi guía y apoyo incondicional en mi camino académico. Gracias por inculcarme valores como la perseverancia, el esfuerzo y la pasión por el conocimiento.

También quiero dedicar este logro a mis amigos cercanos que conocí y quienes han estado conmigo en cada paso del camino, brindándome su amistad, su ánimo y su incondicional apoyo que ahora considero parte de mi familia. Su confianza en mí ha sido fundamental para alcanzar este importante hito en mi vida.

Espero que este trabajo sea una muestra de mi gratitud y un testimonio en mi camino profesional. ¡Gracias por siempre estar ahí!

AGRADECIMIENTOS INSTITUCIONALES

Quiero expresar mi más sincero agradecimiento a las personas que me han brindado su apoyo y colaboración durante la realización de esta investigación.

Agradezco profundamente a todas las personas que han contribuido en mi profesional, por su compromiso, dedicación y valiosa retroalimentación. En especial, agradezco a los docentes y a mis asesores, por su orientación, guía y observaciones.

A Isabel Taracena, por darme la confianza y permitirme colaborar junto a ella durante todo este proyecto, sin ella este trabajo no hubiera sido posible.

No puedo dejar de mencionar a mi padre, quien me ha brindado un apoyo incondicional, ha sido una gran fuente de motivación y aliento en todo momento.

Finalmente, quiero agradecer a mis compañeros de maestría, quienes han sido una parte fundamental en mi formación profesional-personal, y con quienes he compartido grandes experiencias y aprendizajes.

A todos ellos, mi más sincero agradecimiento por su apoyo, paciencia y colaboración durante todo este desarrollo profesional.

ÍNDICE DE CONTENIDO

ÍNDICE DE CONTENIDO.....	7
ÍNDICE DE IMÁGENES, GRAFICOS Y CUADROS.....	12
LISTADO DE SIGLAS.....	17
ÍNDICE DE ANEXOS.....	19
RESUMEN.....	20
INTRODUCCIÓN.....	22
CAPITULO I: ANTECEDENTES (o Estado del Arte).....	24
1.1 Artículos.....	26
1.1.1 How causal machine learning can leverage marketing strategies: Assessing and improving the performance of a coupon campaign.....	26
1.1.2 Tecnologías en la inteligencia artificial para el Marketing: una revisión de la literatura.....	27
1.1.3 Neural Insights for Digital Marketing Content Design.....	27
1.1.4 Análisis y propuesta para el aprovechamiento de herramientas de machine learning en el sector del marketing digital en Netbangers.....	29
1.1.5 Machine learning en marketing: ventajas y ejemplos de aplicación.....	29
1.2 El machine learning y sus aplicaciones de contenido digital en la industria 4.0.....	31
1.2.1 Impactos del uso de machine learning en e-commerce.....	34
1.3 Casos aplicados de Machine learning en el marketing.....	36
1.3.1 Contenido publicitario de JPMorgan generado automáticamente.....	36
1.3.2 Recomendación automática de Netflix.....	37
1.3.3 Algoritmo de uso libre de Airbnb.....	38

1.3.4 Servicio al cliente automatizado por chat	39
CAPÍTULO II. MARCO HISTÓRICO - CONTEXTUAL.....	40
CAPITULO III. PLANTEAMIENTO DEL PROBLEMA	42
3.1 Justificación	44
3.2 Objetivo General.....	48
3.2.1 Pregunta Central de investigación	48
3.3 Variables de investigación	48
3.3.1 Variables para el análisis de geomarketing.....	49
3.4 Hipótesis central de investigación.....	50
3.5 Objetivos específicos	51
3.5.1 Preguntas secundarias de Investigación	53
3.5.2 Supuestos Teóricos secundarios de investigación	54
3.6 Alcances y Limitaciones	56
CAPITULO IV. MARCO TEÓRICO CONCEPTUAL.....	58
4.1 Marketing Digital	58
4.2 Posicionamiento en buscadores.....	60
4.2.1 SEO.....	61
4.2.2 Search quality rating.....	63
4.2.3 El uso de hashtags para posicionamiento en redes sociales.....	64
4.3 Big data	65
4.3.1 Plataformas de recolección y análisis de datos.....	66
4.3.2 Azure y Power BI de Microsoft	67
4.3.3 AWS: Amazon Web Services.....	69

4.3.4 Google analytics.....	70
4.4 Inteligencia artificial	72
4.4.1 Algoritmo de aprendizaje automático.....	73
4.4.2 Algoritmo de redes neuronales	73
4.4.3 Algoritmo de razonamiento lógico.....	74
4.4.4 Algoritmo de optimización	75
4.5 Técnicas de machine learning	75
4.5.1 Aprendizaje supervisado	76
4.5.2 Aprendizaje no supervisado	77
4.5.3 Árbol de Decisiones	77
4.5.4 Regresión	78
4.5.5 Redes neuronales	80
4.5.6 Algoritmo (K-NN) o vecino más cercano	81
4.5.7 Support Vector machines	83
4.5.8 Modelo bayesiano	83
4.5.9 Bosque aleatorio.....	84
4.5.10 Cluster analysis	85
4.5.11 Estimación de densidad Kernel	85
4.6 Aplicación de algoritmos en la predicción de comportamiento de usuarios.....	86
4.7 El proceso del análisis de datos	88
4.8 Funciones de activación	89
CAPÍTULO V. METODOLOGÍA	92
5.1 Tipo de Estudio	93

5.1.1	Diseño de Investigación.....	94
5.2	Sujetos de Estudio o población.....	95
5.3	Muestra	96
5.3.1	Tamaño de muestra en el estudio de geomarketing.	96
5.4	Instrumento de recolección de datos	98
5.5	Instrumentos de análisis de datos.....	98
5.6	Técnicas de machine learning utilizadas	102
CAPÍTULO VI: ANÁLISIS DE DATOS.....		104
6.1	Métricas de la página web (Estudio fotográfico Isabel Taracena) durante la investigación	105
6.1.1	Medición del engagement en las publicaciones en redes sociales.....	108
6.2	Exploración y estudio de mercado aplicado por geomarketing.....	111
6.3	Selección automática de imágenes y correlación de engagement	114
6.3.1	Conclusión de la clasificación automática del tipo de fotografía.....	119
6.3.2	Conclusión de la correlación.....	119
6.3.2	Clasificación (Hierarchical Clustering) agrupadas por etiquetas y el orden automático	120
6.4	Evaluación de diferentes modelos de machine learning en predicción de engagement .	123
6.5	Análisis de la página web con los datos de google analytics.....	128
6.5.1	Análisis por series de tiempo (Adquisición de usuarios y tiempo de interacción)	129
6.5.2	Evaluación de predicciones de múltiples modelos de machine learning	131
6.5.3	Histograma de acciones más realizadas dentro de la página web, secciones más visitadas y países que visualizaron la página web	133
6.6	Análisis de texto en SKAGS y hashtags	136

6.6.1 Análisis con Google trends	139
6.7 Estrategias consideradas e implementaciones realizadas en la campaña de marketing .	142
6.7.1 Estrategia AIDA.....	143
6.7.2 Atracción de público a la página web.....	144
6.7.3 Retención de atención al contenido.....	145
6.7.4 Llamado a la acción	147
6.8 Resumen y comprobación de hipótesis.....	149
CAPÍTULO VII: CONCLUSIONES Y RECOMENDACIONES	150
7.1 Conclusiones generales de los resultados	151
7.2 Recomendaciones generales	151
7.3 Conclusión sobre el proceso realizado	152
CAPÍTULO VIII: PROPUESTA	153
8.1 Propuesta por objetivos específicos.....	154
8.2 Explicación	158
8.3 Fechas	159
8.4 Presupuesto	160
8.5 Cierre de la propuesta	160
8.6 Reflexiones finales	162
BIBLIOGRAFÍA	163
ANEXOS.....	170

ÍNDICE DE IMÁGENES, GRAFICOS Y CUADROS

Figura 1. Gráfica del crecimiento porcentual del Producto Interno Bruto en el comercio electrónico entre 2013 al 2019 en México.....	60
Figura 2. Esquema de un árbol de decisiones, de una problemática se plantean las diferentes decisiones y los diferentes beneficios probables de cada uno.	78
Figura 3. Regresión lineal con una variable dependiente y una variable independiente.	79
Figura 4. Representación de una distribución de nodos en redes neuronales.....	81
Figura 5. Representación gráfica del funcionamiento del modelo K-NN.....	82
Figura 6. Esquema del procesamiento de datos y modelado enfocado en el uso de machine learning.	88
Figura 7. Proceso de extracción de datos, procesado de ETL y carga de la base de datos a servidores para procesarlos (Data warehouse).....	89
Figura 8. Graficación de las principales funciones activación.	91
Figura 9. Captura de la portada principal de la página web.....	92
Figura 10. Captura de las redes sociales a analizar (Instagram)	93
Figura 11. Etapas del procesamiento de datos para su analisis.....	101
Figura 12. Cantidad de visitantes por sección de la página web.....	106
Figura 13. Gráfica de frecuencia por visitas de los usuarios.	106
Figura 14. Resumen por seccion de la pagina web	107
Figura 15. Canales frecuentes donde provienen las visistas.....	107
Figura 16. Resumen de los canales donde provienen los usuarios.....	108
Figura 17. Resumen de las estadísticas generadas en instagram	110
Figura 18. Registro de instalaciones código 711510 DENUE (artistas, escritores y técnicos independientes).....	113

Figura 19. Ubicación de Estudios fotográficos presentes en el directorio de google maps.	114
Figura 20. Proceso de análisis modelado en Orange.	115
Figura 21. Matriz de Distance Map por cada fotografía.	116
Figura 22. Selección de grupos y visualización de imagen seleccionadas.....	117
Figura 23. Primera parte del reporte de correlaciones entre engagement, el orden automático y tipo de fotografías	118
Figura 24. Segunda parte del reporte de correlaciones entre engagement, el orden automático y tipo de fotografías.	118
Figura 25. Resultado de la correlación Pearson y valor p.	119
Figura 26. Flujo de datos para su clasificación hierarchical clustering.	120
Figura 27. Visualización en forma de matriz de la distancia (diferencias) entre variables.....	121
Figura 28. Visualización de grupos que se generaron por similitud.....	122
Figura 29. Visualización del grupo seleccionado en las agrupaciones automáticas.	122
Figura 30. Flujo de datos para la evaluación y predicción de los múltiples modelos de machine learning.....	124
Figura 31. Pruebas por cada modelo en cada variable y matriz de resultados con sus diferencias.	125
Figura 32. Revisión a detalle de las diferencias entre modelos.	126
Figura 33. Grafica de segmentación por tipo de publicación.....	126
Figura 34. Grafica de Regresión lineal de las categorías fotográficas estudiadas.	127
Figura 35. Flujo de datos para el análisis de adquisición e interacciones de los usuarios.....	129
Figura 36. Grafica se series de tiempo en intervalos de semanas sobre usuarios adquiridos en 2022.	130

Figura 37. Grafica se series de tiempo en intervalos de semanas sobre tiempo de interacción realizado en 2022.	131
Figura 38. Resumen de diferencias entre modelos de machine learning por los datos de google analytics.	132
Figura 39. Resumen gráfico de las variables generadas por los modelos y sus promedios.	133
Figura 40. Histograma de las acciones realizadas en la página web.	134
Figura 41. Histograma de las secciones más visitadas en la página web.	134
Figura 42. Histograma de los países donde fue visualizada la página web.	135
Figura 43. Nube de palabras basado en los hashtags utilizados, de las más generales (palabras grandes) a las más específicas (palabras pequeñas) como etiqueta.	136
Figura 44. Aplicación del modelo de tópicos por cada palabra ingresada de los hashtags.	137
Figura 45. Lista de términos desplegados por SKAGS.	137
Figura 46. Visualización de metadatos por combinación de SKAGS.	138
Figura 47. Visualización de metadatos por hashtags.	138
Figura 48. Términos seleccionados entre hashtags y SKAGS en Google trends.	139
Figura 49. Visualización de series de tiempo sobre los términos ingresados (sin la etiqueta cdmx)	140
Figura 50. Visualización de series de tiempo sobre los términos ingresados (con la etiqueta cdmx)	140
Figura 51. Búsqueda a nivel nacional basado en SKAGS más buscados. (Primera parte).....	141
Figura 52. Búsqueda a nivel nacional basado en SKAGS más buscados. (Segunda parte)	141
Figura 53. Representación de un embudo de ventas.....	143
Figura 54. Modelo AIDA.....	144
Figura 55. Sección de la página más visitada en el análisis.	146

Figura 56. Reestructuración del contenido de la página web.....	146
Figura 57. Segunda parte de la reestructuración de la página web.....	147
Figura 58 Apartado de contacto en la pagina web.....	148

ÍNDICE DE TABLAS

Tabla 1. Comparativa entre los años y el incremento de aportación de valor agregado bruto en el comercio electrónico en México	59
Tabla 2. Calculo del engagement de las 8 publicaciones mas populares en instagram de Isabel Taracena.	111
Tabla 3 Usuarios de la pagina web Isabel Taracena por semana, en la semana 3 donde se realizo la primera publicacion en redes sociales con el enlace a la pagina web.	170
Tabla 4 Directorio de empresas de fotografias y sus ubicaciones en ciudad de México que se empleo para el estudio de geomarketing.	171

LISTADO DE SIGLAS

Siglas orden alfabético	Significado
Accuracy	Exactitud
AIDA	Atención, Interés, Deseo y Acción
ARIMA	AutoRegressive Integrated Moving Average
AUC	Area Under the Curve / Área bajo la curva
AWS	Amazon Web Services
B.D.	Big data
CSV	comma-separated values
DENUE	Directorio Estadístico Nacional de Unidades Económicas
ETL	Extract, Transform and Load / Extracción, transformación y carga
IaaS	Infrastructure as a service
IBM	International Business Machines
INEGI	Instituto Nacional de Estadística y Geografía
kNN	(valor objetivo definido como k) K-nearest neighbors

M.L.	Machine learning
MIT	Massachusetts Institute of Technology
PaaS	Platform as a service
Puntuación F1	Valor-F
PwC	(Consultora) PricewaterhouseCoopers
Recall	Exhaustividad
ROI	Retorno de la inversión / Return On Investment
SaaS	Software as a service
SEM	Search Engine Marketing
SEO	Search Engine Optimization
SERP	Search Engine Results Page
SKAGS	Single Keyword Ad Groups
SQL	Structured Query Language
SQR	Search quality rating
SVM	Support vector machines
VAR	Modelos vectoriales autoregresivos

ÍNDICE DE ANEXOS

Tabla 3 Usuarios de la pagina web Isabel Taracena por semana, en la semana 3 donde se realizo la primera publicacion en redes sociales con el enlace a la pagina web.

Tabla 4 Directorio de empresas de fotografias y sus ubicaciones en ciudad de Méxic

RESUMEN

El presente estudio se enfocó en mejorar las estrategias de marketing digital de la microempresa de servicios fotográficos profesionales "Estudio Isabel Taracena" ubicada en Ciudad de México durante el año 2022. El problema a abordar radicaba en las dificultades que enfrentaba la microempresa para alcanzar una mayor visibilidad y alcance en el mercado digital, lo que afectaba su competitividad y crecimiento.

El objetivo principal de esta tesis es mejorar las estrategias de marketing digital de la microempresa de servicios fotográficos profesionales Isabel Taracena mediante el análisis de datos y técnicas de machine learning. Tanto en su página web como su presencia en redes sociales, y se recopilaron datos relevantes sobre el tráfico del sitio, el alcance y el compromiso en redes sociales, tasas de conversión, entre otros.

Con el fin de solucionar este desafío, se realizó una investigación que empleó diversas técnicas y metodologías de análisis de datos, combinadas con el uso de machine learning. Los objetivos del estudio fueron identificar áreas de mejora en las estrategias de marketing digital, proponer recomendaciones específicas para optimizar dichas estrategias y proporcionar una base teórica sólida para comprender la aplicación del machine learning en el marketing digital.

Los resultados del análisis de datos proporcionaron una visión profunda del rendimiento de las estrategias de marketing digital de la microempresa. Se encontraron áreas de mejora, como la optimización del posicionamiento en motores de búsqueda y el contenido en redes sociales, la implementación de estrategias de engagement y la selección adecuada de imágenes para aumentar la interacción con el público.

Como resultado de la investigación, se presentaron recomendaciones concretas para mejorar las estrategias de marketing digital de "Estudio Isabel Taracena". Estas recomendaciones incluyeron la aplicación de técnicas avanzadas de machine learning para personalizar el contenido y las ofertas, la segmentación efectiva del público objetivo y la utilización de análisis de palabras clave y hashtags para optimizar el contenido en redes sociales.

A pesar de los resultados obtenidos, aún queda trabajo por hacer. La implementación práctica de las estrategias propuestas requerirá un seguimiento y análisis continuo para evaluar su impacto en el crecimiento y éxito de la microempresa en el mercado digital. Además, se abre el camino para futuras investigaciones que exploren otras estrategias en el posicionamiento de contenido y otras metodologías para un análisis más profundo y efectivo en el ámbito del marketing digital. La investigación busca contribuir al campo del marketing digital y servir como referencia para otras microempresas interesadas en mejorar su presencia en línea y alcanzar un público más amplio.

Palabras Clave: *Marketing digital, análisis de datos, machine learning, Open source, engagement.*

INTRODUCCIÓN

La presente tesis se centra en la exploración de diversas técnicas y metodologías de análisis de datos con el objetivo de mejorar la toma de decisiones en el ámbito del marketing digital, específicamente en una microempresa de servicios fotográficos profesionales ubicada en Ciudad de México durante el año 2022. El contenido de esta investigación se organiza en ocho capítulos, cada uno de los cuales aborda aspectos clave relacionados con el tema.

Para el desarrollo del presente trabajo se describen los antecedentes y el contexto del estudio realizado en la microempresa en Ciudad de México. Luego se presenta el planteamiento del problema, incluyendo su justificación, objetivos, variables e hipótesis de investigación, así como los alcances y limitaciones del estudio.

El Capítulo I constituye el punto de partida, al analizar los antecedentes relevantes en el campo del marketing digital. Aquí se exploran temas como SEO (optimización para motores de búsqueda), Search Quality Rating (evaluación de calidad de búsqueda), plataformas de análisis de datos, aplicaciones del machine learning y casos prácticos aplicados al marketing.

El Capítulo II se dedica a describir detalladamente el contexto en el que se lleva a cabo el estudio, enfocándose en la microempresa de servicios fotográficos profesionales objeto de investigación. Se presentan aspectos relevantes, como el tamaño de la empresa, su ubicación geográfica y su público objetivo.

El Capítulo III se centra en el Planteamiento del Problema, donde se exponen las motivaciones que impulsaron la investigación, en particular la necesidad de mejorar las estrategias de marketing digital de la microempresa. Aquí se plantea claramente cuál es el desafío principal que se busca abordar.

En el Capítulo IV, se ahonda en el marco teórico contextual del estudio, explorando en profundidad conceptos clave como el marketing digital, el posicionamiento en buscadores, el big data, la inteligencia artificial y las técnicas de machine learning. Este capítulo sienta las bases teóricas necesarias para comprender y contextualizar la investigación.

El Capítulo V se adentra en la metodología empleada durante el estudio. Aquí se detallan aspectos como el tipo de estudio realizado, el sujeto de estudio seleccionado, la muestra

utilizada, el instrumento de recolección de datos empleado y las técnicas específicas de análisis de datos aplicadas en la investigación.

El Capítulo VI se enfoca en los resultados obtenidos del análisis de los datos recopilados. Se examinan diferentes métricas utilizadas durante la investigación, como el estudio de mercado aplicado al geomarketing, la selección automática de imágenes y su correlación con el engagement, la evaluación de modelos de machine learning, el análisis de palabras SKAGS y hashtags, así como las aplicaciones de estrategias de marketing.

El Capítulo VII presenta las conclusiones derivadas del estudio y proporciona recomendaciones concretas para mejorar las estrategias de marketing digital de la microempresa. Aquí se resumen los hallazgos clave y se proponen acciones específicas basadas en los resultados obtenidos.

Finalmente, en el Capítulo VIII se presentan propuestas concretas para mejorar las estrategias de marketing digital de la microempresa en el futuro, tomando como punto de partida los datos analizados y los conocimientos adquiridos a lo largo del estudio.

Para concluir, esta tesis busca cumplir su objetivo principal: mejorar las estrategias de marketing digital de una microempresa de servicios fotográficos profesionales mediante el análisis de datos y el empleo de técnicas de machine learning. Cada capítulo se enfoca en aspectos específicos de la investigación, desde los antecedentes hasta las propuestas para el futuro.

Se espera que esta investigación contribuya al campo del marketing digital al proporcionar una comprensión más profunda de cómo las técnicas de machine learning pueden ser utilizadas para mejorar el rendimiento de las empresas en línea. Además, se espera que los resultados de esta investigación puedan ser útiles para otras microempresas que buscan mejorar su presencia en línea y atraer a un público más amplio.

CAPITULO I: ANTECEDENTES (o Estado del Arte)

A lo largo de esta última década la perspectiva que se tenía de los datos se ha convertido en uno de los modelos de negocios más rentables para las empresas más influyentes en la actualidad. Entre las empresas líderes más rentables del mundo se encuentran Apple, Microsoft, Google, Facebook, IBM y SAP (statista, 2021), todos tienen en común sus operaciones requieren un manejo y análisis de datos que producen sus clientes para realizar un seguimiento.

En palabras de Fausto Ibarra, director de gestión de productos de Google cloud platform “El aprendizaje automático es básicamente una forma en que una computadora busca los recursos de información que un ser humano no puede encontrar”. “Una vez que tiene los datos y la capacitación, e implementa sus modelos, la máquina puede analizar terabytes de datos y hacerse más y más inteligente, básicamente la máquina se entrena, y en última instancia hace predicciones para usted” (Ibarra, 2021).

El aprendizaje de máquinas o también conocido como machine learning ha tenido un impacto y crea en la manera de cómo se trabaja con la información con el fin de obtener más competitividad en el mercado, de acuerdo a una encuesta realizada por el MIT technology review en colaboración con google cloud a finales del 2016-2017 el 45% de las empresas encuestadas que han implementado técnicas de machine learning en sus procesos han alcanzado sus objetivos planteados e inclusive han tenido un retorno de inversión demostrable temprano (MIT Technology Review, 2017).

Dentro de las implementaciones más comunes que tiene el machine learning se encuentra el reconocimiento, clasificación y etiquetado de imágenes, análisis de comportamiento de usuarios y clasificación con minería de texto al igual que el lenguaje natural (Jonathan Benchimol, 2021).

En la encuesta realizada por MIT (MIT Technology Review, 2017) realizó con una muestra de 375 representantes de empresas, a más de 25 países, con una mayoría del sector de tecnologías de la información, servicios empresariales y financieros.

Según los resultados de la encuesta realizada por el MIT Technology Review, que abordó el tema del uso actual del aprendizaje automático en las organizaciones, se identificaron los objetivos buscados al implementar esta tecnología. Entre las respuestas obtenidas, se destacan los siguientes puntos clave:

En primer lugar, el 50% de las empresas encuestadas mencionó que su objetivo principal era obtener un análisis de datos más amplio y acceder a una mayor cantidad de información. Esto refleja la necesidad de utilizar el aprendizaje automático para procesar grandes volúmenes de datos y extraer conocimientos valiosos.

En segundo lugar, un 46% de las organizaciones indicó que buscaba obtener una ventaja competitiva a través del aprendizaje automático. Esto sugiere que reconocen el potencial de esta tecnología para mejorar su posición en el mercado y destacarse frente a la competencia.

Asimismo, un 45% de las empresas manifestó su interés en lograr un análisis de datos más rápido y obtener conocimientos de manera ágil. Esto pone de manifiesto la necesidad de acelerar el procesamiento de datos y reducir los tiempos requeridos para obtener información relevante.

Además, un 44% de las empresas mencionó que buscaba mejorar sus capacidades de investigación y desarrollo mediante el uso del aprendizaje automático, con el objetivo de impulsar la innovación y desarrollar productos de próxima generación.

En el ámbito interno, un 39% de las organizaciones expresó su interés en mejorar la eficiencia de sus procesos operativos a través del aprendizaje automático. Esto demuestra la necesidad de optimizar las operaciones empresariales y reducir los costos asociados.

Por otro lado, un 35% de las empresas mencionó que deseaba obtener una mejor comprensión de sus clientes actuales y potenciales mediante el aprendizaje automático. Esto refleja la importancia de utilizar esta tecnología para analizar y comprender el comportamiento de los clientes, adaptando así las estrategias de marketing y ventas de manera más efectiva.

En términos de reducción de costos, un 31% de las organizaciones consideraba que el aprendizaje automático era una herramienta útil para optimizar los recursos y lograr ahorros en diferentes áreas.

Finalmente, un 26% de las empresas mencionó que buscaba mejorar la eficiencia de los procesos externos, como la cadena de valor, a través del uso del aprendizaje automático. Esto implica utilizar esta tecnología para optimizar las relaciones con proveedores, socios comerciales y otros actores externos.

1.1 Artículos

En la era digital actual, la aplicación de tecnologías avanzadas está transformando la manera en que las empresas abordan sus estrategias de marketing. La convergencia entre el mundo del marketing y la inteligencia artificial ha dado lugar a una revolución en la toma de decisiones y en la comprensión de los consumidores. Uno de los enfoques más notables es la integración del machine learning en el marketing digital. A medida que las organizaciones buscan aprovechar al máximo los datos disponibles y personalizar las experiencias de los clientes, se han explorado diversas formas de implementar el machine learning para obtener ventajas competitivas. En este subtema de antecedentes, se analizarán una serie de artículos relevantes que detallan cómo se ha llevado a cabo esta integración en la práctica, explorando sus éxitos, desafíos y oportunidades. A través de este análisis, se pretende obtener una comprensión más profunda de las estrategias que han demostrado ser efectivas y cómo se han adaptado a diversos contextos empresariales.

1.1.1 How causal machine learning can leverage marketing strategies: Assessing and improving the performance of a coupon campaign

El artículo presenta diferentes métodos de aprendizaje automático causal aplicados a la investigación de marketing y el desarrollo empresarial. Estos métodos se utilizan para evaluar las campañas de cupones de un minorista y optimizar la distribución de los cupones. Los resultados muestran que ciertos tipos de cupones tienen un efecto positivo y significativo en las compras, mientras que otros tipos de cupones pueden reducir el gasto diario de los clientes. Estos métodos también permiten identificar subgrupos de clientes cuyo comportamiento de compra puede ser influenciado de manera más efectiva a través de ciertos tipos de cupones.

El artículo utiliza métodos de aprendizaje automático causal para evaluar el impacto de las campañas de cupones en el comportamiento de compra de los clientes. Específicamente, se utiliza el enfoque de bosques causales desarrollado por el equipo de investigadores.

Este enfoque permite estimar el efecto causal de los cupones al controlar por variables observables que influyen tanto en la recepción de los cupones como en el comportamiento de compra de los clientes. Se aplican modelos de aprendizaje automático para predecir el comportamiento de compra y la recepción de cupones, y luego se integran estas predicciones

en la estimación de los efectos del tratamiento. Los resultados obtenidos permiten identificar los efectos promedio de los cupones en todos los clientes, así como en subgrupos específicos.

En general, estos métodos permiten identificar el efecto promedio del tratamiento, la heterogeneidad de los efectos y las reglas de distribución óptimas de los cupones. El enfoque propuesto puede aplicarse a otras estrategias de marketing y negocios utilizando datos estructurados o no estructurados. (Huber, 2022)

1.1.2 Tecnologías en la inteligencia artificial para el Marketing: una revisión de la literatura

El artículo se trata de una revisión de la literatura sobre las tecnologías de inteligencia artificial aplicadas al marketing. El objetivo del estudio es analizar las contribuciones teóricas, el surgimiento, desarrollo, tendencias, perspectivas, componentes y las contribuciones de estas tecnologías en el campo del marketing. Se evaluaron documentos en español e inglés utilizando bases de datos como Google Académico y Microsoft Academic Search. Se destacan tecnologías como el aprendizaje automático y el procesamiento del lenguaje natural, que pueden potenciar la fidelización de los mercados seleccionados.

El autor presenta una revisión de la literatura sobre las tecnologías de inteligencia artificial (IA) aplicadas al marketing. El objetivo del estudio es proporcionar información relevante tanto para la comunidad científica como para las empresas en la toma de decisiones estratégicas para la publicidad orientada a la teoría histórica como A lo largo de las décadas siguientes, surgieron programas y sistemas que demostraron avances en áreas como reconocimiento de lenguaje natural, resolución de problemas matemáticos y ajedrez.

La creación del lenguaje de programación Lisp y la aparición de sistemas como DENDRAL, capaz de realizar análisis químicos, marcaron hitos en la aplicación práctica de la IA. La década de 1980 vio la introducción de sistemas expertos como R1, que demostraron su valía en la resolución de problemas empresariales y reducción de costos. (Chicaiza, 2020)

1.1.3 Neural Insights for Digital Marketing Content Design

El artículo "Neural Insights for Digital Marketing Content Design" presenta un enfoque basado en redes neuronales para mejorar la eficiencia y efectividad del diseño de contenido de marketing.

El objetivo es extraer información accionable de modelos basados en datos para mejorar el proceso de creación de contenido. El enfoque utiliza técnicas de aprendizaje automático y modelos de redes neuronales profundos para puntuar y extraer ideas de contenido de marketing. El artículo también menciona la importancia en la interpretación de los modelos para que los humanos puedan comprender y utilizar las ideas extraídas. Además, se discuten los desafíos y limitaciones actuales en el diseño de contenido de marketing y se proponen posibles mejoras utilizando modelos de lenguaje y visión de alta capacidad. El artículo concluye destacando el potencial de este enfoque para mejorar el diseño de contenido en diversas áreas, como la atención médica, las finanzas y las ventas bancarias.

La metodología utilizada en esta investigación se basa en el uso de redes neuronales profundas para modelar y mejorar el diseño de contenido de marketing digital. Se propone un enfoque multimodal que utiliza imágenes y texto como entrada para predecir las tasas de éxito y proporcionar información interpretativa para guiar a los especialistas en marketing en la mejora de su diseño existente.

El proceso consta de varias etapas, que incluyen la recopilación de datos de experimentos controlados realizados por especialistas en marketing en diferentes ubicaciones de marketing digital. Estos datos incluyen elementos de contenido, como imágenes, texto y características categóricas, junto con las tasas de éxito correspondientes.

Se utiliza una red neuronal multimodal para predecir la tasa de éxito utilizando la entrada multimodal. Además, se utilizan métodos de atribución de modelos para generar información interpretativa sobre el contenido. Estos métodos permiten identificar los elementos clave en el contenido que contribuyen positiva o negativamente a la tasa de éxito.

La metodología también incluye la evaluación cuantitativa del rendimiento de las ideas generadas y su correlación con los resultados observados. Se utiliza un enfoque de tres pasos para evaluar el rendimiento de las ideas en términos de la correlación entre la aplicación de modificaciones guiadas por las ideas y los resultados observados.

En resumen, la metodología combina el uso de redes neuronales profundas, métodos de atribución de modelos y evaluación cuantitativa para modelar y mejorar el diseño de contenido de marketing digital. (Fanjie Kong, 2023)

1.1.4 Análisis y propuesta para el aprovechamiento de herramientas de machine learning en el sector del marketing digital en Netbangers

Este estudio se enfoca en explorar cómo el machine learning y metodologías derivadas están influyendo en la toma de decisiones estratégicas en marketing digital. Netbangers, una agencia de mercadeo digital, es el escenario para este análisis y propuesta. El objetivo es aprovechar estas tecnologías para mejorar la eficiencia y el impacto de las estrategias comerciales.

En el artículo se examina cómo estas herramientas tecnológicas están asistiendo su forma de trabajar en decisiones complejas. Se investiga el proceso de marketing relacional de Netbangers y se identifica cómo el machine learning puede generar sinergias y optimizar el rendimiento. A través de una revisión detallada de las diversas soluciones de machine learning disponibles, en este estudio se utilizó una herramienta llamada WEKA, estudia el flujo de trabajo que tiene la empresa y determina cuales herramientas se adapta mejor para facilitar el trabajo de Netbangers.

La propuesta se enfoca en implementar estas herramientas tecnológicas en la agencia, resaltando los efectos positivos en el marketing relacional tanto para la agencia como para sus clientes. La tecnología del machine learning se presenta como una valiosa herramienta para optimizar la toma de decisiones y mejorar la satisfacción del cliente.

Este trabajo concluye que el machine learning es altamente beneficioso para las agencias digitales, especialmente en la optimización de estrategias de marketing. El enfoque en satisfacer las necesidades de los clientes y la implementación estratégica de estas herramientas brinda a Netbangers una ventaja competitiva en un entorno dinámico y digital. (Franco, 2020)

1.1.5 Machine learning en marketing: ventajas y ejemplos de aplicación

Este artículo explica cómo el Machine Learning puede mejorar el marketing digital, al permitir analizar grandes cantidades de datos y extraer información relevante para la toma de decisiones. Se enumeran algunas ventajas del Machine Learning en el marketing, como la personalización de ofertas, la segmentación de clientes, la optimización de campañas, la predicción de comportamientos y la generación de contenidos. También se presentan algunos ejemplos de aplicación del Machine Learning en el marketing, como el uso de chatbots, el reconocimiento facial, el análisis de sentimientos y la recomendación de productos.

La eficacia en el procesamiento de datos en constante aumento está intrínsecamente ligada a la calidad de la información recolectada. En la actualidad, cualquier empresa puede acceder a vastas cantidades de datos de sus usuarios, y la clave reside en la habilidad de analizarlos para incorporar conclusiones en la estrategia de marketing. El machine learning ofrece diversas aplicaciones en este campo:

Mejora de la Experiencia del Consumidor en Compras Online: Mediante recomendaciones personalizadas de productos que se ajustan a sus preferencias, garantizando disponibilidad y brindando alternativas si es necesario, se logra una atención continua (24/7) que los clientes valoran positivamente.

Incremento en la Personalización de Mensajes: La personalización de mensajes genera comodidad en los usuarios y fortalece su lealtad hacia la marca.

Fortalecimiento de la Seguridad de Usuarios y Datos: Algoritmos inteligentes permiten la predicción de comportamientos y posibles ataques, aumentando la seguridad de los usuarios y su información.

Detección de Opiniones sobre Productos: El análisis permite detectar opiniones positivas y negativas sobre productos, lo que facilita abordar problemas de reputación en línea.

Reducción de la Tasa de Abandono de Clientes: Identificar usuarios con mayor riesgo de abandonar a la competencia posibilita el envío de mensajes personalizados para recuperar su interés.

Generación de Ingresos Adicionales: Utilizando información del machine learning, es posible predecir las preferencias futuras de los clientes y desarrollar nuevos productos o servicios acorde a esas necesidades.

Publicidad Personalizada: La información del cliente permite ofrecer anuncios adaptados a sus intereses, optimizando la influencia de los anuncios.

El aprendizaje automático está presente en muchas facetas de la vida cotidiana, como el reconocimiento facial y de voz en dispositivos móviles o la predicción del tráfico en sistemas GPS. El análisis de datos en búsqueda de patrones solía ser laborioso, pero el machine learning simplifica el proceso, liberando recursos para otras operaciones.

Para ampliar conocimientos sobre técnicas y herramientas de análisis masivo, el Máster en Marketing Intelligence online de UNIR es una opción para fortalecer habilidades en marketing empresarial. (Unir, 2023)

1.2 El machine learning y sus aplicaciones de contenido digital en la industria 4.0.

El machine learning es una técnica de aprendizaje automático que permite a las máquinas reconocer patrones a partir de datos, algunas técnicas sin que sea necesario intervenir para que el proceso pueda ser autónomo. Esto significa que las máquinas pueden correlacionar y seleccionar grupos de datos y aprender a realizar tareas complejas a partir de ellos, como el reconocimiento de imágenes o analizar textos.

El machine learning es uno de los aportes importantes generados en la industria 4.0, ya que permite a las máquinas realizar tareas complejas y tediosas de manera autónoma, lo que facilita el proceso de automatización y mejora la eficiencia. Además, el machine learning también puede ser utilizado para mejorar la seguridad de las redes y los sistemas informáticos, así como para detectar y prevenir fraudes.

El Machine learning se basa en matemáticas como la estadística, métodos numéricos y computación. Este último no llegó hasta la segunda guerra mundial para solucionar problemas logísticos en tiempos de guerra, pero se destacan tres métodos que establecieron la base de esta técnica.

El Teorema de Bayes es una herramienta matemática para calcular la probabilidad de un evento dado otro evento o conjunto de eventos. Utilizando esta herramienta, el aprendizaje automático puede tomar datos observados y calcular razonablemente la probabilidad de una respuesta dada.

El Ajuste por Mínimos cuadrados: El método de los mínimos cuadrados se utiliza para ajustar los datos a una curva, generalmente una curva lineal. Utilizando este método, el aprendizaje automático puede ajustar los datos a una curva y predecir los resultados sin tener que realizar cálculos complicados.

Esto significa que el algoritmo intentará encontrar los parámetros de una función que minimicen el error entre los datos reales y los datos predichos por la función. Esta técnica se aplica en muchas áreas, como la regresión lineal, la regresión logística y la regresión no lineal.

Cadenas de Markov Es un modelo matemático para predecir qué estado ocurrirá en el futuro a partir de un estado actual. Estos modelos se usan para predecir el comportamiento de muchos sistemas, desde sistemas de recomendación hasta el aprendizaje de lenguaje natural.

Estos modelos se usan ampliamente en el aprendizaje automático para predecir la probabilidad de una secuencia de eventos dada una secuencia inicial. (Kaushik, 2011)

Dentro del departamento de marketing se puede mejorar el proceso de segmentación de mercado y para ello decide utilizar un algoritmo de aprendizaje automático que permita clasificar a los clientes en función de su comportamiento.

El objetivo de este proceso es poder personalizar la oferta de productos y servicios a cada cliente, de forma que se satisfagan mejor sus necesidades.

El algoritmo de aprendizaje automático utilizado es una red neuronal, que ha demostrado ser eficaz en la clasificación de datos.

Algunos de los datos utilizados para el entrenamiento de una red neural son los siguientes:

- Edad
- Sexo
- Ingresos
- Nivel de estudios
- Localización
- Ocupación
- Frecuencia de compra

Los resultados obtenidos son analizados para concluir su efectividad, ya que la red neural es capaz de clasificar correctamente a los clientes en función de sus hábitos de compra u otros comportamientos.

El rendimiento de sus ventas a través de la mejora de la oferta de productos en función de las preferencias de los clientes. Para ello, se utilizará un algoritmo de aprendizaje automático que analice el historial de compras de los clientes para determinar sus gustos e intereses.

El objetivo de mejorar el rendimiento de las ventas a través de la mejora de la oferta de productos en función de las preferencias de los clientes. Un algoritmo de aprendizaje automático que analice el historial de compras de los clientes para determinar sus gustos e intereses.

Este algoritmo está basado en regresión logística, que será capaz de estimar la probabilidad de que un cliente esté interesado en un producto dado sus gustos e intereses. Para mejorar la oferta de productos en función de las preferencias de los clientes, lo que se espera que mejore el rendimiento de las ventas.

Cuando se está buscando mejorar el servicio al cliente a través de la detección de problemas en el proceso de atención al cliente. Para ello, se va a utilizar un algoritmo que analiza las llamadas telefónicas entrantes para detectar problemas en el proceso de atención al cliente.

El proceso de detección de problemas en el proceso de atención al cliente se puede dividir en tres etapas:

Etapa 1: Entrada de datos: se reciben las llamadas telefónicas y se guardan en una base de datos.

Etapa 2: Pre procesamiento: se realiza el pre proceso de las llamadas, que consiste en la extracción de las características de las llamadas.

Etapa 3: Detección de problemas: se utiliza un modelo que permita detectar el problema en el proceso de atención al cliente, clasificarlo, archivar la información y generar soluciones más comunes a ese problema en específico.

Uno de los roles que tienen el machine learning se encuentra el desarrollo de ciudades inteligentes, es la digitalización y optimización para analizar grandes cantidades de datos generados por un una considerable cantidad de usuarios, lo que permite identificar patrones y tendencias que pueden ser útiles para tomar mejores decisiones.

Por ejemplo, el machine learning puede ser utilizado para analizar el tráfico y determinar qué rutas son más congestionadas, o para analizar el consumo de energía y determinar cuáles son

las zonas de la ciudad donde se consume más energía, además de múltiples aplicaciones en logística, negocios y en la gestión de clientes y marketing (Cuadrado, 2019).

El uso de técnicas como Support Vector Machine (SVM), se han utilizado para el análisis financiero sobre las estrategias de marketing a lo largo del tiempo, donde las principales métricas que determinaron el estudio fueron cuatro funciones diferentes de kernel en aplicados al modelo SVM, la función lineal kernel, polynominal kernel, radial gausiana kernel y sigmoidea kerne, donde la función más precisa en ese estudio fue la función radial gausiana kernel para la generación de predicciones dinámicas para el análisis de los datos de marketing.

Esta técnica se utiliza para obtener información acerca del comportamiento de los usuarios en el mercado financiero, debido a que los datos del mercado financiero son muy volátiles, es decir, que cambian constantemente, y con el uso de SVM se puede predecir con mayor precisión el comportamiento de los usuarios con respecto a los datos del mercado.

Otra forma en que SVM se ha utilizado para el análisis financiero es para la detección de fraudes financieros. Esta técnica se ha utilizado para detectar patrones sospechosos en los datos financieros de los usuarios y para predecir con precisión la probabilidad de que una transacción sea fraudulenta. Esto permite a los bancos y otras instituciones financieras tomar medidas proactivas para prevenir el fraude financiero.

En conclusión, el uso de SVM en el análisis financiero se ha demostrado ser una técnica útil para detectar patrones sospechosos, predecir el comportamiento de los usuarios y detectar fraudes financieros. Esta técnica ha demostrado ser una herramienta útil para ayudar a las instituciones financieras a tomar decisiones estratégicas acertadas. (Jiang, 2020)

1.2.1 Impactos del uso de machine learning en e-commerce

De todos los departamentos de una empresa, el marketing es probablemente el que más puede beneficiarse de la inteligencia artificial. Las actividades principales del marketing son comprender las necesidades de los clientes, relacionarlas con productos y servicios y persuadir a las personas para que compren, capacidades que la IA puede mejorar drásticamente. No es de extrañar que un análisis de McKinsey de más de 400 casos avanzados de uso mostrara que el marketing era el ámbito donde la IA aportaría el mayor valor.

Los directores de marketing están cada vez más abrazando la tecnología: una encuesta de la American Marketing Association de agosto de 2019 reveló que la implementación de la IA había

aumentado un 27% en el año y medio anterior. Y una encuesta global de Deloitte de 2020 a los primeros adoptantes de IA mostró que tres de los cinco objetivos principales de la IA eran orientados al marketing: mejorar productos y servicios existentes, crear nuevos productos y servicios y mejorar las relaciones con los clientes.

Si bien la IA se ha introducido en el marketing, esperamos que asuma roles cada vez más importantes en toda la función en los próximos años. Dada la enorme potencialidad de la tecnología, es crucial que los directores de marketing entiendan los tipos de aplicaciones de IA de marketing disponibles hoy en día y cómo pueden evolucionar.

Basándonos en más de una década de experiencia en el estudio de análisis de datos, IA y marketing y asesorando a empresas de todas las industrias al respecto, hemos desarrollado un marco que puede ayudar a los directores de marketing a clasificar los proyectos de IA existentes y planificar la implementación de futuros. Pero antes de describir el marco, veamos el estado actual. La IA de hoy.

Muchas empresas ahora utilizan IA para manejar tareas específicas, como la colocación de anuncios digitales (también conocida como "compra programática"); asistir en tareas amplias, como mejorar la precisión de las predicciones (piense en pronósticos de ventas); y mejorar los esfuerzos humanos en tareas estructuradas, como el servicio al cliente. (Consulte la barra lateral "Aplicaciones de IA bien establecidas en marketing" para obtener una lista de algunas actividades comunes que la IA puede respaldar). Las empresas también emplean IA en todas las etapas del viaje del cliente.

Cuando los clientes potenciales están en la fase de "consideración" e investigando un producto, la IA les dirigirá anuncios y puede ayudar a guiar su búsqueda. Vemos esto sucediendo en el minorista de muebles en línea Wayfair, que utiliza IA para determinar qué clientes son más persuadibles y, en función de sus historiales de navegación, elegir productos para mostrarles. Y los bots habilitados para IA de empresas como Veeva pueden ayudar a los especialistas en marketing a comprender las necesidades de los clientes, aumentar su compromiso en una búsqueda, empujarlos en una dirección deseada (digamos, a una página web específica) y, si es necesario, conectarlos con un agente de ventas humano por chat, teléfono, video o incluso "navegación conjunta" -permitiendo que un agente ayude al cliente a navegar por una pantalla compartida. La IA puede simplificar el proceso de ventas mediante el uso de datos extremos. (Thomas H. Davenport, 2023)

1.3 Casos aplicados de Machine learning en el marketing

El machine learning, una rama de la inteligencia artificial, está revolucionando la forma en que las empresas hacen marketing. Esta tecnología puede analizar grandes cantidades de datos y aprender de ellos para hacer predicciones precisas y automatizar tareas repetitivas.

En el marketing, el machine learning se utiliza para mejorar la segmentación de audiencias, personalizar contenidos, predecir comportamientos de los clientes y optimizar campañas publicitarias, entre otras cosas. A continuación, se hace mención a algunos casos destacados de su aplicación en el marketing.

1.3.1 Contenido publicitario de JPMorgan generado automáticamente

La firma JPMorgan Chase, uno de los gigantes financieros más importantes del mundo, ha equipado a su equipo de marketing con una herramienta de inteligencia artificial (IA) llamada Persado. Esta herramienta utiliza el aprendizaje automático y la tecnología de IA para escribir textos publicitarios que captan la atención de los consumidores y generan un contenido más amigable para los clientes.

La colaboración con Persado ha permitido a JPMorgan Chase obtener tasas de clics más altas con su redacción publicitaria renovada. Esto significa que la IA de Persado está produciendo textos publicitarios que los consumidores encuentran más interesantes y relevantes, lo que a su vez aumenta la efectividad de la publicidad.

La tecnología de Persado utiliza algoritmos y aprendizaje automático para analizar grandes cantidades de datos de los intereses o tendencias del mercado, lo que le permite entender las preferencias y los comportamientos de los consumidores. Con esta información, la IA puede producir mensajes publicitarios personalizados y efectivos para cada audiencia.

Esta colaboración entre JPMorgan Chase y Persado es un ejemplo de cómo la inteligencia artificial está transformando el marketing y la publicidad. Con la ayuda de la IA, las empresas pueden producir contenido publicitario más relevante, aumentar la eficacia de su publicidad y mejorar la experiencia del cliente en general. (Thomas, 2023)

La colaboración con Persado ha permitido a JPMorgan Chase mejorar significativamente las tasas de clics en sus anuncios, gracias a la redacción publicitaria renovada generada por la IA. Esto indica que Persado es capaz de producir textos publicitarios que resultan más interesantes y relevantes para los consumidores, lo que se traduce en una mayor efectividad de la publicidad.

La tecnología de Persado se basa en algoritmos y aprendizaje automático, lo que le permite analizar grandes volúmenes de datos sobre los intereses y tendencias del mercado. Esta capacidad de comprensión de las preferencias y comportamientos de los consumidores permite a la IA de Persado generar mensajes publicitarios personalizados y efectivos para cada audiencia.

El caso de JPMorgan Chase y Persado ejemplifica cómo el machine learning está transformando el campo del marketing y la publicidad. Gracias a la utilización de estas prácticas, las empresas pueden crear contenido publicitario más relevante, aumentar la eficacia de sus estrategias de publicidad y mejorar la experiencia general del cliente.

1.3.2 Recomendación automática de Netflix

El sistema de recomendación de Netflix utiliza el aprendizaje automático para mejorar su eficacia. Esta tecnología también es utilizada por la compañía en investigaciones y pruebas A/B. Además, la empresa utiliza el aprendizaje automático para tomar decisiones sobre qué contenido agregar a su plataforma de streaming y para optimizar la producción de sus series y películas originales a través de su estudio de cine.

El sistema de recomendación de Netflix es uno de los más reconocidos en el mundo del entretenimiento. Este utiliza el aprendizaje automático para analizar el historial de visualización de cada usuario, así como las calificaciones y comentarios que ha hecho sobre los contenidos que ha visto. Con esta información, el sistema genera recomendaciones personalizadas para cada usuario, lo que mejora su experiencia de usuario y aumenta su satisfacción.

Además, el aprendizaje automático también es utilizado por Netflix para analizar los datos y tomar decisiones sobre el contenido que se agregará a su plataforma. Al utilizar algoritmos de aprendizaje automático, Netflix puede identificar patrones en la audiencia y prever qué contenido será exitoso. Esto les permite tomar decisiones más informadas sobre qué películas y series originales producir y lanzar.

En resumen, Netflix utiliza el aprendizaje automático para mejorar la experiencia de usuario de sus clientes y para tomar decisiones más informadas sobre la producción y adición de contenido a su plataforma de streaming. Esta tecnología es un ejemplo de cómo las empresas están utilizando el aprendizaje automático para mejorar sus servicios y productos en el mundo digital. (Thomas, 2023)

Además de Netflix, otras compañías reconocidas como Amazon y Google también aplican el aprendizaje automático en sus plataformas para ofrecer recomendaciones personalizadas a sus usuarios.

Amazon utiliza técnicas de machine learning en su sistema de recomendación para sugerir productos de consumo a sus clientes. El algoritmo analiza los patrones de compra, navegación y preferencias de cada usuario, así como las interacciones de otros usuarios con productos similares. De esta manera, el sistema establece correlaciones entre usuarios y proporciona recomendaciones de productos relevantes que puedan resultar de interés para cada persona.

Del mismo modo, Google emplea el aprendizaje automático en varios de sus servicios para brindar recomendaciones adaptadas a cada usuario. Por ejemplo, en su motor de búsqueda, utiliza algoritmos que analizan el historial de búsqueda y navegación de los usuarios, así como los patrones de comportamiento de otros usuarios con búsquedas similares. De esta forma, Google ofrece resultados de búsqueda más relevantes y sugiere contenido relacionado que se ajuste a las necesidades e intereses individuales de cada usuario.

Estos ejemplos de Amazon y Google demuestran cómo el uso del aprendizaje automático en los sistemas de recomendación se ha vuelto común en las plataformas digitales. Estas empresas emplean análisis de datos y correlaciones entre usuarios para ofrecer sugerencias personalizadas, mejorando así la experiencia del usuario y aumentando la relevancia de los productos o contenido presentados.

1.3.3 Algoritmo de uso libre de Airbnb

Airbnb es conocida por ser una empresa líder en la aplicación de la inteligencia artificial y el machine learning en sus productos y servicios. Uno de estos productos es Aerosolve, una biblioteca de machine learning de código abierto.

Esta biblioteca es ideal para ejecutar algoritmos con características fácilmente interpretables, como la búsqueda de palabras clave y filtros para alquileres vacacionales. Aerosolve es la misma biblioteca que Airbnb utiliza para predecir la probabilidad de reserva y sugerir precios a sus anfitriones. Con el uso de esta biblioteca, Airbnb ha logrado mejorar la experiencia de sus usuarios y aumentar la eficiencia en su plataforma. (Thomas, 2023)

El uso de este algoritmo para diversas tareas, como la búsqueda de palabras clave y la aplicación de filtros en la plataforma de alquileres vacacionales. Esta biblioteca permite a Airbnb predecir la probabilidad de reserva y proporcionar sugerencias de precios a sus anfitriones. Gracias al uso

de Aerosolve, Airbnb ha logrado mejorar significativamente la experiencia de sus usuarios, así como aumentar la eficiencia en su plataforma.

La utilización de Aerosolve es solo uno de los ejemplos en los que Airbnb ha aplicado el machine learning para optimizar sus servicios. La empresa continúa explorando nuevas formas de aprovechar esta tecnología para personalizar las recomendaciones, mejorar la precisión de la búsqueda y ofrecer experiencias más relevantes a sus usuarios.

1.3.4 Servicio al cliente automatizado por chat

Ada es una plataforma de interacción de marca que permite a las empresas brindar un servicio al cliente de alta calidad y consistente a través de una función de chat impulsada por inteligencia artificial (IA). La plataforma de Ada utiliza modelos de aprendizaje automático para analizar el texto en más de 100 idiomas y predecir las necesidades de los clientes.

La tecnología proactiva de Ada permite a las empresas reducir el tiempo necesario para resolver problemas y proporcionar respuestas precisas y rápidas a las preguntas de los clientes. Con su capacidad para analizar los datos de conversación, la plataforma de Ada puede proporcionar a las empresas una comprensión más profunda de las necesidades de los clientes y ayudar a mejorar sus servicios y productos.

En resumen, la plataforma de chat de Ada impulsada por IA ofrece una forma eficiente y efectiva para que las empresas brinden un servicio al cliente excepcional y de calidad, lo que puede mejorar la satisfacción del cliente y la fidelidad a la marca. (Thomas, 2023)

La implementación de servicios de atención al cliente automatizados por chat impulsados por machine learning no es algo exclusivo de un solo negocio, sino que varias empresas han utilizado esta tecnología con éxito, es una práctica que cada vez más empresas implementan (Amazon, Uber, Spotify, Mercado libre y Steam) por su facilidad de dar seguimiento a cada caso de reclamaciones en menos tiempo, aumentando así la satisfacción del cliente y la credibilidad hacia la empresa.

CAPÍTULO II. MARCO HISTÓRICO - CONTEXTUAL

El estudio se realizó en la microempresa Isabel Taracena donde ofrece servicios fotográfico profesional donde inicio operaciones de manera autónoma en el año 2018 en Ciudad de México, que tiene prácticas en marketing digital a través de una página web y redes sociales.

Entre los servicios que ofrece se encuentra cubrir eventos y ofrecer obras tanto de su propia autoría como solicitado por el cliente, además de haber colaborado con diferentes editoriales y revistas como Gourmet, Playboy, Futbol total, Open Woman of 10.

A lo largo del 2022 se llevó a cabo el estudio a través de la recopilación de datos y su análisis para obtener conclusiones y recomendaciones que aporten a su desempeño en los medios digitales, los principales datos obtenidos fue sobre la actividad que tenía la página web (<https://www.isabeltaracena.com>) y redes sociales (Instagram: @isa.taracena) para identificar los momentos puntuales de mayor interacción y desempeño de los medios digitales. Se obtuvo el comportamiento de los usuarios, el tipo de contenido que más les interesa, las horas de mayor actividad y los mejores días para publicar.

A través de estos medios digitales da a conocer su trabajo y el llamado a la acción es contactar o agendar el servicio solicitado, además de realizar sus propios proyectos de corte artísticos en forma de ensayos y collages, aunque ella está familiarizada con el uso de medios digitales no suele adentrarse al uso de las estadísticas de sus medios para una toma de decisiones.

Esto con el fin de mejorar el posicionamiento de su página web para mejorar su visibilidad, así como mejorar la estrategia de contenido para los medios digitales para asegurar una mayor interacción a partir de sus publicaciones de contenido de manera periódica y en los días y horas de mayor actividad de los usuarios para asegurar una mayor interacción, a continuación, se redactará a más profundidad el contexto del entorno del planteamiento del problema.

Principales productos y servicios que se ofrecen: Fotografías de naturaleza artística, cubrir Eventos y creación de catálogos de productos además de impartir talleres relacionados a la fotografía. Los desafíos específicos que se le presenta a la empresa en estudio se observaron las siguientes: dificultades para equilibrar la creatividad artística con la estrategia digital, la falta de tiempo (ya que es el trabajo de una sola persona que es Isabel Taracena) y recursos que limitan su capacidad para realizar un seguimiento detallado de las estadísticas y tendencias digitales.

Para los datos principales obtenidos se incluye la identificación de momentos de mayor interacción y desempeño de los medios digitales, el comportamiento de los usuarios, el tipo de contenido más atractivo y las horas y días óptimos para la publicación. Sin embargo, a pesar de su familiaridad con el uso de medios digitales, la empresa no ha aprovechado plenamente las estadísticas y tendencias digitales para la toma de decisiones.

En el siguiente capítulo se detallan de manera específica las diferentes problemáticas que presenta el estudio de fotografía Isabel Taracena. Se explorarán en profundidad los desafíos que enfrenta la empresa, como la dificultad de equilibrar la creatividad artística con la estrategia digital. Además, se analizarán el tipo de contenido más atractivo y los horarios y días óptimos para la publicación. A pesar de la familiaridad de la empresa con el uso de medios digitales, no se ha aprovechado completamente las estadísticas y tendencias digitales para la toma de decisiones. Por lo tanto, en el siguiente capítulo se proporcionará una visión completa de las problemáticas específicas que enfrenta el estudio de fotografía Isabel Taracena.

CAPITULO III. PLANTEAMIENTO DEL PROBLEMA

El problema de posicionamiento de una página web ha evolucionado con el surgimiento de tecnologías de almacenamiento y procesamiento de datos, así como el crecimiento exponencial del comercio electrónico. En los primeros años de internet, las técnicas de posicionamiento se centraban en la optimización de palabras clave y enlaces, pero en la actualidad, el comportamiento de los usuarios revela que la mayoría de las visitas se concentran en los primeros resultados de búsqueda. Por lo tanto, si una página web no logra posicionarse de manera efectiva, sus posibilidades de ser visualizada y atraer tráfico orgánico se reducen significativamente.

Para abordar este problema, se han desarrollado técnicas de posicionamiento, como el SEO (Search Engine Optimization), que utilizan un conjunto de métricas para mejorar el ranking de una página web en los resultados de búsqueda de Google. Sin embargo, en un entorno tan competitivo, estas técnicas pueden no ser suficientes para garantizar la relevancia y visibilidad deseada.

En el caso de la empresa en estudio, este problema de posicionamiento se vuelve aún más relevante debido a su participación en el mercado electrónico y la necesidad de llegar a un público amplio en línea. La capacidad de captar la atención de los usuarios y destacarse frente a la competencia se convierte en un factor crucial para su éxito, a través del conjunto de métricas que se utilizan para mejorar el ranking de una web en los resultados que emite el propio google (Google Analytics), y este permita tener una mayor aceptación por parte del algoritmo de búsqueda y posicionamiento conocido como PageRank.

En la página web no presenta un posicionamiento en su medio por falta de optimización SEO, teniendo nulas visitas a la página web teniendo un promedio mensual de 2 visitas en las 2 primeras semanas antes de realizar las primeras implementaciones de posicionamiento SEO Tabla 3. De esta manera a medida que el mercado digital continúa creciendo, es crucial que la empresa aborde este problema para aumentar su visibilidad y captar la atención de un público amplio en línea.

Sin embargo, no siempre es suficiente para ser relevante en un entorno tan competitivo, el mercado electrónico siguen en auge, en 2021 en México se reportó un incremento de 7.3% de

nuevos usuarios en internet, que representa un mercado potencial de 4,660 millones de personas (Naranjo, 2021)

Es fundamental que la empresa en estudio considere no solo la optimización técnica de su página web, como los tiempos de carga y el diseño responsivo, sino también la coherencia entre el contenido ofrecido y el término buscado. Además, es importante aprovechar las herramientas de machine learning disponibles, que permiten gestionar y optimizar el contenido de manera más efectiva, así como utilizar herramientas de análisis digitales para comprender mejor los datos y maximizar su utilidad.

El problema de posicionamiento en motores de búsqueda representa un desafío significativo para la empresa en estudio en el contexto actual del comercio electrónico en México. Para superar este desafío, es necesario comprender el desarrollo histórico del problema, implementar técnicas de posicionamiento efectivas y utilizar herramientas de análisis y optimización adecuadas. De esta manera, la empresa podrá aumentar su visibilidad en línea y aprovechar al máximo las oportunidades que brinda el uso eficiente de la información.

Es fundamental entender el alcance de las técnicas básicas de machine learning y su aplicación en el campo de la mercadotecnia digital. Además, es importante unificar las herramientas de análisis de datos y comprender cómo influyen en la práctica. También se busca obtener información válida más allá de los datos proporcionados por Google. El objetivo es poder realizar pronósticos o crear modelos óptimos que permitan mejorar el enfoque del contenido web y alcanzar de manera más efectiva al público objetivo (Saura, 2020).

En una página web, la cantidad de tráfico e interacciones que tiene es un factor determinante al momento de monetizar el contenido, ya sea a través de publicidad, suscripción, venta o donativos directos. Para medir a los usuarios y las interacciones, se utilizan técnicas de clúster o agrupaciones de usuarios y sus datos, con el fin de realizar estrategias al mercado meta o perfil de usuarios más relevantes (Muhammad Ridwan Andi Purnomo, 2020).

Los registros deben cumplir características específicas, y la pérdida de datos puede responder a problemas de la persona que ingresó los datos o a la generación de información por parte de las empresas en sus procesos internos. Por lo tanto, gestionar esta información se vuelve importante en el mundo empresarial, ya que crea un valor agregado directa o indirectamente.

La manera en que se gestiona esta información se ha vuelto más compleja debido a la cantidad inmensa de datos que no se pueden analizar de manera analógica. Además, dependiendo de los resultados o conclusiones que se requieran obtener de estos datos, no siempre es posible procesarlos con métodos convencionales. Por lo tanto, el auge de las ciencias de datos, inteligencia artificial y aprendizaje de máquina ha adquirido una mayor relevancia en este ámbito.

Hasta el momento, el machine learning se ha desarrollado principalmente para mejorar el servicio al consumidor, desde analizar el patrón de consumo de los clientes hasta ser implementada como parte de un motor de búsqueda. Sin embargo, el aprendizaje de máquina representa un paso aún mayor en la interacción entre los usuarios y la información. Por lo tanto, si se desarrollaran estos modelos en un entorno empresarial, se lograría una optimización que muchas empresas buscan o necesitan implementar para tener una mayor efectividad en sus sistemas internos y externos.

3.1 Justificación

La presente investigación se centra en el problema del posicionamiento deficiente de la página web de la empresa en estudio. Esta problemática ha sido identificada como una barrera significativa que afecta su visibilidad, alcance en el mercado y, en última instancia, sus ingresos. Al abordar este problema, se espera contribuir de manera sustancial a la mejora de su desempeño en el entorno digital y a su competitividad en el mercado actual.

Es crucial destacar la importancia de este problema en el contexto actual, donde el comercio electrónico ha experimentado un crecimiento exponencial en México, ya que México es el segundo país con mayor número de usuarios en internet en América Latina por debajo de Brasil con 96,87 millones de usuarios en el año 2022 de mercado potencial en el mercado electrónico y el consumo a través del comercio electrónico en México ha crecido exponencialmente en los últimos años, debido a que esta modalidad de compra ofrece numerosas ventajas para los consumidores, tales como: la comodidad de comprar desde cualquier lugar, la facilidad para comparar precios y productos, la seguridad en las transacciones y el ahorro de tiempo (statista, 2022).

La perspectiva del comercio electrónico en la economía nacional son históricas, ya que se tiene una expectativa positiva para el 2025, donde se considera un crecimiento proyectado del comercio electrónico en un 226% y el 24% de las ventas de las tiendas físicas se realizarán por este medio para el año 2025, el mercado mexicano amplía sus opciones de compra al

ecommerce, ya que hay una gran variedad de tiendas en línea que ofrecen productos y servicios de excelente calidad a precios muy competitivos. Por esta razón, cada vez son más los consumidores que optan por realizar sus compras por internet (Nuguez, 2021).

“El 78% de las pymes tiene presencia en redes sociales, y el 56% de ellas, vende a través de las mismas; mientras que el 58% de las pymes ya tiene su propio sitio web y el 28% lo utiliza para vender”; “El siguiente paso para estar en la nube tiene que ver con el procesamiento de gran cantidad de datos, con los beneficios y las capacidades de machine learning para tener información descriptiva y predictiva del usuario; y el tercer paso que cualquiera puede tomar, son las herramientas colaborativas a costos muy razonables” asegura Juan Carlos Gómez director de comercio minorista y multicanal de Google México (López, 2021).

El mercado en México presenta un incrementado total de 88.2 millones en el año 2020 siendo los dispositivos móviles (teléfonos inteligentes) el medio más utilizado para conectarse a internet, se estima que hay 84.1 millones de usuarios de internet en 2020, lo que representa el 72.0% de la población con una proporción del 71,3% de mujeres y el 72,7% de hombres en uso doméstico.

A través del censo realizado por el INEGI se estima que el 78,3% de la población en áreas urbanas son usuarios, en comparación con el 50,4% en áreas rurales. Usuarios urbanos en 2019 Se estima en 76,6%, y en las zonas rurales se estima en 47,7%. Las tres principales formas en que los usuarios se conectan a Internet en 2020 son: teléfonos inteligentes 96,0%, computadoras portátiles 33,7% y televisores inteligentes en un 22,2% (INEGI, 2021).

En promedio, una página web puede monetizar a través de publicidad entre \$1 y \$2 (dólares) por cada 1,000 visitantes únicos. Si se realizan buenas prácticas de SEO, se puede incrementar esta cifra en hasta un 50%, sin embargo, la actividad económica que ejerce en la fotografía es variante, desde cursos estimados en \$800 a \$1,800 pesos mexicanos dependiendo la cantidad de horas que dure el curso y la especialización de los temas, secciones de fotografía entre \$500 a \$2,000 pesos mexicanos dependiendo del evento.

De las principales actividades que realizan los internautas en 2020 se encuentra las plataformas de comunicación (correo, video llamadas, entre otros) 93,8%, la búsqueda de información 91,0% y el acceso a redes sociales 89,0% (INEGI, 2021).

Se estima que el crecimiento del comercio electrónico no extinguirá el mercado de las tiendas físicas, se espera un el crecimiento de un 58% de las ventas proyectadas entre el año 2021 y

2025, estas condiciones considera que la población joven tendrá mejor preparación en el uso de tecnología, la población adulta tendrá un incremento en sus ingresos y puedan utilizar las TIC's (tecnologías de la información y comunicaciones) sin tantas barreras técnicas especializadas por el desarrollo de interfaces aún más intuitivas.

Conforme más empresas se suman a esta tendencia, la lucha por destacar y captar la atención de los consumidores se vuelve cada vez más intensa. En este sentido, el posicionamiento eficiente de una página web se convierte en un factor crítico para lograr una ventaja competitiva sostenible.

La magnitud del problema se refleja en la pérdida de oportunidades de negocio que resulta de un posicionamiento deficiente. Estudios demuestran que una gran proporción de los usuarios de Internet confía en los primeros resultados de búsqueda para obtener información y realizar compras en línea. Por lo tanto, una empresa con un bajo posicionamiento en los motores de búsqueda sufre una disminución significativa en su visibilidad y, en consecuencia, pierde potenciales clientes y oportunidades de generar ingresos, la mayor cantidad de visitas que son visualizadas pertenecen a las primeras tres opciones de búsqueda, por lo que raramente se visualizan las últimas opciones de búsqueda (MIT Technology Review, 2017).

La importancia que tiene conocer la influencia del motor de búsqueda más popular es considerar que la empresa multinacional estadounidense detrás de Google (Alphabet) reporta ingresos de 74,900 millones de dólares en el primer trimestre del año 2021 convirtiendo a Google en la empresa que más ingresos genera a través de anuncios en internet además de los beneficios compuestos por los servicios del cómputo en la nube y publicidad en YouTube, siendo una de las múltiples empresas con mayor facturaciones tiene junto con las empresas multimillonarias detrás de las redes sociales más populares (Reuters, 2022).

La relevancia de abordar este problema trasciende los límites de la empresa en estudio. El mejoramiento del posicionamiento web no solo impactará positivamente su crecimiento y rentabilidad, sino que también contribuirá al desarrollo del comercio electrónico en México. Al fortalecer la presencia en línea de las empresas nacionales, se impulsará la economía del país, generando empleo y fomentando la competitividad a nivel nacional e internacional.

Si bien existen técnicas como el SEO y el sistema de PageRank que buscan mejorar el posicionamiento web, su aplicación no siempre es sencilla ni garantiza resultados óptimos. Aquí es donde la incorporación de técnicas de Machine Learning en esta investigación adquiere

relevancia. El uso de algoritmos de aprendizaje automático permitirá desarrollar modelos y estrategias más efectivas para mejorar el posicionamiento de la página web de la empresa, considerando sus características específicas, el comportamiento de los motores de búsqueda y las preferencias de los usuarios.

Algunas de las limitaciones técnicas que un usuario que genera información o contenido pueda tener se consideran las siguientes: el usuario no tiene conocimiento de programación ni tampoco pueda realizar una interpretación a un código o base de datos, tecnología obsoleta o no tiene acceso a herramientas especializadas para el manejo de datos ya sea por desconocimiento o por falta de presupuesto, que las actividades laborales de los usuarios no tenga tiempo para realizar un análisis de la información generada, no se le da un seguimiento a la información generada y solo se archiva como datos históricos, al realizar un análisis de la información no tener un objetivo, un flujo de información confuso o meta clara para orientar los resultados del análisis y falta de políticas para el uso de los sistemas de información (Ana Prieto, 2004).

Otro aspecto a considerar es que dentro de una empresa es importante que los colaboradores puedan conocer la naturaleza de la información que ellos generan, además de recalcar e incentivar el buen manejo o buenas prácticas de la información (no alterar ni falsificar datos), esto con el fin de favorecer la confidencialidad de la información, así como la toma de decisiones ante las estrategias más oportunas para su uso.

Además, las herramientas de machine learning de uso libre se han popularizado cada vez más incluso teniendo una integración con otras herramientas de pago (Tableau y Azure) por lo que son el instrumento principal para múltiples proyectos de análisis, para gestionar el contenido de una página web junto con su optimización SEO para ser mejorado. Además de ser planteada con una funcionalidad compatible con las herramientas de análisis digitales, para que sea accesible, fácilmente interpretado por los programas de análisis y los usuarios que no tenga un conocimiento profundo de estadística o informática puedan tener un mejor entendimiento de los datos.

Este trabajo busca comprender las limitaciones existentes en las estrategias de posicionamiento actuales y ofrecer soluciones más efectivas y personalizadas.

3.2 Objetivo General

Plantear mejores prácticas de marketing para la microempresa de fotografía “Estudio Isabel Taracena) mediante el análisis de página web, redes sociales y técnicas de machine learning

3.2.1 Pregunta Central de investigación

- ¿Cómo puede el uso de algoritmos de Machine Learning mejorar la efectividad de las campañas publicitarias en redes sociales de la microempresa en términos de aumento de conversiones y reducción de costos (ROI)?
- ¿Cuáles son las herramientas de Machine Learning más efectivas para analizar el comportamiento de los consumidores en línea y obtener información relevante para la microempresa en cuestión?
- ¿Qué estrategias específicas de marketing digital pueden usarse para aumentar el alcance y la visibilidad de la microempresa en su mercado objetivo y cómo pueden ser mejoradas con técnicas de Machine Learning?
- ¿Cuáles son los factores clave que influyen en la experiencia de usuario en línea y cómo puede el uso de Machine Learning ayudar a mejorarla para aumentar la visualización de la microempresa?
- ¿Qué tipos específicos de contenido son más efectivos para atraer a un mayor número de visitantes y aumentar el tráfico en línea de la microempresa, y cómo puede esto ser medido a través del análisis de datos con Machine Learning?
- ¿Cuáles son las características específicas que definen al cliente objetivo de la microempresa, y cómo puede ser definido de manera más objetiva y precisa a través de técnicas de Machine Learning?

3.3 Variables de investigación

En este apartado se presentan las variables utilizadas en el estudio, las cuales se consideran relevantes para el análisis de los datos y están directamente relacionadas con el problema, objetivo e hipótesis de la investigación.

Variables dependientes:

- Datos recopilados: Los datos recopilados representan una variable dependiente en este estudio, ya que estos datos se utilizan para llevar a cabo análisis destinados a mejorar

las prácticas de marketing digital, basándose en las visitas al sitio web y las interacciones en redes sociales. Esta variable es esencial para comprender el impacto de las estrategias de marketing digital en la microempresa.

- Técnicas de machine learning utilizadas: Esta variable se refiere a las técnicas específicas de machine learning seleccionadas y aplicadas en el análisis de datos y en la toma de decisiones dentro del ámbito del marketing digital. Estas técnicas incluyen algoritmos de clasificación, regresión, agrupación, recomendación y otros enfoques de aprendizaje automático utilizados para mejorar las estrategias de marketing y maximizar el rendimiento de la microempresa. Estas técnicas son esenciales para el análisis de datos y para lograr los objetivos del estudio.

Variables independientes:

Las variables que seleccionadas son indicadores para medir las interacciones y alcance de audiencia a la página web y publicaciones en redes sociales son las siguientes:

- Número de visitas al sitio web y redes sociales: Esta variable representa la cantidad de usuarios que acceden y visitan la página web y perfiles en redes sociales de la microempresa. Se busca determinar si la implementación de técnicas de machine learning tiene un efecto significativo en el aumento de visitas y en la retención de los usuarios a lo largo del tiempo.
- Engagement: Esta variable busca evaluar la posición y reconocimiento de la microempresa en las redes sociales a través de la interacción con las publicaciones (reacciones y comentarios). Se analizará el grado de influencia y visibilidad alcanzados, considerando el uso de técnicas de machine learning en las prácticas de marketing digital.

3.3.1 Variables para el análisis de geomarketing

Se han identificado las siguientes variables de control para considerar posibles factores externos que pueden influir en los resultados obtenidos:

Variables dependientes:

- Tipo de industria: Se define como "artistas, escritores y técnicos independientes" según la clasificación DENU (711510). Esta variable busca tener en cuenta las particularidades propias de la industria a la que pertenece la microempresa. Presupuesto destinado a marketing y publicidad: Esta variable permite considerar el nivel de recursos económicos

destinados por la microempresa para llevar a cabo sus actividades de marketing y publicidad, lo cual puede afectar la implementación y efectividad de las técnicas de machine learning.

Variable independiente:

- Nivel de competencia en el mercado: Se evaluará el grado de competencia presente en el mercado en el que se desenvuelve la microempresa, ya que puede influir en el impacto de las prácticas de marketing digital y las técnicas de machine learning.

En estas variables, se analiza la relación entre las variables independientes y dependientes, las cuales se reflejan en el problema, objetivo e hipótesis de la investigación. Además, se consideran variables intervinientes y de control que permiten contextualizar y controlar posibles influencias externas en los resultados obtenidos.

3.4 Hipótesis central de investigación

Hipótesis Central: Con la implementación de técnicas de machine learning y estrategias de marketing en redes sociales y sitio web se obtendrá un impacto positivo en el aumento del tráfico del sitio web <https://www.isabeltaracena.com/> y mejorará el posicionamiento de la microempresa de fotografía (Estudio Isabel Taracena) en el mercado objetivo.

Variable Dependiente: Esta variable engloba dos aspectos: aumento del tráfico del sitio web y mejora en el posicionamiento. Estos aspectos están relacionados con el objetivo general de aumentar la visibilidad y el impacto de la microempresa en el mercado de la fotografía y el arte visual.

Variable Independiente: Esta variable representa las estrategias que se están implementando, como técnicas de machine learning y estrategias de marketing en redes sociales y el sitio web.

Explicación:

Considerando el problema de tráfico y posicionamiento en el mercado que enfrenta la microempresa de fotografía, se plantea la hipótesis de que al aplicar técnicas de machine learning y estrategias de vinculación en redes sociales, realizar un análisis con el fin de lograr un aumento significativo en el tráfico del sitio web y una mejora en el posicionamiento de la empresa en el mercado objetivo.

La hipótesis se basa en la premisa de que las técnicas de machine learning permitirán un análisis más preciso y eficiente de los datos generados por el sitio web y las redes sociales, lo que conducirá a una optimización de las estrategias de vinculación y, como resultado, a un aumento en el tráfico del sitio web.

Además, se espera que esta mayor visibilidad y participación en línea contribuyan a fortalecer el posicionamiento de la microempresa en el mercado objetivo, generando un mayor impacto y reconocimiento entre su audiencia objetivo.

La investigación se centrará en recopilar y analizar datos relevantes, comparando el tráfico y el posicionamiento antes y después de la implementación de las técnicas de machine learning y estrategias de vinculación en redes sociales. Los resultados obtenidos permitirán evaluar la validez de la hipótesis planteada y determinar si estas acciones son efectivas para abordar el problema identificado y alcanzar el objetivo establecido.

3.5 Objetivos específicos

1 Evaluar el impacto del uso de algoritmos de Machine Learning en las campañas publicitarias de redes sociales de "Estudio Isabel Taracena" para aumentar el número de conversiones y reducir los costos asociados, mejorando el retorno de la inversión (ROI).

¿Cuál es el impacto del uso de algoritmos de Machine Learning en las campañas publicitarias de redes sociales para aumentar el número de visualizaciones y así reducir los costos asociados, mejorando el retorno de la inversión?

2 Identificar las herramientas de Machine Learning más adecuadas para analizar el comportamiento de los consumidores en línea y obtener información relevante que permita optimizar las estrategias de marketing digital de "Estudio Isabel Taracena".

¿Qué herramientas de Machine Learning son adecuadas para medir el comportamiento de los consumidores en línea?

3 Diseñar estrategias específicas de marketing digital basadas en técnicas de Machine Learning para aumentar el alcance y la visibilidad de "Estudio Isabel Taracena" en su mercado objetivo, maximizando el impacto y la interacción con el público.

¿Cómo se pueden diseñar estrategias específicas de marketing digital que facilite llegar e interactuar con su mercado objetivo?

4 analizar cómo la implementación de técnicas de Machine Learning influye en la retención de usuarios y el engagement en línea de "Estudio Isabel Taracena", identificando los factores clave que contribuyen a una mayor interacción con los clientes.

¿De qué manera influye la implementación de técnicas de Machine Learning en la retención de usuarios en redes sociales?

5 Mejorar la experiencia del usuario en línea de "Estudio Isabel Taracena" mediante la aplicación de técnicas de Machine Learning, optimizando la navegación, el diseño y el contenido para aumentar la satisfacción y el tiempo de permanencia en la página web y redes sociales.

¿Cómo se puede mejorar la experiencia del usuario en línea, como optimizar la navegación para que los usuarios se sientan atraídos al contenido?

6 Identificar qué tipos específicos de contenido generan un mayor tráfico y engagement en línea para "Estudio Isabel Taracena" mediante el análisis de datos con técnicas de Machine Learning, con el fin de optimizar la estrategia de contenidos.

¿Cuáles son los tipos específicos de contenido que generan un mayor tráfico para poder hacer una mejor segmentación entre el contenido y usuarios?

7 Definir de manera más objetiva y precisa al cliente objetivo de "Estudio Isabel Taracena" a través de técnicas de Machine Learning, segmentando y analizando datos demográficos y patrones de comportamiento, para adaptar estrategias de marketing de manera efectiva.

¿Cómo se puede definir al cliente objetivo de "Estudio Isabel Taracena" y que técnicas o herramientas de análisis serían útiles para identificarlos de manera oportuna?

8 Identificar las herramientas de Machine Learning adecuadas para aplicar en ventas mediante el análisis de las principales herramientas open source.

¿Cuáles son herramientas open source de Machine Learning?

9 Evaluar las plataformas que permiten crear una página web con funciones de marketing digital y la posibilidad de aplicar técnicas de Machine Learning como administrador.

¿Cuáles son las plataformas para crear una página web que tenga funciones de seguimiento de marketing?

10 Determinar las métricas actuales de la página web seleccionada y calcular su media de visitas antes de aplicar las técnicas de Machine Learning.

¿Cuáles son las métricas y estadísticas actuales de la página web antes de aplicar las técnicas de Machine Learning?

11 Seleccionar y aplicar las técnicas de Machine Learning más apropiadas para analizar las métricas y el historial de visitas de la página web seleccionada.

¿Qué técnicas de Machine Learning son apropiadas para analizar las métricas y el historial de visitas?

12 Comparar las métricas y el historial de visitas con las métricas obtenidas mediante las técnicas de Machine Learning y evaluar los resultados para comprobar la hipótesis planteada.

¿Cuál es la diferencia entre las métricas obtenidas mediante las técnicas de Machine Learning para comprobar la hipótesis planteada (si el Machine Learning puede ayudar a orientar el contenido de Isabel a su público objetivo)?

3.5.1 Preguntas secundarias de Investigación

¿Cómo afecta el uso de algoritmos de Machine Learning en las campañas de redes sociales para aumentar conversiones y reducir costos (ROI)?

¿Cuáles son las ventajas y desventajas de cada técnica de machine learning identificada para su aplicación en ventas?

¿Cuáles son las herramientas de Machine Learning más adecuadas para optimizar estrategias de marketing digital?

¿Cuáles son las características y funcionalidades clave de una página web que permita la implementación de técnicas de machine learning?

¿Cómo diseñar estrategias específicas de marketing digital con técnicas de Machine Learning para incrementar el alcance y visibilidad?

¿Cuál es el perfil demográfico y comportamiento de los visitantes actuales de la página web antes de aplicar las técnicas de machine learning?

¿Cómo se comparan las diferentes técnicas de machine learning en términos de precisión y eficacia al aplicarlas a las métricas-historial de la página web seleccionada?

¿Cuáles son los cambios significativos en las métricas y media de visitas después de aplicar las técnicas de machine learning en la página web analizada?

¿Cuál es el impacto de la aplicación de técnicas de machine learning en términos de aumento de visitas y mejora del posicionamiento de la página web en el mercado?

¿Cuáles son las técnicas de Machine Learning más apropiadas para analizar el comportamiento de los consumidores en línea y obtener información relevante para la microempresa en cuestión?

¿Qué estrategias específicas de marketing digital pueden usarse para aumentar el alcance y la visibilidad de la microempresa en su mercado objetivo y cómo pueden ser mejoradas con técnicas de Machine Learning?

¿Qué tipos específicos de contenido son más efectivos para atraer a un mayor número de visitantes y aumentar el tráfico en línea de la microempresa, y cómo puede esto ser medido a través del análisis de datos con Machine Learning?

¿Cuáles son las características específicas que definen al cliente objetivo de la microempresa, y cómo puede ser definido de manera más objetiva y precisa a través de técnicas de Machine Learning?

3.5.2 Supuestos Teóricos secundarios de investigación

¿Cuáles son las ventajas y desventajas de cada técnica de machine learning identificada para su aplicación en ventas?

Hipótesis alternativa: Existen diferencias significativas en términos de precisión y eficacia entre las diferentes técnicas de machine learning para su aplicación en ventas.

Hipótesis nula: No existen diferencias significativas en términos de precisión y eficacia entre las diferentes técnicas de machine learning para su aplicación en ventas.

¿Cuáles son las características y funcionalidades clave de una página web que permita la implementación de técnicas de machine learning?

Hipótesis alternativa: La presencia de características y funcionalidades específicas en una página web con funciones de marketing digital facilita la implementación exitosa de técnicas de machine learning.

Hipótesis nula: No hay relación entre las características y funcionalidades de una página web con funciones de marketing digital y la implementación de técnicas de machine learning.

¿Cuál es el perfil demográfico y comportamiento de los visitantes actuales de la página web antes de aplicar las técnicas de machine learning?

Hipótesis alternativa: Existen diferencias significativas en el perfil demográfico y comportamiento de los visitantes actuales de la página web antes de aplicar las técnicas de machine learning.

Hipótesis nula: No existen diferencias significativas en el perfil demográfico y comportamiento de los visitantes actuales de la página web antes de aplicar las técnicas de machine learning.

¿Cómo se comparan las diferentes técnicas de machine learning en términos de precisión y eficacia al aplicarlas a las métricas-historial de la página web seleccionada?

Hipótesis alternativa: Al comparar las diferentes técnicas de machine learning, se observarán diferencias significativas en términos de precisión y eficacia al aplicarlas a las métricas-historial de la página web seleccionada.

Hipótesis nula: No existen diferencias significativas en términos de precisión y eficacia al aplicar las diferentes técnicas de machine learning a las métricas-historial de la página web seleccionada.

¿Cuáles son los cambios significativos en las métricas y media de visitas después de aplicar las técnicas de machine learning en la página web analizada?

Hipótesis alternativa: Después de aplicar las técnicas de machine learning en la página web analizada, se observarán cambios significativos en las métricas y media de visitas.

Hipótesis nula: No se observarán cambios significativos en las métricas y media de visitas después de aplicar las técnicas de machine learning en la página web analizada.

¿Cuál es el impacto de la aplicación de técnicas de machine learning en términos de aumento de visitas y mejora del posicionamiento de la página web en el mercado?

Hipótesis alternativa: La aplicación de técnicas de machine learning tiene un impacto significativo en el aumento de visitas y mejora del posicionamiento de la página web en el mercado.

Hipótesis nula: No hay relación entre la aplicación de técnicas de machine learning y el aumento de visitas y mejora del posicionamiento de la página web en el mercado.

3.6 Alcances y Limitaciones

En el presente apartado se analizará el alcance y las limitaciones que se observaron a lo largo de la realización de este trabajo cuyo objetivo principal es incrementar la visibilidad y el impacto de la microempresa en el mercado de la fotografía y el arte visual.

Alcances:

- El estudio se enfocará en una microempresa de fotografía que vende sus propias obras, además de realizar sesiones fotográficas solicitadas para cubrir eventos personales y empresariales.
- Se utilizarán técnicas de machine learning para analizar los datos recopilados de la microempresa y mejorar sus estrategias de marketing digital.
- El objetivo principal es incrementar la visualización e impacto de la microempresa en el mercado de la fotografía y arte visual.
- Se propondrán recomendaciones específicas para la microempresa en cuestión, basadas en los resultados obtenidos.

Limitaciones:

- El estudio se limita a una sola microempresa de fotografía y puede no ser generalizable a otras empresas de otros sectores.
- Los resultados obtenidos dependerán en gran medida de la calidad y cantidad de los datos recopilados, así como de la capacidad de las técnicas de machine learning para analizarlos.
- La empresa al no ser un ecommerce (no realiza sus ventas directas desde medios digitales) no es posible garantizar de manera directa un aumento en las ventas o en los ingresos de la microempresa, ya que otros factores externos pueden influir en estos resultados.

El estudio se enfoca en técnicas de machine learning específicas y si bien este estudio presenta valiosas oportunidades para mejorar las estrategias de marketing digital de la microempresa de fotografía, es importante reconocer sus limitaciones.

A pesar de estas restricciones, se considera que las recomendaciones propuestas serán un punto de partida para fortalecer el posicionamiento y presencia de la microempresa "Estudio Isabel Taracena".

CAPITULO IV. MARCO TEÓRICO CONCEPTUAL

El presente trabajo de investigación se centra en el análisis y desarrollo de estrategias de marketing digital a través del uso de herramientas tecnológicas como el posicionamiento en buscadores aplicado en aprendizaje automático. El objetivo principal de este estudio es comprender cómo estas herramientas pueden ayudar a las empresas a mejorar su desempeño en el mercado actual y cómo pueden ser aplicadas de manera efectiva para alcanzar sus objetivos comerciales.

Para lograr este objetivo, se investigaron diversos temas aportando al marco teórico relacionado con el marketing digital, el posicionamiento en buscadores y su aplicación en las diferentes técnicas - algoritmos del aprendizaje automático para sustentar la hipótesis propuesta. Se explorarán las diferentes teorías y enfoques relacionados con estos temas para comprender su funcionamiento y su impacto actual.

La investigación se enfocará en analizar las diferentes herramientas y técnicas que las empresas pueden utilizar para mejorar su presencia en línea, aumentar su visibilidad en los motores de búsqueda y obtener información valiosa sobre sus clientes a través de la recopilación y análisis de datos. También se explorará el uso de técnicas de aprendizaje automático para desarrollar modelos predictivos y recomendaciones personalizadas que puedan ayudar en la toma de decisiones.

En definitiva, este marco teórico desarrollara los temas basados en las estrategias de marketing digital y su relación con el uso de herramientas tecnológicas avanzadas. A través de la comprensión y el análisis de estos temas, se espera obtener una comprensión más profunda de cómo las empresas pueden aprovechar los beneficios de estas tecnologías emergentes.

4.1 Marketing Digital

El marketing digital es un conjunto de técnicas con medios electrónicos, que tienen como objetivo posicionar y visualizar la marca de una empresa con el propósito de tener un mejor posicionamiento respecto a la competencia (Karina Bricio Samaniego, 2018). En México cerca de un promedio de 70 millones de usuarios utilizan internet, el Considerando aún que México lleva varios años de atraso en el desarrollo de plataformas digitales y acceso a internet a comparación de otros países, en estos últimos años se ha reportado un incremento de

consumidores en el mercado digital, las empresas han optado por tener presencia online así habilitan los canales digitales como medio para estar más en contacto con sus consumidores y evitando tantos intermediarios en el proceso.

Las tecnologías de la información y comunicaciones se han integrado tanto a la rutina diaria de millones de personas alrededor del mundo que a tendencias como redes sociales y motores de búsqueda más eficientes. Esto da la oportunidad de tener acceso a una gran cantidad de datos /productos donde las empresas tienen un mayor alcance e influencia en sus mercados meta.

En 2019, México alcanzó un valor agregado Bruto del 5.8% del Producto Interno Bruto (PIB) que es proporcional a \$1, 462,583 pesos con una tendencia a la alza (ver Tabla 1. Comparativa entre los años y el incremento de aportación de valor agregado bruto en el comercio electrónico en México y Figura 1. Gráfica del crecimiento porcentual del Producto Interno Bruto en el comercio electrónico entre 2013 al 2019 en México.) (Inegi, 2021).

*Tabla 1.
años y el
de valor
comercio*

Año	Precios corrientes en pesos mexicanos
2013	481,232
2014	595,556
2015	662,367
2016	818,696
2017	1,013,258
2018	1,175,809
2019	1,462,583

*Comparativa entre los
incremento de aportación
agregado bruto en el
electrónico en México*

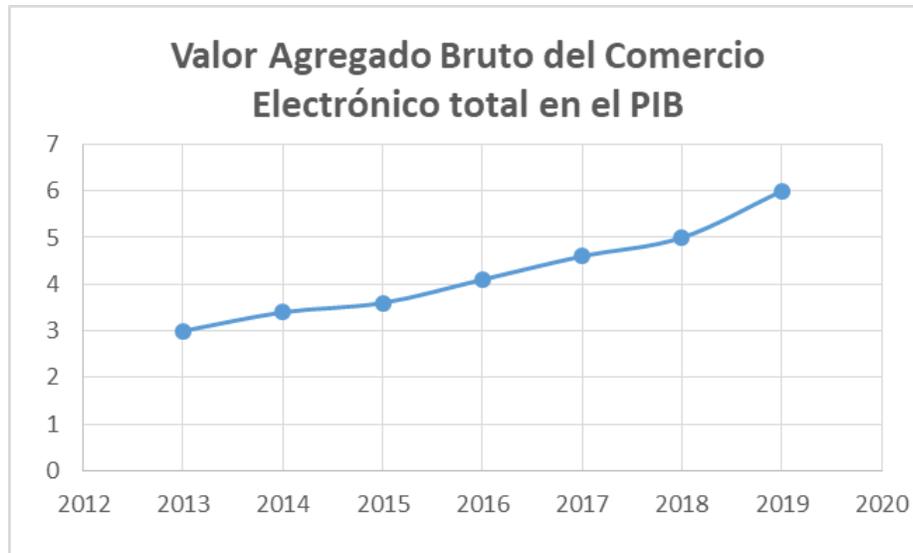


Figura 1. Gráfica del crecimiento porcentual del Producto Interno Bruto en el comercio electrónico entre 2013 al 2019 en México.

El mercado digital ha traspasado la barrera física y ha permitido que pequeños negocios puedan ser visibles y poder comerciar con un mercado que de otra manera no se tendría acceso, por lo que los negocios que no son visible en el mercado digital pierdan clientes potenciales (Pareja, 2021).

4.2 Posicionamiento en buscadores

Existen dos tipos de posicionamiento en los motores de búsqueda para lanzar resultados que trata del posicionamiento que se consigue a través de los resultados de búsqueda naturales de los motores de búsqueda. Es el más deseable, ya que es el que proporciona un mayor tráfico cualificado a una web.

SEO: El posicionamiento SEO es una técnica utilizada en Internet para mejorar el ranking de una página en los resultados de un buscador. La finalidad de mejorar el posicionamiento es conseguir que la página aparezca en las primeras posiciones de los resultados de búsqueda, lo que traerá un aumento en el tráfico web a la página.

SEM: El objetivo de SEM es mejorar el posicionamiento de una página web en los motores de búsqueda a través de la optimización de sus elementos para que aparezca en las primeras posiciones de los resultados cuando un usuario busque una palabra clave relacionada con su actividad a través de publicidad de pago.

El buscador considera que contenido sea valioso relevante, optimizado para los motores de búsqueda y los recursos del sitio web (Pareja, 2021). Para el posicionamiento en motores de búsqueda es crucial para lograr una mayor visibilidad y tráfico en una página web. Tanto hacer uso de SEO O SEM son técnicas importantes para el éxito de una estrategia de marketing digital y requieren de la optimización de contenido relevante y valioso, así como de recursos en la página web para lograr una mejor visibilidad en los resultados de búsqueda.

4.2.1 SEO

El SEO es uno de los aspectos más importante de la estrategia de marketing online porque es responsable de hacer que un contenido sea más visible y mantenerlo por delante de otros resultados. Mejora la visibilidad del sitio web y redes sociales, aumentar el tráfico de usuarios y llegar al mercado objetivo, aumenta la tasa de conversión entre interacciones-ventas y el optimiza el retorno de la inversión (ROI) de una campaña de marketing digital. Como parte de cualquier estrategia de marketing digital, la optimización de motores de búsqueda (SEO) es esencial para llevar los negocios a sus clientes a través de plataformas en línea. El marketing eficaz requiere una serie de herramientas y prácticas profesionales pueden ayudar a que el contenido tenga una clasificación más alta en los resultados de los motores de búsqueda, lo que generará más tráfico a su sitio web y posiblemente más negocios (Pareja, 2021).

Beneficios que ofrece SEO

- Bajo costo de implementación en los mejores puestos de búsqueda.
- Intento de compra.
- Rentabilidades a largo plazo al consolidarse entre los sitios más populares.

La solución más eficiente para procesar y clasificar datos es a través de clusters, una de las maneras es utilizando diferentes herramientas que se utilizan en la programación como: Python, Scala, Java o R son algunos ejemplos. Donde es posible utilizarlo de forma interactiva mediante los intérpretes, o utilizando otras herramientas de análisis donde ya están integradas como Google data studio, recursos accesibles gracias a las tecnologías de la información (Muhammad Ridwan Andi Purnomo, 2020).

Últimamente la mejor manera de posicionar un contenido relevante es a través de un posicionamiento SEO (Search Engine Optimization), este consiste en aparecer en los primeros resultados de un motor de búsqueda. Google es el destacado por ser el motor de búsqueda más

utilizado a nivel mundial, manteniéndose en una media del mercado de un 88% respecto a sus principales competidores (Yahoo, Baidu y Bing) entre enero del 2014 y enero del 2021 (statista, 2021).

Componentes de SEO

- El motor de búsqueda.
- El algoritmo de búsqueda.
- SERP (Search Engine Results Page).

El SEO, si se hace bien, es una de las formas más rentables de marketing. La audiencia estará buscando el producto o servicio en concreto una vez consolidado en un nicho. Aunque el SEO es una práctica con pasos concretos, las pautas y las técnicas de optimización se pueden comprender y aprender a través del “Centro de la Búsqueda de Google”. Integrar las técnicas de SEO con los requisitos de marketing es un gran desafío en el marketing digital.

Conocer las mejores prácticas para SEO (Pareja, 2021).

- Integrar técnicas de SEO, técnicas de marketing y machine learning.
- Para dirigir un contenido de nicho a un perfil demográfico adecuado en la etapa adecuada.
- Comprenda cómo operan los buscadores como Google y Bing para tener una ventaja competitiva.
- Conocer el uso y explotar el potencial de palabras clave y comprenda la importancia de usar la palabra clave correcta.
- Conocer las variables que influyen en una página web para optimizar cada aspecto de su contenido.
- Al crearse una nueva cuenta en Google analytics, se obtiene una matrícula (ID) de seguimiento. Después se debe insertar este código en la página web que se pretende analizar.
- Google analytics tiene la facilidad de establecer una interfaz donde se visualizan datos generados por los usuarios y se segmenta en: fechas, regiones geográficas e interacciones concretas.

En conclusión, el SEO y su importancia en la estrategia de marketing online es muy relevante en la actualidad. El SEO es responsable de hacer que el contenido sea más visible y mantenerlo por delante de otros resultados, lo que mejora la visibilidad del sitio web y redes sociales, aumenta el tráfico de usuarios y llega al mercado objetivo. Además, mejora la tasa de conversión entre interacciones, lo que una buena campaña tendrá un gran margen óptimo del retorno de la inversión (ROI). El rol de los motores de búsqueda es esencial para llevar los negocios a sus clientes a través de plataformas en línea, y el uso de herramientas sumado a buenas prácticas profesionales puede ayudar a que el contenido tenga una clasificación más alta en los resultados de los motores de búsqueda, lo que generará más tráfico a su sitio web y posiblemente más negocios.

La implementación de técnicas de SEO puede tener un bajo costo y rentabilidades a largo plazo al consolidarse entre los sitios más populares. Además, el uso de clusters es una de las maneras más eficientes de procesar y clasificar datos en SEO, y se pueden utilizar diferentes herramientas de programación como Python, Scala, Java o R. Integrar las técnicas de SEO con los requisitos de marketing es un gran desafío en el marketing digital, pero conocer las mejores prácticas y entender cómo operan los motores de búsqueda como Google y Bing pueden dar una ventaja competitiva.

4.2.2 Search quality rating

Los diferentes factores de posicionamiento que tiene google están correlacionados de manera directa, desde la página esta optimizada para una rápida descarga de contenido (Estructura, texto, imágenes, videos o cualquier otro tipo de contenido multimedia), si tiene una verificación de seguridad SSL y si este es relevante al igual que confiable.

Estos aspectos de medición se le conocen como "Page Quality Rating" que mide la calidad sobre cantidad de contenido, metadatos (datos de los datos) y la optimización técnica del sitio web, que es una interfaz limpia, sin contenidos demasiado pesados para cargar e interconectado su contenido entre si u otras páginas relevantes con links o hipervínculos, además de generar contenido constantemente.

La gestión de recursos, certificaciones de seguridad de la página web puede denominarse como el margen de confiabilidad que maneja esa página, por ejemplo, información académica: donde Wikipedia pierde relevancia para temas más especializados, ya que cualquier persona puede crear y alterar contenido dentro de esa plataforma, además de contar con una certificación una

certificación de seguridad SSL, que al introducir dato esta se ve cifrada por algún código o algoritmo (Pareja, 2021).

Esto se trata de uno de los factores más importantes en el posicionamiento de una página web en los resultados de búsqueda de Google. La calidad de una página web es un aspecto clave en el SEO y puede ser evaluada a través del Page Quality Rating, que toma en cuenta aspectos como la relevancia y confiabilidad del contenido, la optimización técnica del sitio web (El tiempo que tarde en cargar los recursos como imágenes u otros recursos multimedia) y la gestión de recursos.

Es importante destacar que Google utiliza evaluadores humanos para determinar la calidad de una página web, lo que significa que las páginas deben ser atractivas tanto para los usuarios como para los motores de búsqueda. Además, los criterios de calidad de Google están en constante evolución, por lo que es importante mantenerse actualizado en las mejores prácticas de SEO para asegurarse de que una página web esté optimizada y tenga una buena posición en los resultados de búsqueda.

4.2.3 El uso de hashtags para posicionamiento en redes sociales

La implementación de hashtags ha demostrado ser una herramienta efectiva para mejorar el posicionamiento de las publicaciones en redes sociales y, por ende, aumentar la visibilidad y el alcance de la marca, sin embargo, se debe realizar las siguientes consideraciones.

Al realizar un óptimo uso de hashtags es importante investigar los más populares dentro del sector de la marca. Esto permite conocer los temas más relevantes y los intereses de los usuarios en la red social.

Posteriormente, se debe identificar el contenido más popular asociado a esos hashtags. Esto proporcionará información sobre el tipo de contenido que funciona mejor y tener un mejor entendimiento de las tendencias del mercado.

Es importante determinar qué hashtags son más efectivos para llegar al público objetivo de la marca. En este sentido, se pueden utilizar herramientas de análisis de redes sociales para conocer los intereses de los seguidores.

Para lograr equilibrio entre los hashtags de mayor alcance contra los que son de nicho, ya que los hashtags populares tienen un mayor volumen de búsquedas, pero también son más

competitivos, mientras que los de nicho son más específicos, por lo que tienen menos competencia.

Es fundamental experimentar con diferentes combinaciones de hashtags para evaluar su efectividad. Buscando los hashtags que funcionan mejor que otros dependiendo del contenido, la red social y el público objetivo.

Por último, es importante controlar los resultados y ajustar la estrategia del contenido digital según sea necesario. Esto permitirá optimizar el uso de los hashtags y lograr un mayor alcance y engagement en las redes sociales.

4.3 Big data

En la actualidad, el big data es una herramienta clave para el éxito en el marketing digital. Gracias a esta tecnología, se pueden recopilar, analizar y utilizar grandes cantidades de datos en tiempo real, lo que permite tomar mejores decisiones en relación a la estrategia a seguir en marketing. Por ejemplo, el big data puede ayudar a identificar cuáles son los momentos en que los consumidores están más receptivos a los mensajes, cuáles son sus gustos y preferencias en cuanto a productos o servicios, o cuáles son las redes sociales en las que están más activos. De esta forma, el marketing digital se puede adaptar a las necesidades y preferencias de cada consumidor, lo que aumenta las posibilidades de éxito.

El big data es una tecnología con un auge reciente, que permite procesar y analizar grandes cantidades de datos para encontrar patrones de comportamiento. Esto requiere una poderosa capacidad de computación, que no estaba disponible hasta hace poco y que es esencial para el procesamiento y análisis de datos a gran escala.

El aprendizaje automático y el internet de las cosas son de las tecnologías que más se benefician del big data. Esta tecnología tiene como objetivo proporcionar a las computadoras capacidades de aprendizaje y comunicación a través de una programación de algoritmos altamente inteligentes. Este es un proceso de inducción de conocimiento, que se nutre de los datos generados para lograr la mejora continua del sistema.

Entre las innumerables aplicaciones del aprendizaje automático se encuentran el diagnóstico médico (como la detección de cáncer o la predicción de algunas enfermedades a través de los síntomas), los traductores automáticos (Ya sea por un dispositivo físico o en la nube) o la robótica

(sea capaz de analizar su entorno y lograr evitar obstáculos sin necesidad de ingresar comandos específicos para evitarlos) (Asociación de internet, 2020)

Las principales funcionalidades del big data son la capacidad de recolectar, almacenar, procesar y analizar una gran cantidad de datos en tiempo real. Estas funcionalidades se aplican en Hadoop (plataforma de software libre que proporciona una infraestructura para el análisis de datos masivos) mediante la utilización de frameworks como Spark, que permite ejecutar operaciones de procesamiento en paralelo sobre grandes conjuntos de datos.

En la estructura de Hadoop se conforma de una manera sencilla, consta de un núcleo y una interfaz de usuario. El núcleo de Hadoop es una colección de Java virtual machines (JVM) que ejecutan el código de Hadoop, mientras que la interfaz de usuario es una capa de abstracción que permite a los usuarios interactuar con los datos, además proporciona una distribución de datos segura y eficiente para el análisis de datos en paralelo. Análisis de datos en paralelo: que proporciona una infraestructura para el análisis de datos simultáneos y permite aprovechar al máximo la potencia de la computación en la nube y en la computación de centros de datos.

El software permite hacer uso de la computación en la nube y en la computación de centros de datos, facilita la administración de datos para gestionar y monitorizar el análisis de datos, incluye herramientas de minería de datos, análisis estadístico y visualización de datos (While, 2012). Además de contar con una amplia compatibilidad (Windows, Mac OS X y Linux).

4.3.1 Plataformas de recolección y análisis de datos

El estudio del patrón de los datos es fundamental para la mercadotecnia digital, ya que a través de ellas nos facilita la tarea de analizarlos de manera automática la a través de la información que generan las acciones de los consumidores en Internet y tomar decisiones acertadas en materia de estrategias de marketing. Gracias a estas herramientas, se pueden identificar las tendencias en el comportamiento de los usuarios en línea, así como las preferencias y necesidades de estos últimos. De esta forma, se pueden generar campañas publicitarias más efectivas y orientadas al público objetivo.

Las plataformas de análisis de datos son fundamentales para entender el comportamiento de los consumidores en Internet. Es posible correlacionar con sus intereses, qué productos o servicios prefieren, predecir cuándo están dispuestos a comprar y mucho más. De esta forma, las empresas pueden elaborar estrategias de marketing más efectivas, dirigidas a los segmentos de

mercado que presentan mayores posibilidades para concretar una venta; Entre las diferentes plataformas de análisis de datos líderes del mercado se encuentran los siguientes:

4.3.2 Azure y Power BI de Microsoft

Microsoft Azure es una plataforma de computación en la nube que ofrece una amplia gama de servicios en la nube, incluidos servicios en la nube como son de IaaS (Infraestructura como servicio), PaaS (Plataforma como servicio) y SaaS (software como servicio). Azure también incluye una serie de herramientas analíticas que le permiten analizar y visualizar datos en tiempo real y a su vez ofrece una variedad de herramientas analíticas, incluidas herramientas de minería de datos, análisis de datos en tiempo real, análisis predictivo y visualización de datos, además es una plataforma con una gama amplia de servicios donde se cotiza una anualidad dependiendo de la personalización de los servicios que se lleguen a utilizar (Azure, 2021).

Al igual que Azure, Power BI es una herramienta analítica de Microsoft que permite visualizar y analizar datos en tiempo real. Además, Power BI ofrece una serie de funcionalidades que permiten mejorar la productividad a la hora de analizar datos, como la posibilidad de crear gráficos y tablas dinámicas, o de realizar consultas en tiempo real a través de una interfaz intuitiva (BI, 2021).

La diferencia entre estas herramientas analíticas que ofrece Microsoft es que Azure es una plataforma de nube pública de Microsoft que ofrece una variedad de servicios de computación en la nube, almacenamiento en la nube, plataformas de aplicaciones y redes. Power BI es un servicio de análisis de datos en la nube que permite a los usuarios analizar y visualizar datos de forma interactiva (Azure, 2021).

Azure ofrece una variedad de herramientas para analizar datos, incluyendo Azure Data Lake Store, Azure Data Lake Analytics, Azure Machine Learning, Azure SQL Data Warehouse, Azure Stream Analytics y Azure Data Factory. Power BI proporciona una herramienta para analizar datos y presentar informes, ambas tienen funcionalidades similares en el análisis de datos, como la capacidad de importar datos de fuentes externas, realizar un análisis de datos en tiempo real y generar reportes. Sin embargo, tiene una funcionalidad adicional que permite visualizar datos en un mapa.

Azure modela los experimentos creados en el área de trabajo de Azure Machine Learning y representa la forma principal de habilitar un enfoque iterativo para desarrollar rápidamente

soluciones de Azure Machine Learning. En cada experimento de Azure Machine Learning, se testea y brinda un espacio de trabajo visual interactivo para crear, probar e iterar fácilmente sus experimentos de análisis predictivo. Estos experimentos se pueden enviar a Azure ML Studio para que se ejecuten. Los experimentos de Azure Machine Learning son altamente iterativos. Puede crear, editar, probar, guardar y volver a ejecutar modelos de información fácilmente.

El uso de Azure Machine Learning Experiments está diseñado específicamente para permitir que los científicos de datos modernos de hoy en día "fallen rápidamente" al evaluar nuevos modelos predictivos, al tiempo que brindan la capacidad de mejorar los comentarios de los modelos predictivos para mejorarlos.

Azure ML Studio Es el principal base de trabajo del análisis predictivo interactivo al que se accede desde un espacio de trabajo de Azure Machine Learning para permitir al analista de datos crear experimentos de Azure Machine Learning mediante una interfaz de diseño visual de arrastrar y soltar. El acceso a un entorno único de Azure ML Studio se rige desde un espacio de trabajo de Azure Machine Learning.

Además de permitir la creación de nuevos experimentos, Azure ML Studio también tiene enlaces a experimentos de muestra de Azure Machine Learning. Estos se proporcionan para que pueda aprender fácilmente de otros a medida que avanza en su viaje de ciencia de datos y aprovechar algunas de las mejores técnicas de procesamiento y herramientas de la industria para ayudar a lograr sus objetivos específicos de análisis predictivo (Barnes, 2015).

Estas plataformas son unas excelentes opciones cuando se quiere migrar la información (por ejemplo base de datos, cartera de clientes o gastos realizados dentro de un periodo específico) de una empresa a la nube y máquinas virtuales, con la ventaja de que todos los colaboradores de la empresa puedan tener acceso a ella sin necesidad de estar en las instalaciones laborales, además de contar con funciones avanzadas e inteligencia artificial para gestionar la información de manera eficiente con Azure functions.

El manejo de la base de datos en el marketing digital es parte importante en el seguimiento de clientes, desde conocer el estatus que tiene el pedido del cliente e incluso poder seguir ofreciendo servicios una vez que se concluya la venta con el cliente.

También es utilizado para la toma de decisiones creando un entorno que facilita la visualización de la información en tiempo real, para realizar un seguimiento oportuno de los indicadores y

metas que la empresa se proponga. En el caso del marketing digital es importante conocer la conversión de visitas en ventas y considerar el retorno de inversión para destinar un presupuesto a una campaña digital.

4.3.3 AWS: Amazon Web Services

AWS (Amazon Web Services) es una plataforma de nube pública que proporciona una serie de servicios de computación en la nube, almacenamiento en la nube, bases de datos en la nube, análisis en la nube, gestión de contenidos en la nube, invocación de servicios en la nube y redes en la nube. AWS es un servicio de Amazon.com, Inc.

AWS proporciona una serie de herramientas de análisis que le permiten medir el rendimiento de sus aplicaciones y servicios en la nube. Además, AWS ofrece una gran gama de funcionalidades dentro de su servicio de análisis que le permiten medir el rendimiento de sus aplicaciones y servicios en la nube. Algunos de estos servicios son: Amazon CloudWatch, Amazon CloudTrail, Amazon Inspector, Amazon Macie y Amazon Kinesis Firehose (AWS, 2021).

Amazon web Services ofrece un año de uso gratuito, este brinda muchos servicios y la flexibilidad de experimentar y de ver cuáles servicios son los que mejor se adaptan a las necesidades comerciales o personales dentro de la amplia variedad de servicios que cubre la capa gratuita de servicios algunos tienen diferentes tipos de límites.

Amazon elastic Cloud computing consiste básicamente en iniciar sesión dentro de entornos informáticos virtuales en un servidor conocidos Como instancias; se utilizan 750 horas de instancias de uso mensual gratuito.

Amazon web Services Todas y cada una de las cuentas nuevas de AWS, Eso no significa que ese límite estas cuentas para adquirir otros servicios por eso es muy importante tener en cuenta que no todos los servicios están cubiertos por dicha capa gratuita que ofrece AWS

Se puedes escoger servidores de diferentes opciones de disco duro y puede subirla a este servidor y el servidor puede subir también aplicaciones en cualquier lenguaje

Existen más de 100 productos que ofrece Amazon Web Services de los cuales su aplicación de machine learning en el marketing digital podría beneficiarse. Algunos de estos servicios son:

- Amazon Elastic Compute Cloud (EC2): Ofrece capacidad de computación en la nube, escalabilidad y una gran variedad de opciones de software.
- Amazon Simple Storage Service (S3): Proporciona almacenamiento en la nube seguro, escalable y asequible.
- Amazon DynamoDB: Es una base de datos NoSQL en la nube que proporciona una alta disponibilidad, escalabilidad y fiabilidad.
- Amazon Elastic MapReduce (EMR): Proporciona una manera sencilla de usar Hadoop (entorno de trabajo opensource para trabajar con grandes cantidades de datos) en la nube para analizar grandes cantidades de datos.
- Amazon Lex: permite diseñar y crear chatbots para automatizar el servicio al cliente a través de mensaje o voz.
- Amazon Machine Learning: Ayuda a los desarrolladores a crear modelos de aprendizaje automático para mejorar el rendimiento de las aplicaciones.
- Amazon CloudFront: Proporciona una red de distribución de contenido en la nube.
- Amazon CloudWatch: Monitorea y analiza el rendimiento de aplicaciones y servicios en la nube.

Las metas en el marketing digital pueden lograrse a través de Amazon Web Services, que facilita su accesibilidad a través de la computación en la nube que ofrece servicios de computación, almacenamiento, bases de datos, redes y aplicaciones. AWS es una excelente manera de escalar el poder de cómputo y almacenamiento para satisfacer las necesidades de una empresa en crecimiento. AWS también ofrece una amplia gama de servicios de seguridad que ayudan a proteger las aplicaciones y datos (AWS, 2021).

4.3.4 Google analytics

Google Analytics es una herramienta gratuita de análisis web que ofrece información sobre el tráfico de un sitio web y sus visitantes. La herramienta proporciona datos sobre el origen de las visitas, el comportamiento de los visitantes en el sitio, el número de páginas vistas, el tiempo que permanecen en el sitio, etc.

En esta plataforma se muestran los datos que permite a los administradores detrás del sitio web medir el éxito de su estrategia de marketing digital. Principalmente ofrece la información sobre el tráfico web, el comportamiento de los usuarios y el rendimiento de las campañas de marketing digital. Además, Google Analytics también permite crear informes personalizados y analizar el rendimiento de las páginas web.

La solución que propone google analytics para las páginas web es que se puedan ser vincular y estén disponibles online, se obtiene todo tipo de información sobre las visualizaciones e interacciones que tuvieron los usuarios con la página web, esta herramienta ofrece una gran cantidad de información - métricas para estimar la actividad y relevancia del contenido (Orlando Troisi, 2019).

La vinculación entre google analytics y algunas plataformas que facilita el desarrollo de contenido web es muy accesible, utilizando un código que se genera en google analytics, los datos recolectados tienen la finalidad de analizarse comprender la interacción de los usuarios y optimizar la información que generan las visitas.

La información de los usuarios también muestra que características tienen (sexo, ubicación y si son visitantes frecuentes), por lo que además plantea el público objetivo que consume el contenido Esta información puede utilizarse para mejorar el diseño y el contenido del sitio web y para optimizar las campañas de marketing.

Google Analytics es gratuito y de interfaz amigable. Es una herramienta que se puede integrar con otros servicios de Google, como AdWords y Google Search Console. También se puede integrar con otros servicios de terceros, como MailChimp, Optimizely y Salesforce. De igual manera se correlaciona con otras plataformas que tiene google como sería google console search (Es el panel de administración de búsqueda de google, donde se pueden gestionar las propiedades de búsqueda, analizar el rendimiento de las campañas de búsqueda, y realizar cambios en el ranking de las webs) y google data studio (una herramienta de análisis que permite visualizar los datos de google analytics, YouTube o Adwords en forma para crear reportes y gráficos de la actividad en línea).

Para utilizar los informes de Google Analytics en la mercadotecnia digital, debe registrarse para una cuenta de Google Analytics y agregar el código de seguimiento en todas las páginas del sitio web. Luego, debe analizar los informes de Google Analytics para identificar las áreas del sitio web y las campañas de marketing que funcionan mejor y las que necesitan mejoras.

4.4 Inteligencia artificial

La inteligencia artificial es una herramienta que forma parte de la rama de la inteligencia computacional, cada vez más trascendental en la vida laboral y cotidiana (como la medicina, la industria, la agricultura, la ingeniería, el transporte, etc). Gracias a ella, se pueden analizar grandes cantidades de datos en poco tiempo para tomar mejores decisiones en cuanto a estrategias de marketing. Además, la inteligencia artificial puede ayudar a personalizar las campañas de marketing en función de las necesidades y preferencias de los consumidores.

Las principales tecnologías y herramientas que se utilizan para el desarrollo de la inteligencia artificial son:

- Tecnologías de la información: bases de datos, inteligencia artificial, minería de datos, redes neuronales.
- Herramientas de software: lenguajes de programación, entornos de desarrollo.
- Hardware: ordenadores, servidores, dispositivos de entrada y salida.

La inteligencia artificial está basada en imitar la inteligencia humana utilizando técnicas de robótica y software para emularlo. Los autores Stuart Russell y Peter Norvig (Norvig, 2009) distinguieron cuatro tipos de inteligencias en 2009: sistemas que piensan como humanos, como redes neuronales artificiales. Sistemas que funcionan como humanos. Sistemas que usan lógica racional como se interpreta en los sistemas expertos y sistemas que usan un comportamiento racional, capaz de tomar decisiones a partir de datos históricos o empíricos. (Serrano Sandoval, 2018)

Los algoritmos de inteligencia artificial se pueden clasificar en dos grandes grupos:

- Los basados en la estadística, que buscan la optimización de un objetivo a través de la evaluación de un gran número de posibilidades.
- Los basados en la lógica, que buscan la resolución de problemas a través de la deducción de reglas.

A continuación, se describen algunos de los algoritmos de inteligencia artificial más comunes:

4.4.1 Algoritmo de aprendizaje automático

El aprendizaje automático está basado el uso de múltiples algoritmos con el fin de encontrar un patrón en los datos y realizar proyecciones para mejorar su rendimiento a medida que adquieren más información. Este enfoque se utiliza en una amplia variedad de aplicaciones, incluyendo la clasificación de correos electrónicos no deseados, la detección de fraudes en transacciones financieras, la clasificación de imágenes, la recomendación de productos y servicios, y la traducción automática de idiomas, entre otras.

Existen diferentes tipos de algoritmos de aprendizaje automático, como los algoritmos supervisados, no supervisados y de refuerzo. Los algoritmos supervisados se utilizan para predecir valores de salida a partir de datos de entrada ya establecidos previamente, mientras que los algoritmos no supervisados se utilizan para identificar patrones o grupos en los datos de entrada no definidos. Los algoritmos de refuerzo se utilizan para tomar decisiones secuenciales a medida que reciben retroalimentación de su entorno, por lo que en ese proceso se modifica el algoritmo a uno que se adapte de manera más óptima.

El aprendizaje automático se basa en conceptos matemáticos y estadísticos como la optimización, la regresión, la clasificación y la agrupación. A medida que los algoritmos de aprendizaje automático se vuelven más sofisticados, también se están explorando nuevas técnicas como el aprendizaje profundo y las redes neuronales, que tienen la capacidad de procesar grandes cantidades de datos para extraer patrones más complejos. El aprendizaje automático es una herramienta basada en experimentación, de constante pruebas que permite su evolución, esto le da un gran potencial de transformar la forma en como las organizaciones abordan los desafíos de manera más óptima. (Serrano Sandoval, 2018)

4.4.2 Algoritmo de redes neuronales

El desarrollo de las redes neuronales es capaz de aprender y reconocer patrones en los datos de entrada, lo que las convierte en una herramienta valiosa en campos como el reconocimiento de imágenes y el procesamiento del lenguaje natural. Estas redes están formadas por capas de nodos que procesan la información y la transmiten a través de conexiones interconectadas. El aprendizaje se produce a través de la adaptación de los pesos de las conexiones en función de los resultados obtenidos en la tarea que se está realizando.

Existen varios tipos de redes neuronales, como las redes neuronales de alimentación directa, las redes neuronales recurrentes y las redes neuronales convolucionales. Cada tipo de red neuronal

se adapta mejor a un tipo específico de tarea, y su eficacia depende en gran medida de la calidad y cantidad de los datos de entrenamiento.

En la actualidad, las redes neuronales son ampliamente utilizadas en aplicaciones de reconocimiento de voz, reconocimiento de imágenes, procesamiento del lenguaje natural, robótico y control de procesos, entre otras. Además, el uso de algoritmos de redes neuronales está en constante evolución y mejora, lo que hace que su potencial en diversos campos sea cada vez mayor (Larranaga, 1997).

4.4.3 Algoritmo de razonamiento lógico

Los algoritmos de razonamiento lógico son un tipo de algoritmo que se basa en la deducción de reglas para resolver problemas. La teoría de la lógica proposicional sostiene que cualquier problema se puede descomponer en una serie de proposiciones simples y que estas proposiciones se pueden combinar para llegar a una solución. Los algoritmos de razonamiento lógico utilizan esta teoría para establecer reglas y restricciones (constantes y límites) de un problema.

Estos algoritmos se utilizan en diversas áreas, principalmente en las diferentes técnicas aplicadas a la inteligencia artificial e informática teórica. En la inteligencia artificial, los algoritmos de razonamiento lógico suelen usarse para construir sistemas expertos, que pueden tomar decisiones y proporcionar soluciones en áreas específicas. En la informática teórica, estos algoritmos se utilizan para analizar y probar la corrección de los programas.

Los algoritmos de razonamiento lógico suelen ser utilizados en situaciones donde se requiere una respuesta precisa y lógica. Por ejemplo, en el campo de la medicina, estos algoritmos se utilizan para analizar los síntomas de un paciente y llegar a un diagnóstico preciso. También se utilizan en sistemas de seguridad informática para analizar y prevenir posibles amenazas. En resumen, los algoritmos de razonamiento lógico son una herramienta importante en la resolución de problemas complejos para casi cualquier campo de estudio.

4.4.4 Algoritmo de optimización

Los algoritmos de optimización se utilizan para buscar la solución óptima a un problema. Estos algoritmos se basan en la teoría de la optimización, que afirma que todos los problemas se pueden reducir a una serie de variables y parámetros.

La inteligencia artificial tiene el potencial de cambiar significativamente nuestra vida diaria. Por ejemplo, podría mejorar significativamente nuestro trabajo al ayudarnos a realizar tareas de forma más eficiente. También podría mejorar nuestra vida social al permitirnos comunicarnos con otros de forma más efectiva. Además, podría mejorar nuestro estilo de vida al ayudarnos a mantenernos saludables y activos.

Por otra parte las inteligencias artificiales son una amenaza para algunos empleos, pero también son una oportunidad para mejorar y automatizar procesos. Según un informe de la consultora PwC México, las inteligencias artificiales serán una de las diez tecnologías más disruptivas en 2020. (pwc, 2022)

Las empresas están empezando a implementar soluciones de inteligencia artificial para mejorar la eficiencia y la productividad. Según un estudio de la consultora Capgemini, el 63% de las empresas ya están utilizando o planean utilizar inteligencia artificial en el corto plazo.

4.5 Técnicas de machine learning

Las técnicas de aprendizaje automático son un conjunto de métodos matemáticos que permiten que una computadora aprenda a realizar tareas complejas, como reconocer imágenes, clasificar objetos o predecir el comportamiento de sistemas complejos, a partir de datos.

Las técnicas de aprendizaje automático se dividen en tres categorías principales:

Supervisado: En esta categoría, el aprendizaje se realiza a partir de un conjunto de datos de entrada que se conocen a priori y en los que se incluyen las respuestas correctas. El objetivo es que la computadora aprenda a identificar patrones en estos datos y a generar las respuestas correctas.

No supervisado: En esta categoría, el aprendizaje se realiza a partir de un conjunto de datos de entrada que no incluyen las respuestas correctas. El objetivo es que la computadora aprenda a identificar patrones en estos datos y a generar las respuestas correctas.

Semi-supervisado: En esta categoría, el aprendizaje se realiza a partir de un conjunto de datos de entrada, en el que se incluyen tanto las respuestas correctas como las incorrectas. El objetivo es que la computadora aprenda a identificar patrones en estos datos y a generar las respuestas correctas.

Las técnicas de aprendizaje automático se pueden aplicar a una gran variedad de problemas, como el reconocimiento de caras, la clasificación de objetos, el análisis de textos, la predicción del comportamiento de sistemas complejos o menos predecibles. (Daniel Santos, 2020)

4.5.1 Aprendizaje supervisado

En el aprendizaje supervisado en machine learning aplicado al marketing digital, el ordenador aprende a reconocer patrones en una gran cantidad de datos, como por ejemplo el historial de las interacciones de un usuario con una marca en una red social, para después poder predecir qué acciones pueden resultar en una mejor respuesta de ese usuario.

Se utilizan conjuntos de datos conocidos para crear un modelo que luego puede hacer predicciones. Los conjuntos de datos conocidos se denominan e incluyen elementos de datos de entrada junto con valores de respuesta conocidos.

El aprendizaje supervisado es una técnica de aprendizaje automático que requiere la existencia de una serie de ejemplos de entrenamiento, o datos de entrada, junto con una serie de respuestas correctas o resultados deseados, o datos de salida. La tarea del aprendizaje supervisado es aprender a predecir las respuestas correctas a partir de los datos de entrada.

Existen diferentes modelos de aprendizaje supervisado, cada uno con sus propias características. Los modelos de aprendizaje supervisado más comunes son el modelo de regresión lineal, el modelo de redes neuronales, el modelo de árboles de decisión y el modelo de support vector machines (SVM).

Los modelos de aprendizaje supervisado se pueden utilizar para resolver problemas de clasificación y de predicción. Los problemas de clasificación consisten en determinar a qué clase pertenece un objeto determinado, mientras que los problemas de predicción consisten en predecir el valor de una variable dependiente a partir de una serie de valores de una o más variables independientes.

El aprendizaje supervisado se ha utilizado con éxito en la inteligencia artificial para resolver problemas como el reconocimiento de caras, el reconocimiento de objetos, la clasificación de textos, la predicción de la demanda de energía eléctrica y la predicción del precio de las acciones.

A partir de estos conjuntos de datos de entrenamiento, los algoritmos de aprendizaje supervisado de datos, los algoritmos de aprendizaje supervisado intentan construir un nuevo modelo que pueda hacer predicciones basadas en nuevos valores de entrada junto con el aprendizaje supervisado puede dividirse en dos categorías generales de algoritmos. (Barviera, 2017)

4.5.2 Aprendizaje no supervisado

En machine learning aplicado al marketing digital, el aprendizaje no supervisado se refiere a la capacidad de una máquina para aprender a través de datos históricos, sin la intervención humana. Este tipo de aprendizaje es particularmente útil en el contexto del marketing digital, ya que permite a las máquinas identificar patrones en grandes cantidades de datos, lo que les permite realizar predicciones sobre el comportamiento de los consumidores y tomar decisiones estratégicas en base a esas predicciones.

Este proceso se basa en la capacidad de la máquina para identificar patrones en los datos y extrapolarlos a situaciones nuevas. El aprendizaje no supervisado es una técnica que se utiliza como alternativa en el aprendizaje automático, y tiene la ventaja de que no requiere la intervención humana para identificar y corregir errores. Sin embargo, tiene el inconveniente de que la máquina puede aprender erróneamente y no ser capaz de generalizar los patrones a nuevas situaciones. (Tello, 2007)

4.5.3 Árbol de Decisiones

El árbol de decisiones es una técnica incorporada en el desarrollo metodológico y algunos Sistemas de razonamiento para la investigación en inteligencia artificial, la programación en informática y su aplicación, por su estructura, en su interpretación en comprensión y análisis, en su uso diario se puede proporcionar desde un diagnóstico médico, pronóstico del tiempo, control de calidad, etc. Procesos que requieren análisis de datos y toma de decisiones. Predice valores objetivos de una función de una serie de variables predictivas (Calancha, 2021).

Las principales fortalezas de usar árboles de decisión en un proceso de análisis que permite analizar rápidamente las consecuencias de todas las decisiones posibles. Ya que se muestra de

manera gráfica en su arquitectura y cuantificar el costo de los resultados y la posibilidad de diferentes resultados a tomar la decisión del mejor escenario probable (ver Figura 2. Esquema de un árbol de decisiones, de una problemática se plantean las diferentes decisiones y los diferentes beneficios probables de cada uno. **¡Error! No se encuentra el origen de la referencia.**).

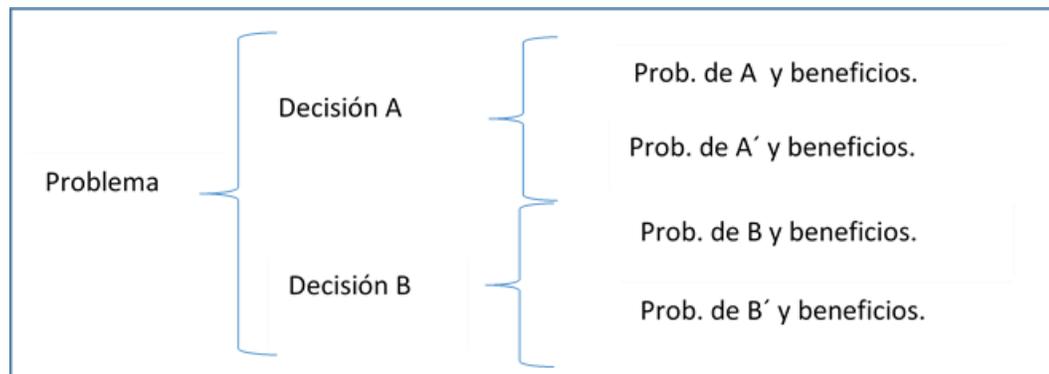


Figura 2. Esquema de un árbol de decisiones, de una problemática se plantean las diferentes decisiones y los diferentes beneficios probables de cada uno.

El árbol de decisiones es de utilidad en la mercadotecnia digital porque ayuda a tomar decisiones con base a criterios preestablecidos. De esta forma, se reducen las posibilidades de error y se asegura que las medidas que se tomen con la información disponible. Además, el árbol de decisiones facilita la visualización de múltiples alternativas a un problema, ya que agrupa las opciones en categorías y permite que cada miembro del grupo de trabajo comprometidos.

4.5.4 Regresión

La regresión lineal es una técnica de modelado estadístico que se utiliza para describir una variable de respuesta continua en función de una o más variables predictores. Particularmente útil para comprender y predecir el comportamiento de sistemas complejos o analizar datos experimentales, financieros y biológicos. La función de regresión lineal le permite crear modelos lineales. Esto para describir la relación entre variables dependientes contra variables independientes y obtener predicciones a través de los patrones que se obtienen (Ma. Jiménez, 2000).

Existen varios tipos de regresión lineal (Mathworks, 2021)

- Regresión lineal simple: modelado de un vector.

- Regresión lineal múltiple: ingresa más vectores entre la dispersión de factores.
- Regresión lineal multivariable: múltiples resultados “y”
- Regresión lineal múltiple multivariable: Múltiples resultados ordenados en forma matricial.

En la práctica: si Y es una variable definida en la misma población que X, es posible determinar si existe alguna relación entre las modas de X e Y gráficamente, un diagrama de dispersión o nube de puntos nos permite obtener información sobre el tipo de relación entre X e Y y es útil para detectar posibles valores atípicos o extremos.

El coeficiente de correlación determina el grado de asociación lineal entre X e Y, en lugar de determinar a priori cualquier direccionalidad de la relación entre las dos variables. En cambio, la regresión lineal simple permite cuantificar cómo varía el nivel promedio de la variable Y con la variable X, asumiendo implícitamente que X es la variable explicativa o independiente e Y es la variable de respuesta o dependiente, Se estiman los parámetros a y b en la creación de la pendiente donde a es el intercepto y b es la pendiente de X como predictor (Figura 3. Regresión lineal con una variable dependiente y una variable independiente.).

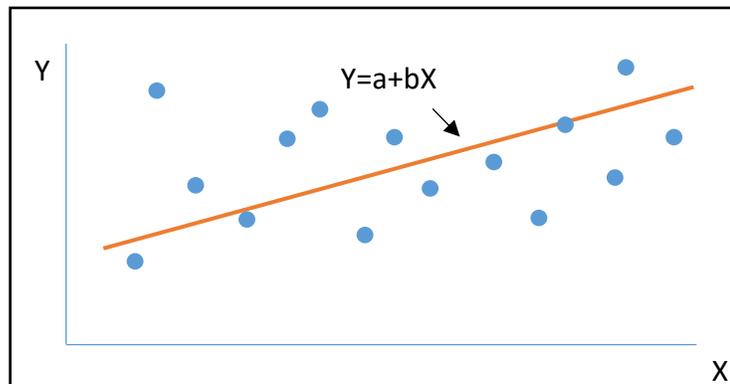


Figura 3. Regresión lineal con una variable dependiente y una variable independiente.

Cuando el coeficiente de correlación es significativo ($p < 0,05$), es decir, que existe una relación entre las dos variables.

Para validar la covarianza en una regresión lineal, se debe comprobar que la matriz de covarianza es diagonal, es decir, que todas las correlaciones entre las variables son iguales a

cero. Si esto no es así, significa que hay algún tipo de correlación entre las variables, lo que invalidaría la regresión lineal.

En mercadotecnia electrónica, la regresión lineal se usa para predecir el consumo de un producto o servicio basado en el comportamiento pasado de los consumidores. Por ejemplo, una empresa podría usar la regresión lineal para estimar el número de visitantes únicos a su sitio web basándose en el número de visitas registradas en el pasado para poder realizar proyecciones a futuro (Laguna, 2020).

4.5.5 Redes neuronales

Una red neuronal es un modelo simplificado que comprende el funcionamiento de como el cerebro humano procesa la información: el modelo funciona combinando una gran cantidad de unidades de procesamiento (datos de entrada) interconectadas que parecen versiones abstractas de neuronas interconectadas entre sí por módulos.

La primera red neuronal artificial fue creada en 1951 por Marvin Minsky y Dean Edmonds bajo el nombre SNARC (Stochastic Neural Simulation Re-Information Computer), la cual fue diseñada para imitar a una rata cuando aprende a salir de un laberinto. En 1952, Arthur Samuel se propuso aprender a jugar a las damas a partir de sus propios errores, y en 1957 diseñó una red neuronal que podía reconocer sistemáticamente a los personajes de una partida de damas. La primera red neuronal artificial que pudo aprender por sí misma fue creada en 1969 por Robert Hecht-Nielsen (Maisueche Cuadrado, 2019)

Las redes neuronales artificiales son una técnica informática para simular el funcionamiento de los cerebros humanos y animales. Se basan en el modelo de una neurona, que es una célula nerviosa que recibe información del mundo exterior a través de sus dendritas y la procesa para emitir una señal a través de su axón. La red neuronal artificial consta de una serie de estas neuronas, cada una conectada a otras a través de una sinapsis. La función de la red neuronal artificial es la de recibir información de una entrada, procesarla y emitir una respuesta.

Las unidades de procesamiento están organizadas en capas. Una red neuronal generalmente consta de tres partes: una capa de entrada, donde las unidades representan campos a estudiar; una o más capas ocultas de procesamiento con funciones o modelos que transforman los datos; y una capa de salida, donde una o más unidades representan campos de destino. Las unidades están conectadas por diferentes fuerzas de conexión (ver Figura 4. Representación de una

distribución de nodos en redes neuronales.4). De todas las variables se arroja un resultado predictivo, considerando todas las variables y este se va ajustando dependiendo de las necesidades (IBM, 2021).

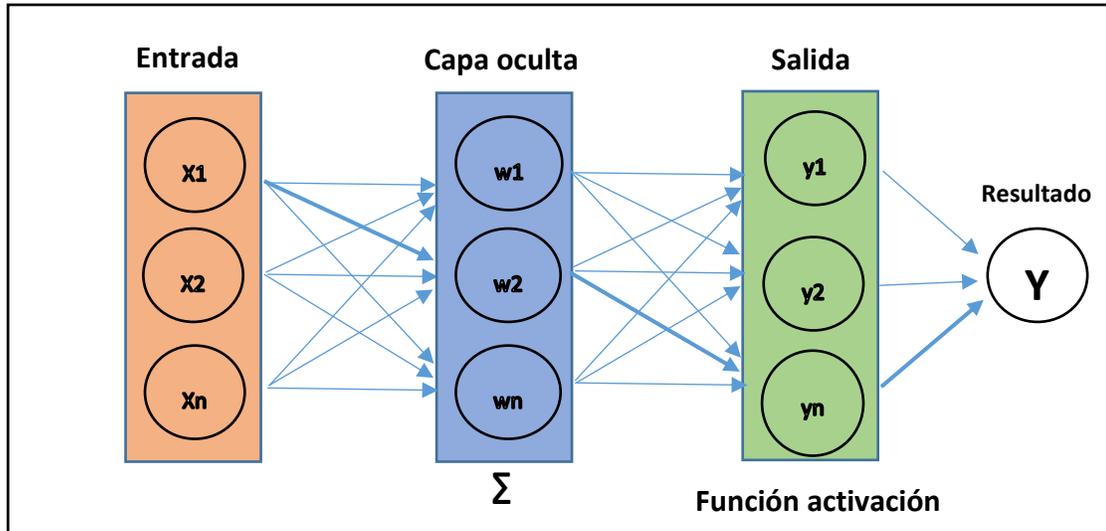


Figura 4. Representación de una distribución de nodos en redes neuronales.

Las redes neuronales artificiales se pueden clasificar en tres tipos:

- Redes neuronales simples: son redes formadas por una sola capa de neuronas.
- Redes neuronales multilayer: son redes formadas por varias capas de neuronas.
- Redes neuronales recurrentes: son redes formadas por neuronas conectadas en forma de cíclica.

Las redes neuronales artificiales (RNA) se pueden utilizar para entrenar modelos predictivos, que pueden ser aplicados en la mercadotecnia digital para mejorar la efectividad de las campañas publicitarias. Por ejemplo, se pueden utilizar para identificar patrones de comportamiento en los usuarios, y así desarrollar estrategias de marketing más efectivas.

4.5.6 Algoritmo (K-NN) o vecino más cercano

El algoritmo KNN es uno de los algoritmos de clasificación más prácticos y aun puede dar resultados muy completos que aportan al análisis de variables. Este modelo es considerado

como parte del campo del aprendizaje supervisado este se puede utilizar para el reconocimiento de patrones, la extracción de datos y la detección de intrusiones.

El modelo K-NN selecciona un valor deseado o buscado, y empieza una correlación con los datos más cercanos al valor de K creando un clúster de datos objetivos para aprender a distinguir entre pares de objetos en función de la proximidad en el espacio de variables. Para ello, se calcula la distancia entre cada objeto y todos los demás objetos de la muestra. Los objetos más próximos se clasificarán en la misma categoría (Figura 5. Representación gráfica del funcionamiento del modelo K-NN.5).

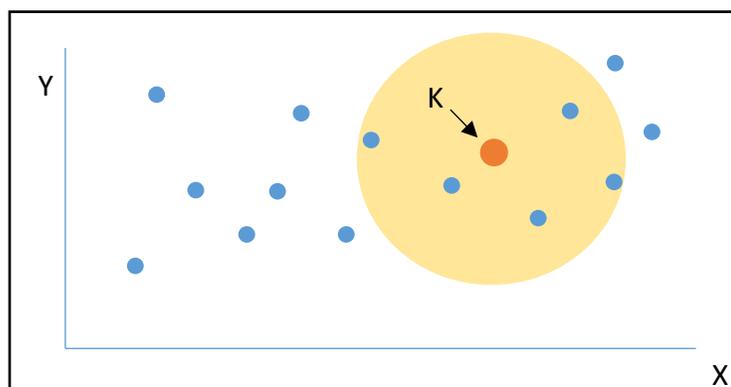


Figura 5. Representación gráfica del funcionamiento del modelo K-NN.

Es un clasificador potente y versátil, que a menudo se utiliza como punto de referencia para clasificadores más complejos, como redes neuronales artificiales y vectores de soporte. K-NN puede superar a los clasificadores más potentes y se puede utilizar en diversas aplicaciones, como la previsión económica, la compresión de datos y la genética.

El algoritmo K-NN se puede aplicar en la mercadotecnia digital de dos maneras:

1. Para predecir el comportamiento de los clientes en función de su historial de compras.
2. Para personalizar las recomendaciones de productos y servicios en función de las preferencias de cada cliente.

En conclusión, el algoritmo K-NN es una técnica simple pero efectiva en el campo del aprendizaje automático supervisado. El principio básico del K-NN es encontrar los k vecinos más cercanos a un punto de datos y utilizar sus etiquetas para predecir la etiqueta de ese punto. El valor de k y la medida de distancia son los dos parámetros clave que influyen en el rendimiento del modelo K-NN. (Moreno, 2006)

El algoritmo K-NN se puede utilizar en una variedad de aplicaciones, incluyendo clasificación de imágenes, reconocimiento de voz, análisis de sentimientos y recomendación de productos. Si bien el algoritmo K-NN tiene sus limitaciones, es una herramienta valiosa en el conjunto de herramientas de aprendizaje automático y puede ser un buen punto de partida para problemas de clasificación simples.

4.5.7 Support Vector machines

Las support vector machines (SVMs, por sus siglas en inglés) son un tipo de algoritmo de aprendizaje automático que se utiliza para clasificar datos. Funcionan a partir de una función de error, que es el objetivo que se busca minimizar. La SVM busca dividir los datos en dos clases lo más finamente posible, de forma que la función de error sea lo más baja posible para cada una de las clases.

Parte de los modelos consiste en la separación de las diferentes clases en las variables de estudio, los datos que se encuentran en un rango de margen o una tendencia son partidos por un vector que los separa y diferencia entre los diferentes grupos de datos, para SVM es un algoritmo iterativo. Esto significa que se puede modificar la malla (un conjunto de puntos que se sitúan en el espacio de los datos) a medida que se van obteniendo nuevos datos. De esta forma, se puede mejorar la clasificación de los datos.

Apreciar esa diferencia de datos separándolos en grupos sirve para estimar en su análisis una implementación de correlaciones entre factores, el modelo más común considerado es la regresión no lineal. Usar SVM en marketing digital incluyen su uso para identificar y dirigirse a clientes potenciales, predecir el comportamiento de los clientes y mejorar las estrategias de marketing (Meyer, 2015).

4.5.8 Modelo bayesiano

El promedio del modelo bayesiano (BMA) utiliza con el promedio de varios modelos para hacer predicciones, con datos dados por la probabilidad posterior de cada modelo en función de los resultados. El BMA suele dar una mejor respuesta que un solo modelo, mediante regresión por pasos, especialmente cuando modelos muy diferentes tienen casi el mismo rendimiento en el conjunto de entrenamiento, pero pueden tener un rendimiento diferente. El teorema de Bayes es una prioridad con cada modelo sea óptimo para un propósito determinado.

En el aprendizaje supervisado, la aplicación del modelo bayesiano tiene como objetivo mejorar la precisión de las predicciones realizadas por el algoritmo de aprendizaje. Para ello, se utilizan las probabilidades bayesianas para estimar la confianza en las predicciones del algoritmo. De esta forma, se pueden identificar y corregir los errores de predicción, mejorando el rendimiento del algoritmo de aprendizaje.

El modelo abreviado BMA se puede utilizar con cualquier paquete de R (software estadístico) que utilizan la prioridad implícita en los Criterios de información bayesianos la sustitución de prioridad implícita en otros criterios (George Box, 2011).

4.5.9 Bosque aleatorio

El bosque aleatorio es una técnica de aprendizaje automático que se utiliza para resolver problemas de regresión y clasificación. Utiliza el aprendizaje por conjuntos, que es una técnica que combina clasificadores para proporcionar soluciones a problemas complejos. El algoritmo de bosque aleatorio consta de muchos árboles de decisión.

El bosque generado por el algoritmo de bosque aleatorio se entrena mediante la agregación de guías. El modelo está regido por un algoritmo establecido que puede mejorar la precisión de los demás algoritmos de aprendizaje automático. El algoritmo genera el resultado basándose en la predicción del árbol de decisiones. Predecir tomando el valor promedio o el promedio de varios árboles. Aumentar el número de árboles aumentará la precisión de los resultados (Cutler Adele, 2012).

En el aprendizaje supervisado, el modelo de bosque aleatorio se puede aplicar para clasificar una muestra de entradas en una categoría dada. El modelo consta de una serie de nodos, cada uno de los cuales representa una clase o categoría. Cada nodo tiene una probabilidad asociada a él, que indica la probabilidad de que una entrada se clasifique en esa categoría. La importancia de cada nodo se determina mediante una función de peso, que asigna un peso a cada nodo en función de la importancia relativa de la clase a la que pertenece.

Los nodos se conectan entre sí mediante una serie de arcos, cada uno de los cuales representa una probabilidad de conexión. La longitud del arco representa la probabilidad de que la conexión se produzca. Cuanto mayor es la probabilidad, mayor es la longitud del arco.

Las entradas se introducen en el modelo a través de una serie de arcos que conectan el nodo inicial con el nodo final. La entrada se clasifica en la categoría asociada al nodo final al que llega.

La idea esencial del bosque aleatorio es promediar muchos modelos de datos aproximadamente imparciales, reduciendo así la variación. Los árboles son modelos ideales para el agrupamiento de datos, porque pueden registrar estructuras de interacción complejas en los datos y, si crece lo suficientemente profundo, tienen un sesgo relativamente bajo. Debido a que los árboles son notoriamente ruidosos (una gran dispersión entre los datos), se benefician enormemente del promedio.

Cada árbol se construye usando el siguiente algoritmo: se designa la variable N , el número de casos de prueba y M el número de variables en el clasificador. Sea m el número de variables de entrada utilizadas para determinar las decisiones en un nodo determinado; m debe ser mucho menor que M en un conjunto de entrenamiento para este árbol y el resto de los casos de prueba para estimar el error. Para cada nodo del árbol, se eligen aleatoriamente m variables como base para la decisión. Se Calcula la partición óptima del conjunto de entrenamiento a partir de m variables (Chacon, 2017).

4.5.10 Cluster analysis

Metodología para el análisis de clúster y tiene una aplicación de agrupación de datos que comparten una misma tendencia, además de los beneficios del análisis de conglomerados, existen algunas desventajas. Analizar La agrupación es una técnica descriptiva, no teórica y sin razonamiento. El análisis de conglomerados no tiene una base estadística para extraer inferencias estadísticas La población de la muestra es un método basado en criterios geométricos, aplicado a una técnica exploratoria, descriptiva más que explicativa.

La solución no es única, siempre que sea miembro de un grupo empresarial El número de soluciones depende de muchos elementos del programa seleccionado. Por otra parte, La solución de clúster depende completamente de las variables utilizadas, la adición o destrucción de variables. La correlación puede tener un impacto significativo en la solución final (Edwards Antony, 1995).

Los análisis de cluster en la mercadotecnia digital pueden ayudar a las empresas a identificar grupos de consumidores potenciales que podrían estar interesados en sus productos o servicios. Además, estos análisis pueden ayudar a las empresas a determinar las características que podrían atraer a estos consumidores. Por ejemplo, las empresas pueden descubrir que hay un grupo de consumidores que está interesado en productos de alta gama, o que hay un grupo de consumidores que está interesado en productos ecológicos.

4.5.11 Estimación de densidad Kernel

La estimación de densidad de kernel es una técnica que se utiliza para medir la distribución de puntos en una imagen. Esta técnica tiene aplicaciones en la mercadotecnia digital, ya que permite medir la concentración de puntos en una imagen y analizar el patrón de distribución de estos puntos. Esto permite identificar zonas de alta concentración de puntos en una imagen y analizar la colocación de estos puntos. Esto puede ayudar a identificar patrones de comportamiento en los consumidores y diseñar estrategias de mercadotecnia en base a estos patrones.

El modelo de densidad de kernel se usa en el aprendizaje supervisado para estimar la probabilidad de una clase dada a partir de una muestra de datos y poder calcular la distancia entre dos puntos en un espacio determinado, y la densidad de este modelo se usa para calcular la probabilidad de una clase dada a partir de una muestra de dato, utilizado en geomarketing para realizar análisis de viabilidad comercial en el desarrollo urbano (cuando el mercado objetivo forma parte de una clase social o pueda encontrarse en una parte específica de la ciudad).

Este Modelo que es utilizado para mostrar la distribución de una población de datos, a diferencia de un histograma, este no tiene una distribución normal, y además muestra la concentración de las variaciones de la población suavizado de la escala entre separaciones, para la estimación de una tendencia poblacional (Silverman, 2018).

4.6 Aplicación de algoritmos en la predicción de comportamiento de usuarios

En la generación de información que los usuarios se obtienen a través de sus interacciones en la web, esta información es particularmente valiosa para entender sus necesidades y comportamientos. (Las herramientas como pixel de seguimiento es un ejemplo de este proceso) Para aprovechar esta información, es necesario aplicar técnicas de análisis de datos, como el machine learning.

Para implementar esta técnica, es necesario seguir ciertos pasos clave. Primero es necesario seleccionar un algoritmo de machine learning adecuado para la predicción de comportamientos de los usuarios en la página web. Es importante seleccionar un algoritmo que ya se conoce su funcionalidad para el que esté diseñado en ese propósito específico.

Una vez seleccionado el algoritmo, se deben identificar los datos necesarios para hacer predicciones. Estos pueden incluir el historial de navegación de los usuarios en el sitio web, las características del sitio web, la demografía de los usuarios y otros datos relevantes.

Después de recopilar los datos necesarios, se debe entrenar el algoritmo con estos datos. Es importante asegurarse de que el conjunto de datos utilizados para el entrenamiento sea representativo de la población de usuarios del sitio web.

Finalmente, una vez que se ha entrenado el algoritmo, se puede utilizar para hacer predicciones sobre una base de datos generado por los usuarios en la web. Estas predicciones pueden ser utilizadas para mejorar la experiencia de usuario, personalizar el contenido y optimizar el diseño del sitio web.

En conclusión, la aplicación de algoritmos de machine learning para la predicción del comportamiento de los usuarios. Es importante seguir los pasos adecuados para seleccionar el algoritmo adecuado, recopilar los datos necesarios y entrenar el modelo correctamente para obtener resultados precisos y útiles para la optimización de la experiencia del usuario en el sitio web.

Para los negocios digitales. Una nueva opción se ha puesto a disposición de las empresas: AISHII, un algoritmo que predice el comportamiento de los usuarios de cualquier producto digital. Este sistema estudia las primeras interacciones del usuario en una web o aplicación y las analiza a través de modelos estadísticos, inteligencia artificial y aprendizaje automático para adivinar si es un potencial comprador o no, si va a realizar más visitas o qué productos y servicios le han llamado la atención.

Según Diego Cuadrado, el responsable de este proyecto, se trata de una combinación de psicología y analítica de mercados. El porcentaje de acierto en la predicción del comportamiento de los usuarios es superior al 75%, pero es necesaria una cantidad significativa de datos para que el algoritmo realmente pueda ser predictivo. Por eso, cuantos más datos pueda recoger la empresa interesada en el servicio, mejor (Catalunya vanguardista, 2022).

En este proyecto se ofrece a las empresas tecnológicas interesadas una prueba gratuita para analizar el comportamiento de hasta 10.000 usuarios. El caso de AISHII es uno de muchos desarrollos algorítmicos para anticipar el comportamiento de los usuarios y mejorar las estrategias empresariales en el mundo digital.

4.7 El proceso del análisis de datos

El proceso del análisis de datos es un paso fundamental en el desarrollo de aplicaciones de machine learning. Comienza con la definición del objeto de estudio, la recolección de datos relevantes, seguido de la preparación y tratamiento de los mismos mediante técnicas de extracción, transformación y carga.

Luego, se procede a la selección y entrenamiento de un modelo de machine learning que pueda hacer predicciones precisas basadas en los datos SQL recopilados, Dentro del desarrollo de una aplicación del machine learning, podemos definir las variables de las métrica, después del lanzamiento de la página web, para la solvencia que brinda google analytics retroalimenta la etapa de la recolección de datos, etapa previa a la minería y selección de datos (ver Figura 6. Esquema del procesamiento de datos y modelado enfocado en el uso de machine learning.6).

- Objeto de estudio
- Preparación y tratamiento de datos (Extracción, transformación y carga)
- Modelo de entrenamiento
- Testeo del modelo
- Mejoras y entrenamiento

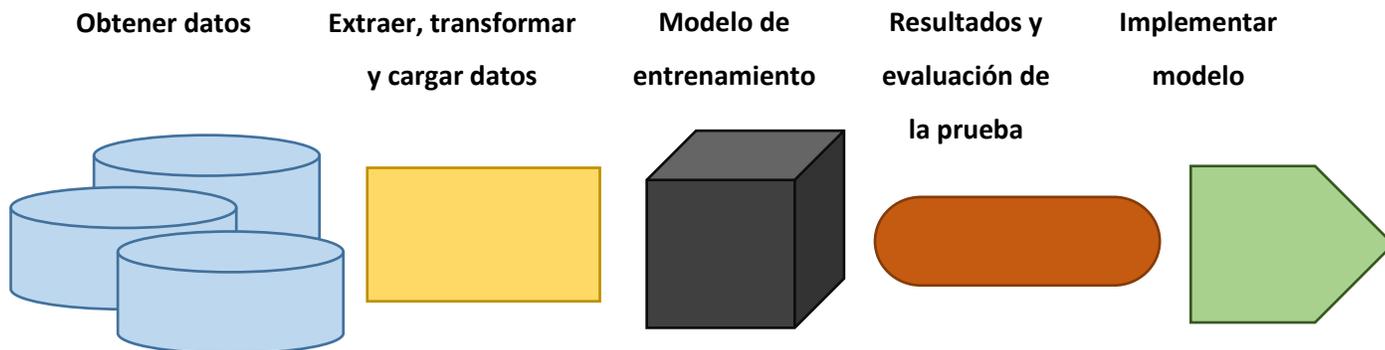


Figura 6. Esquema del procesamiento de datos y modelado enfocado en el uso de machine learning.

El modelo se prueba y se ajusta según sea necesario para mejorar su precisión, lo que puede incluir la eliminación de variables irrelevantes o la incorporación de nuevos datos. En última

instancia, el análisis de datos es un proceso iterativo que requiere una estrecha colaboración entre los científicos de datos y los expertos en el tema para lograr resultados precisos y útiles.

En el proceso de ETL (ver Figura 7) es importante considerar la naturaleza de los datos y poder agruparlos por clúster (K-Mean) o grupos para una mejor gestión, transformar los datos o llenar los datos faltantes con datos neutros y cargarlos en un documento que un orden para que no se crea conflicto al exportarlo en un documento .CSV o exportado de una base de datos SQL y no crea sea rechazado por otros programas de análisis estadísticos o almacenarlos (Géron, 2017).

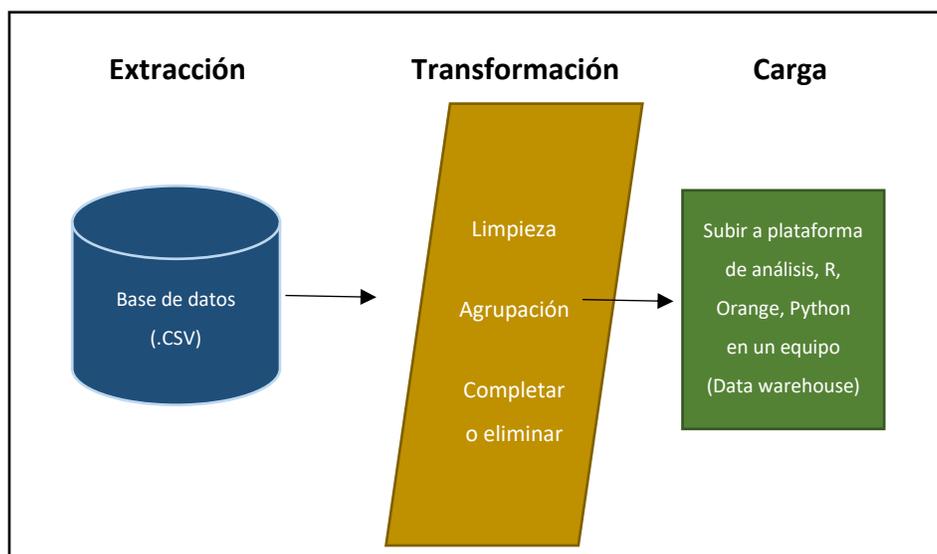


Figura 7. Proceso de extracción de datos, procesado de ETL y carga de la base de datos a servidores para procesarlos (Data warehouse).

En el apartado de los modelos de entrenamiento o también conocido como funciones de activación posibilita la personalización de los modelos y el formato de los resultados deseados ya que las funciones de activación permiten la personalización de un modelo de entrenamiento para satisfacer las necesidades específicas de un usuario.

4.8 Funciones de activación

Las funciones de activación son esenciales en las redes neuronales artificiales, ya que permiten la transferencia de información entre las neuronas, lo que a su vez permite el aprendizaje y la memoria de la red. Estas funciones determinan qué patrones de entrada para adaptarse a patrones de salida, así como la fuerza de la asociación entre ellos. Es importante seleccionar la

función de activación adecuada, ya que esto afecta significativamente el rendimiento y la capacidad de generalización del modelo de aprendizaje automático.

Las funciones de activación dentro de una red neuronal artificial se integran para permitir la comunicación entre las neuronas, así como la capacidad de aprendizaje y memoria de la red. La activación permite la transferencia de información entre las neuronas, lo que permite que la red aprenda y se adapte a nuevas situaciones para determinar qué patrones de entrada activan qué patrones de salida, y para determinar el grado de activación de cada uno de ellos. Las funciones de activación también determinan la fuerza con la que cada patrón de entrada se asocia con cada patrón de salida.

Las funciones de activación son importantes en los modelos de machine learning porque determinan cómo el modelo aprende y generaliza de los datos. Hay muchos tipos diferentes de funciones de activación (ver Figura 8), cada uno con sus propias características. Algunos funcionan mejor para determinados tipos de problemas, mientras que otros funcionan bien para la mayoría de los problemas. Seleccionar la función de activación correcta es una parte clave en el éxito del modelo de aprendizaje automático (Ver Figura 8. Graficación de las principales funciones activación.8). A continuación se recopila las funciones más comunes. (Matich, 2001)

Función identidad: también conocida como función lineal, esta función devuelve el mismo valor que se le proporciona como entrada. Esta función es útil cuando se desea que la salida de la neurona sea proporcional a su entrada.

Función escalón: esta función produce una salida binaria (0 o 1) en función de si la entrada está por encima o por debajo de un umbral determinado. Esta función es útil para la clasificación binaria.

Función lineal a tramos: también conocida como ReLU (unidad lineal rectificadora), esta función devuelve el valor de entrada si es positivo, 0 si es negativo. La función ReLU es popular en las redes neuronales convolucionales debido a su simplicidad y eficacia en el procesamiento de imágenes.

Función sigmoidea: esta función tiene una forma de "S" y devuelve una salida en el rango de 0 a 1. Es útil para problemas de clasificación binaria y para modelar relaciones no lineales.

Función gaussiana: esta función se basa en la distribución gaussiana (también conocida como distribución normal) y produce una salida en forma de campana. Es útil para modelar relaciones no lineales complejas.

Función tangente: también conocida como tangente hiperbólica, esta función tiene una forma similar a la función sigmoidea pero su rango es de -1 a 1. Es útil para problemas de clasificación binaria y para modelar relaciones no lineales. (Antonio J. Serrano, 2009-2010)

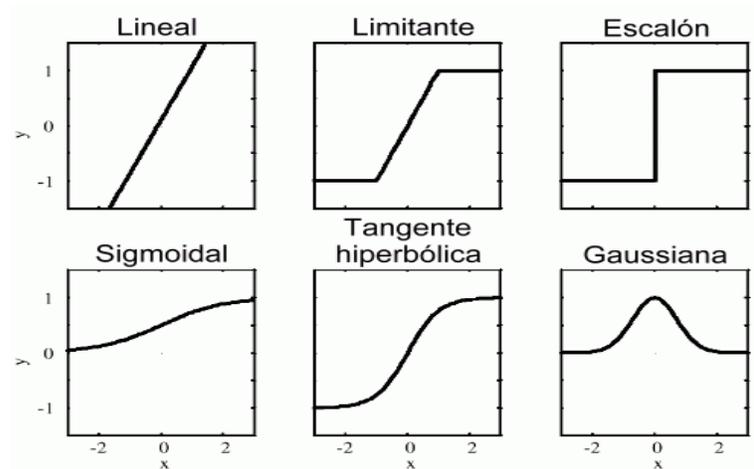


Figura 8. Graficación de las principales funciones activación.

Las últimas dos fases se basan en analizar los resultados de estas funciones y realizar ajustes si es necesario y fortalecer el modelo, Las funciones de activación en el aprendizaje automático son importantes para que el aprendizaje sea efectivo. La activación puede ser iniciada por el estímulo mismo o guiada por el usuario para ajustar el comportamiento de la red. La activación es importante porque permite que el aprendizaje sea más efectivo y que se mantenga por más tiempo.

Existen diferentes tipos de funciones de activación, y cada una tiene sus propias características y aplicaciones. La función identidad, la función escalón y la función lineal a tramos son funciones de activación lineales simples, mientras que la función sigmoidea, la función gaussiana y la función tangente son funciones de activación no lineales.

Después de aplicar estas funciones, se deben analizar los resultados y realizar ajustes si es necesario para fortalecer el modelo de aprendizaje automático. La activación es importante para

el aprendizaje efectivo, ya que permite que la red pueda adaptarse y aprender de nuevas situaciones, y mantener el aprendizaje por un período más largo.

CAPÍTULO V. METODOLOGÍA

Dado que el objetivo principal es aplicar herramientas de Machine Learning al marketing digital, la metodología está orientada hacia la resolución de problemas prácticos en el campo del marketing. Por lo que la metodología es considerada ciencia aplicada, ya que busca aplicar el conocimiento teórico y práctico del Machine Learning para mejorar estrategias y resultados reales en el marketing digital.

Se realizará un análisis exhaustivo de la página web de Isabel Taracena y sus redes sociales para obtener una comprensión completa de su estrategia actual de marketing digital. Además, se llevará a cabo una entrevista con la propietaria del estudio de fotografía para obtener más información sobre su experiencia en marketing digital y su disposición a implementar herramientas de machine learning en su estrategia.

La metodología empleada en esta investigación consiste en identificar una página web (ver Figura 9) que permitiera aplicar técnicas de Machine Learning y que tuviera funciones de marketing digital y acceso de administrador. Se seleccionó una microempresa de servicios fotográficos profesionales con sede en Ciudad de México que utiliza prácticas de marketing digital a través de su página web y redes sociales (ver Figura 10).

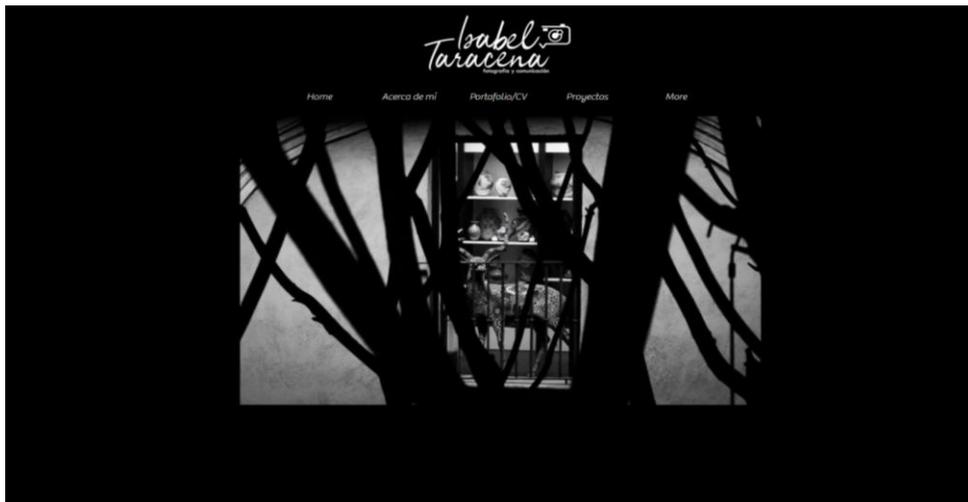


Figura 9. Captura de la portada principal de la página web.

Se vinculó la página web (<https://www.isabeltaracena.com>) con Google Analytics para obtener datos de tráfico y se accedió a las métricas de Instagram para medir el desempeño de las publicaciones. A partir de estos datos, se utilizó la técnica de Machine Learning para analizar el

comportamiento de los usuarios, identificar el tipo de contenido que más les interesa, las horas de mayor actividad y los mejores días para publicar.

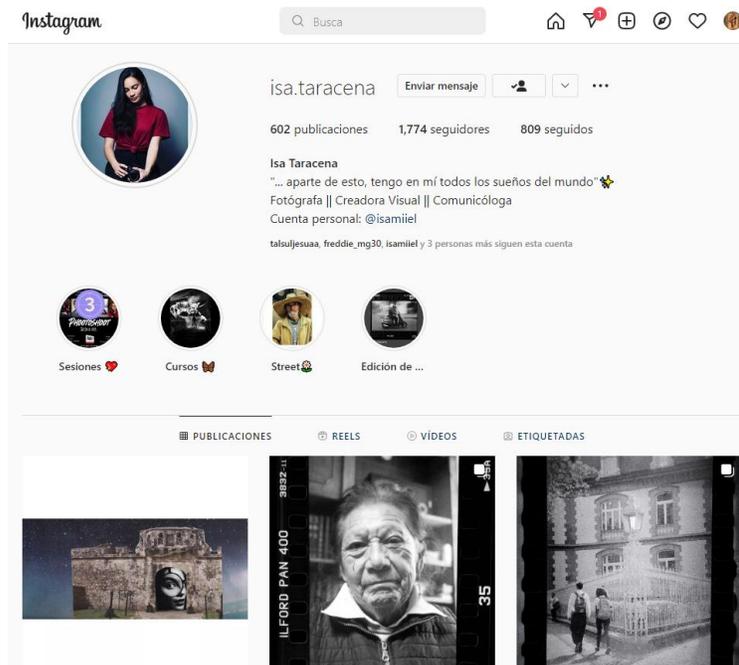


Figura 10. Captura de las redes sociales a analizar (Instagram)

Durante los primeros dos meses, el análisis de Machine Learning fue limitado debido a la baja cantidad de visitas generadas en la página web, lo que no permitía un análisis objetivo. Por lo tanto se optó acciones de marketing específicas para poder analizar el cada cambio y observar reacciones.

5.1 Tipo de Estudio

El tipo de estudio se empleará tanto métodos cuantitativos como cualitativos. El estudio será retrospectivo, analítico, experimental y de intervención. Esto con la intención de explicar la relación causal entre la implementación de herramientas de machine learning en la página web y el aumento en interacciones.

Categorías en las que recae el tipo de investigación:

- Retrospectivo y prospectivo: Se utilizarán datos retrospectivos de las ventas anteriores, y se realizará un seguimiento prospectivo de las visitas después de implementar las herramientas de machine learning.
- Transversal: El estudio se realizará en un momento específico, comparando las métricas de la página web / redes sociales antes y después de implementar las herramientas de machine learning. (seguimiento en Instagram desde 3 enero 2022 al 19 de diciembre del 2022, mientras que el seguimiento de la página web 14 de Marzo del 2022 al 20 de noviembre del 2022)
- Analítico: Se analizarán las métricas de la página web / redes sociales para determinar si las herramientas de machine learning han mejorado su rendimiento. (engagement y visitas a la página web en el tiempo)

Tipo de estudio: Explicativo/causal, ya que se busca explicar la relación causal entre la implementación de herramientas de machine learning en la página web y el aumento en interacciones. Además, se buscará identificar los factores que tengan una correlación y cómo afectan los resultados finales (palabras clave, presentación del contenido y tipo de fotografías que tengan una mejor recepción al público).

En conclusión, el estudio que se llevará a cabo será de tipo mixto, ya que se emplearán tanto métodos cuantitativos como cualitativos para analizar las métricas de la página web y redes sociales antes y después de implementar las herramientas de machine learning. Se espera que los resultados del estudio permitan mejorar el rendimiento de la página web y redes sociales, y contribuyan a la toma de decisiones estratégicas en futuras publicaciones que se realicen.

5.1.1 Diseño de Investigación

El diseño de investigación de esta tesis es no experimental y de tipo descriptivo transaccional. El objetivo es describir la relación entre el uso de herramientas de machine learning aplicado a un contenido digital para incrementar el engagement generado por los usuarios.

Se considera no experimental debido a que el objetivo principal es describir la relación entre las variables dadas a los modelos de machine learning y el comportamiento de los usuarios, por lo que no se manipula ninguna variable en las condiciones existentes.

Además, se considera que el diseño es transaccional, porque se recopilarán datos en un determinado tiempo (a lo largo del 2022) y se analizarán esos datos recopilados para establecer

la relación entre las variables de interés. En este caso, se recopilarán datos tanto cuantitativos como cualitativos sobre el uso de herramientas de machine learning por la actividad generada en la página web y el engagement generado por los usuarios en las publicaciones-contenido digital.

En términos cuantitativos, se recopilarán datos de las plataformas de análisis (Google Analytics) para analizar estadísticamente las métricas de la página web. Además, se utilizará un enfoque cualitativo para recopilar información sobre la percepción de los usuarios sobre el uso del contenido soportado por las herramientas de machine learning.

Para garantizar la privacidad y la confidencialidad de los datos, se obtendrán los permisos necesarios de la empresa para llevar a cabo la investigación. Además, se adoptarán medidas de seguridad adecuadas para proteger los datos de los clientes y de la empresa en todo momento.

5.2 Sujetos de Estudio o población

La población objeto de estudio son consumidores potenciales de arte visual en la Ciudad de México, tanto locales como extranjeros, que frecuentan galerías de arte y zonas turísticas de la ciudad. Además, se estima que el mercado está conformado de la siguiente manera: tener 25 años o más, y pertenecer a los estratos socioeconómicos de clase media, media-alta o alta. Se espera que tengan un nivel de educación profesional o posgrado y sean activos en redes sociales.

- Edad aproximada a 25 años en adelante
- Pertenecer a la clase media, media-alta o alta
- Residir en la Ciudad de México o ser extranjero interesado en obras visuales
- Tener interés en obras visuales
- Ser activo en redes sociales
- Poseer nivel de estudios profesionales o posgrados

Para el estudio de geomarketing el universo está definido como pequeñas y medianas empresas con presencia en internet y que utilizan marketing digital.

- Empresas de comercio electrónico que cuenten con plataforma en línea para ventas al consumidor final.
- Que operen en cualquier estado o ciudad de México

Para la muestra seleccionada en esta investigación sobre el geomarketing se conforma por una recolección de pequeñas y medianas empresas que operan en México recopiladas a través del directorio de empresas presentes en google y DENUE, con presencia en internet y que utilizan marketing digital, pero no tienen acceso a herramientas especializadas de machine learning para la optimización de recursos digitales. Además, se consideró una población de consumidores potenciales de arte visual en la Ciudad de México, tanto locales como extranjeros, con características demográficas específicas. La selección rigurosa de la población seleccionada garantiza que los resultados obtenidos sean representativos y generalizables, lo que contribuirá a obtener conclusiones sobre la relación entre el uso de herramientas de machine learning aplicado a un contenido digital y el engagement generado por los usuarios.

5.3 Muestra

Para la muestra se seleccionó de manera intencional y está compuesta por la página web del estudio de fotografía "Isabel Taracena Fotografía" (T., 2021) (<https://www.isabeltaracena.com/>), ubicado en la ciudad de México. Dicha microempresa cuenta con página web y redes sociales personales que fueron seleccionadas por su presencia en línea y por su oferta de cursos de fotografía y servicios de fotografía para diversos eventos.

Para la muestra en redes sociales se tomó la totalidad de 1866 seguidores que mantiene Isabel Taracena en sus redes sociales, mismo que se utilizó para medir el engagement.

5.3.1 Tamaño de muestra en el estudio de geomarketing.

El objetivo principal del estudio es analizar cómo una pequeña empresa en línea puede mejorar su estrategia de marketing digital mediante el uso de herramientas de machine learning para optimizar sus recursos digitales. La muestra es representativa de pequeñas empresas que tienen presencia en línea y utilizan herramientas de marketing digital, pero que carecen de presupuesto y conocimientos técnicos específicos para acceder a herramientas especializadas de machine learning.

Donde:

E = margen de error

Z = valor crítico de la distribución normal estándar para el nivel de confianza deseado

Sigma = desviación estándar de la población

n = tamaño de muestra

Para conocer cómo se conforma el mercado, se consideró una población aproximada de 70 estudios fotográficos en Ciudad de México y se realizó un muestreo aleatorio simple de 55 estudios fotográficos para la investigación (Tabla 4). Se eligió un tamaño de muestra de ese tamaño para tener un margen de error del 10% con un nivel de confianza del 95%, lo cual permite una precisión adecuada en los resultados de la investigación. La muestra se obtuvo mediante el directorio de estudios fotográficos de google para generar una lista completa de los mismos en la ciudad de México. Se utilizó la selección aleatoria simple para garantizar que todos los estudios fotográficos tuvieron la misma probabilidad de ser elegidos para la muestra.

Se recopiló información del DENUÉ y del directorio de empresas en Google Maps. Aplicando a la ecuación una población aproximada de 70 estudios fotográficos en Ciudad de México (N), considerando una muestra de 55 estudios fotográficos para la investigación (n). Y tener un margen de error del 5% con un nivel de confianza del 95%.

La población es finita, la fórmula para calcular el tamaño de muestra sería:

$$n = N * (Z^2 * p * (1-p)) / ((N-1) * E^2 + Z^2 * p * (1-p))$$

Donde:

N = tamaño de la población

n = tamaño de muestra buscado

Z= Parámetro estadístico que depende el nivel de confianza (Nivel de confianza del 95% es igual a Z alfa de 1.96)

E= Error de estimación máximo aceptado

P= Probabilidad de que ocurra el evento estudiado

(1-p)= Probabilidad de que no ocurra el evento estudiado

$$n = 70 * (1.96^2 * 0.5 * 0.5) / ((70-1) * 0.1^2 + 1.96^2 * 0.5 * 0.5)$$

n = 55.3

Por lo tanto, el tamaño de muestra de 55 se ajusta a la población finita de 70 estudios fotográficos para tener un margen de error del 10% con un nivel de confianza del 95%.

5.4 Instrumento de recolección de datos

La recolección de datos para este estudio se realizó a lo largo del año 2022, utilizando diversas fuentes confiables. Para la obtención de las métricas de redes sociales se utilizó la plataforma Meta, específicamente para el seguimiento de las cuentas de Instagram de Isabel Taracena. Para el análisis de la página web se empleó Google Analytics, una herramienta ampliamente utilizada en la medición de tráfico y estadísticas de páginas web.

Para considerar el estudio de mercado de la Ciudad de México, se utilizó la plataforma de estadísticas del INEGI, más específicamente el apartado DENU (Directorio Estadístico Nacional de Unidades Económicas), que ofrece información actualizada sobre empresas y establecimientos en todo el país. Adicionalmente, se utilizó el directorio de empresas que ofrece Google Maps como una fuente complementaria.

Todas las fuentes de datos utilizadas fueron vinculadas directamente a la cuenta de Instagram de Isabel Taracena, por lo que se consideran confiables y están directamente relacionadas con la microempresa objeto de estudio. La recolección de datos se llevó a cabo de manera completamente en línea, sin necesidad de realizar una investigación de campo en la Ciudad de México debido a limitaciones de presupuesto y tiempo.

5.5 Instrumentos de análisis de datos

El análisis de datos desempeña un papel fundamental en la toma de decisiones informadas en una amplia variedad de campos, desde la ciencia y la tecnología hasta los negocios y la investigación académica. Para realizar un análisis de datos efectivo, es crucial contar con las herramientas adecuadas. Estas herramientas, que pueden ser software, técnicas estadísticas o plataformas en línea, permiten a los analistas explorar, visualizar y comprender conjuntos de datos complejos.

A continuación, se menciona y explica brevemente las herramientas de análisis utilizadas:

- Microsoft Excel: Excel es una herramienta de hoja de cálculo ampliamente utilizada para el análisis de datos. Permite realizar cálculos, crear gráficos y tablas dinámicas, y es útil para tareas de manipulación y limpieza de datos.
- Python: Python es un lenguaje de programación versátil y muy utilizada en análisis de datos. Librerías como Pandas, NumPy y Matplotlib proporcionan capacidades avanzadas para cargar, procesar y visualizar datos.
- Anaconda Python: Anaconda es una distribución de Python que se utiliza comúnmente en la ciencia de datos y el análisis de datos. Proporciona un entorno de desarrollo integrado (IDE) llamado "Anaconda Navigator" que incluye numerosas bibliotecas y herramientas preinstaladas para el análisis de datos, como NumPy, Pandas, Matplotlib y Jupyter Notebook. Una de las ventajas clave de Anaconda es su sistema de gestión de paquetes, que facilita la instalación y gestión de paquetes y entornos de Python. Esto es particularmente útil cuando se trabaja en proyectos de ciencia de datos que requieren versiones específicas de paquetes.
- R: R es un lenguaje de programación y un entorno específicamente diseñado para la estadística y el análisis de datos. Tiene una amplia gama de paquetes y librerías para tareas estadísticas y gráficas.

Una vez que se obtuvo la información empírica mediante las fuentes de recolección de datos mencionadas en la sección anterior, se procedió a realizar el proceso de extracción, transformación y carga de los datos en un archivo de formato .csv utilizando el software Microsoft Excel. Este proceso permitió ordenar la información y descartar o rellenar los espacios que obstaculiza el análisis de datos, aunque también podría utilizarse open office como una alternativa de uso libre para esta parte del proceso.

Para Identificar la aplicación de técnicas de aprendizaje automático en ventas mediante el análisis de las principales herramientas de código abierto para el aprendizaje automático.

Posteriormente, se utilizó la distribución Anaconda, que consiste en un conjunto de herramientas integradas a Python, que es un lenguaje de programación de alto nivel que permite la facilidad de lectura, escritura y mantenimiento de código. Anaconda es un entorno de desarrollo integrado que provee una amplia variedad de herramientas para Python.

Las principales ventajas de usar anaconda como principal herramienta para el aprendizaje automático son:

- Anaconda es una herramienta versátil, sirve tanto para analizar y transformar los datos como para implementar los modelos precargados de machine learning de manera directa
- Anaconda es una herramienta de código abierto que está disponible para todos.
- Anaconda ofrece una gran cantidad de paquetes pre-instalados que facilitan el uso de las técnicas de aprendizaje automático.
- Anaconda es una herramienta fácil de usar que tiene una interfaz gráfica intuitiva.
- La comunidad que utiliza anaconda también está conformada por investigadores e instituciones que aportan más funcionalidades experimentales y constante retroalimentación a estas herramientas de análisis.

Las principales desventajas de usar anaconda como principal herramienta para el aprendizaje automático son:

- Anaconda no es una herramienta tan especializada en el aprendizaje automático y, por lo tanto, no ofrece todas las funcionalidades que otras opciones si ofrecen para este fin.
- Anaconda tiene una curva de aprendizaje un tanto elevada por la amplia gama de programas que contienen, lo que puede dificultar el uso de esta herramienta para personas que no estén familiarizadas con el aprendizaje automático.

Recopilar y comparar las distintas técnicas y elementos que componen las técnicas de machine learning, para plantear los recursos de análisis que se requieren para su implementación a través de la distribución anaconda de Python poder hacer uso de las librerías que ya incluyen los modelos de algoritmos de Machine learning.

Antes de comenzar a colocar los datos dentro de un modelo primero se debe realizar un tratamiento a los datos en su formato apropiado para que sean utilizables, para un óptimo análisis de datos. A través de un procesamiento ETL (Extract, Transform and Load) que consiste en preparar los datos para que sea compatible con el modelo propuesto y no sean rechazados al momento de ejecutar las pruebas.

Presentar la evolución de los modelos y cómo estos afectan a los resultados de manera directa al posicionamiento SEO siguiendo el presente flujo de información (ver Figura 11).

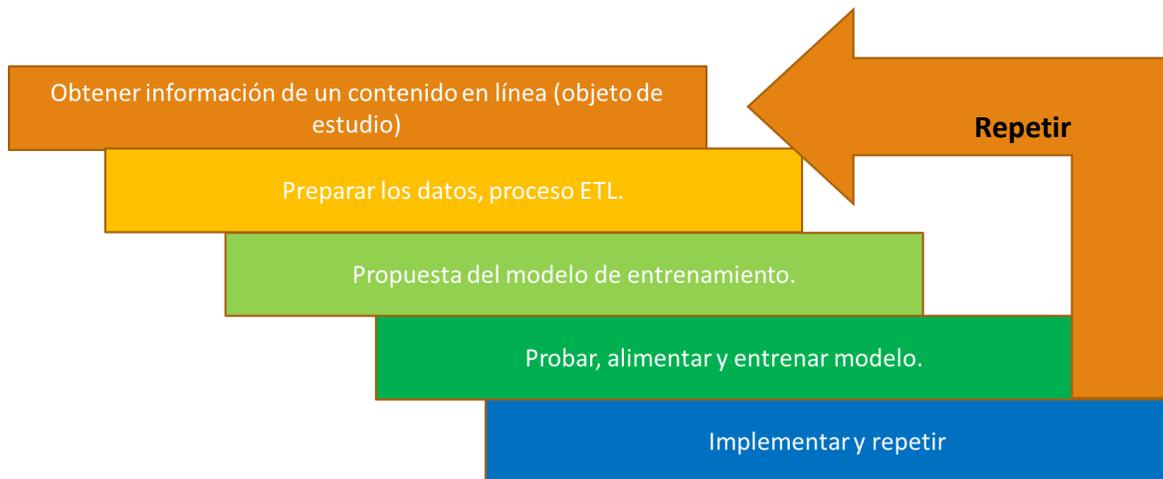


Figura 11. Etapas del procesamiento de datos para su análisis.

El proceso del análisis de datos es el siguiente:

- Se obtiene datos producidos por las visitas en google analytics, estos son exportados en formato .CSV
- Ordenar la información y limpiar campos vacíos o irrelevantes, para compilar la base de datos sin complicaciones (Proceso de ETL).
- Cargar la base de datos en el software estadístico, modelar y acomodar los datos con las funciones de activación predictivas.
- Entrenar los modelos con datos y constatar los resultados con google analytics.
- Implementar los datos en la página web y medir el alcance de los cambios.
- Repetir el proceso con los nuevos datos generados en google analytics.
- Registrar las métricas y media de visitas con machine learning de la página web analizada.

Proponer los modelos matemáticos o también conocidos como funciones de activación, en el que consiste en determinar qué función queremos que realicen es esa operación en concreto y qué resultados nos lanzaría de esos datos.

Entrenar los modelos matemáticos y realizando ajustes, alimentando constantemente con nuevos datos (feedback de la página web) para obtener mejores rendimientos en los resultados al momento de implementarlos en la estrategia de posicionamiento SEO.

Realizar la comparación de las métricas-historial y media de visitas con las métricas-ML para evaluar los resultados y hacer la comprobación de la hipótesis.

Implementar cambios en la página web, basada en los resultados de los análisis, medir el alcance de los cambios y obtener una retroalimentación de las métricas para realizar ajustes pertinentes en los análisis.

Para la selección de las diferentes herramientas que ofrece la distribución Anaconda, se destaca el uso del programa Orange como herramienta principal de minería de datos y exploración. Orange permitió aplicar diversas técnicas de análisis de machine learning para el estudio de las variables y conceptos seleccionados en la investigación.

También se utilizaron puntualmente otros softwares estadísticos para pruebas específicas de datos. Se utilizaron los lenguajes de programación R y Python (a través de Jupyter Notebook) para realizar análisis específicos, tales como regresión lineal y correlación de datos. Estas pruebas permitieron validar y complementar los resultados obtenidos con la herramienta Orange y enriquecer el análisis de los datos.

5.6 Técnicas de machine learning utilizadas

Para llevar a cabo el análisis de la web, se requiere que el usuario se identifique con una cuenta de Google. Una vez que el usuario está identificado, aparece una barra lateral en la que se listan todas las propiedades de las que el usuario es propietario. Debajo de esta lista, se encuentra el enlace "Análisis de la Web", el cual, al hacer clic, nos llevará a la página de inicio de Google Analytics, esto es nuestro principal instrumento de análisis de la página web de Isabel Taracena.

En esta página, el usuario puede seleccionar la propiedad a analizar, ya sea a través de una lista de property IDs o de una lista de nombres de dominio. Una vez seleccionada la propiedad, aparecerá una pantalla en la que se listarán todas las variables de análisis.

Las variables de análisis se clasifican en dos grandes grupos:

- Variables de medición: Estas variables se utilizan para medir el rendimiento de la página web. Por ejemplo, el número de visitas, el tiempo medio de visita, interacción en redes sociales (engagement), el porcentaje de rebote, etc.

- Variables de segmentación: Estas variables se utilizan para analizar el comportamiento de los visitantes de la página web. Por ejemplo, el origen de las visitas, el dispositivo utilizado, el navegador, etc.

Medir las variables actuales que tiene la página web, analizar su comportamiento a lo largo del tiempo, media de visitas y plantear predicciones sin Machine learning.

Para seleccionar las técnicas de machine learning apropiada para aplicar las siguientes técnicas para analizar a través de los datos generados en Google Analytics:

1. Regresión lineal: Esta técnica es adecuada cuando se desea predecir el valor de una variable dependiente a partir de una o más variables independientes, asumiendo una relación lineal entre ellas. Puede ser útil para estimar métricas como el número de visitas o el tiempo medio de visita en función de otras variables.
2. Regresión logística: Se utiliza cuando se desea predecir el valor de una variable dependiente binaria (por ejemplo, sí o no) en función de variables independientes. Puede ser útil para analizar variables como la tasa de conversión o la probabilidad de que un usuario realice una acción específica en la página web.
3. Clasificación: Se emplea cuando se desea dividir un conjunto de datos en grupos o categorías distintas. Esta técnica puede ser útil para analizar variables como el origen de las visitas (por ejemplo, búsqueda orgánica, tráfico directo, referencia, etc.) o para clasificar usuarios en grupos demográficos o de comportamiento.
4. Clustering (KNN, tree, random forest): Es una técnica que agrupa los datos en conjuntos similares sin etiquetas previas. Puede ser útil para descubrir patrones ocultos en los datos de Google Analytics, identificar segmentos de usuarios con características comunes o agrupar páginas web según su comportamiento.
5. Redes neuronales: Son modelos de machine learning más complejos que pueden capturar relaciones no lineales entre variables. Pueden ser útiles cuando se desea realizar análisis más avanzados, como la detección de anomalías en el comportamiento de los usuarios, la predicción de futuros patrones de visitas o el análisis de texto generado por los usuarios.

Al analizar las métricas que nos ofrece Google analytics para conocer la naturaleza de los datos y crear modelos estadísticos, con el fin de utilizarlos en los análisis e implementar los cambios en la página web.

Aplicar la técnica de machine learning seleccionada previamente en la investigación de las diferentes técnicas que solucione mejor las variables del análisis digital, obtenidas por google analytics, en el siguiente capítulo se presentara todo el desarrollo de la metodología, partiendo desde los datos que se obtuvieron de primera mano con las herramientas digitales, además de cómo fueron procesados e implementados en los modelos de machine learning.

CAPÍTULO VI: ANÁLISIS DE DATOS

La era digital ha abierto una gran cantidad de posibilidades para la recopilación y análisis de datos en diferentes ámbitos. En particular, en el campo del marketing, los datos se recolectan con ciertas herramientas clave para comprender el comportamiento del consumidor y mejorar las estrategias comerciales.

En este contexto, el análisis de datos se ha convertido en una disciplina fundamental para la toma de decisiones basada en hechos y no en suposiciones. La investigación en este campo ha llevado a la creación de diversas técnicas y metodologías que permiten extraer información valiosa de grandes conjuntos de datos.

En este apartado se abordará el análisis de datos en el ámbito del marketing digital, centrándose en siete subtemas clave. En primer lugar, el subtema 6.1 Métricas de la página web, se presenta la recolección de los datos de primera mano en las plataformas digitales de google analytics e Instagram (6.1.1)

Se plantearán las diferentes métricas que se pueden utilizar durante la investigación para medir el éxito de una campaña publicitaria y evaluar el rendimiento del negocio.

Luego, en el subtema 6.2 (Exploración y estudio de mercado, se explorará el uso del geomarketing) para la realización de estudios de mercado, analizando la ubicación de los consumidores y la forma en que interactúan con los productos y servicios ofrecidos.

Posteriormente en el capítulo 6.3 (selección automática de imágenes y correlación de engagement) donde se aborda el tema de la selección automática de imágenes y su correlación con el engagement del consumidor en las redes sociales, con el objetivo de comprender cómo se puede mejorar la interacción con los clientes a través del contenido visual.

En el capítulo 6.4 (Evaluación de diferentes modelos de machine learning en predicción de engagement), se examinará el rendimiento de los diferentes modelos de machine learning para predecir el engagement y determinar cuál es el más adecuado para el contexto en cuestión.

Por último, se analizarán los datos de Google Analytics en el capítulo 6.5 (Análisis de la página web con los datos de machine learning) donde visualizaremos las interacciones de los visitantes en la página web a través de series de tiempo, buscando identificar patrones y tendencias que permitan mejorar la toma de decisiones en el futuro, en este apartado se comprueba si las métricas de machine learning tiene un impacto en el aumento de visitas.

En resumen, esta tesis tiene como objetivo explorar diferentes técnicas y metodologías de análisis de datos para mejorar la toma de decisiones en el marketing digital. Se espera que este estudio contribuya al conocimiento de la disciplina y ofrezca información valiosa para la empresa como “Isabel Taracena” que buscan mejorar su presencia en línea y aumentar el engagement de sus clientes.

6.1 Métricas de la página web (Estudio fotográfico Isabel Taracena) durante la investigación

Durante el proceso de investigación, se utilizó la herramienta de análisis web de Google Analytics para obtener información valiosa sobre el comportamiento de los usuarios en la página web. Al analizar las métricas, se observó que la sección de proyectos es la más visitada, seguida de cerca por la sección de portafolio. De los 114 visitantes de la página web dentro de un periodo de 12 semanas 95 usuarios visitaron el apartado de proyectos, lo que representa un 83% de los visitantes están interesados en su portafolio en conocer más sobre los proyectos y trabajos realizados por el estudio fotográfico Isabel Taracena, lo que sugiere la importancia de destacar esta sección en la página web y mantenerla actualizada con información relevante.

La frecuencia de visitas a estos apartados de la página web puede ser utilizada como una métrica importante para evaluar el rendimiento de la misma. Asimismo, esta información puede ser utilizada por la propietaria del estudio fotográfico para identificar patrones de comportamiento y preferencias de los usuarios, lo que propondría a ayudar a mejorar la experiencia de usuario en la página web (ver Figura 12).

Es importante destacar que estas métricas no deben ser tomadas como la única fuente de información para la toma de decisiones. Se deben considerar otros factores relevantes, como la

segmentación de los usuarios y la evaluación de los resultados obtenidos a través de la implementación de diferentes estrategias y mejoras en la página web.

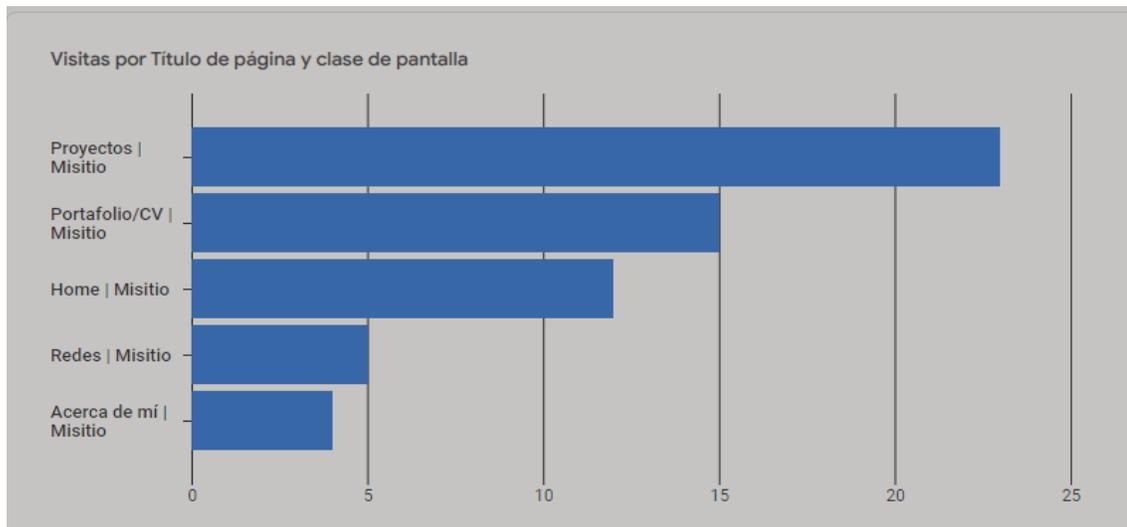


Figura 12. Cantidad de visitantes por sección de la página web

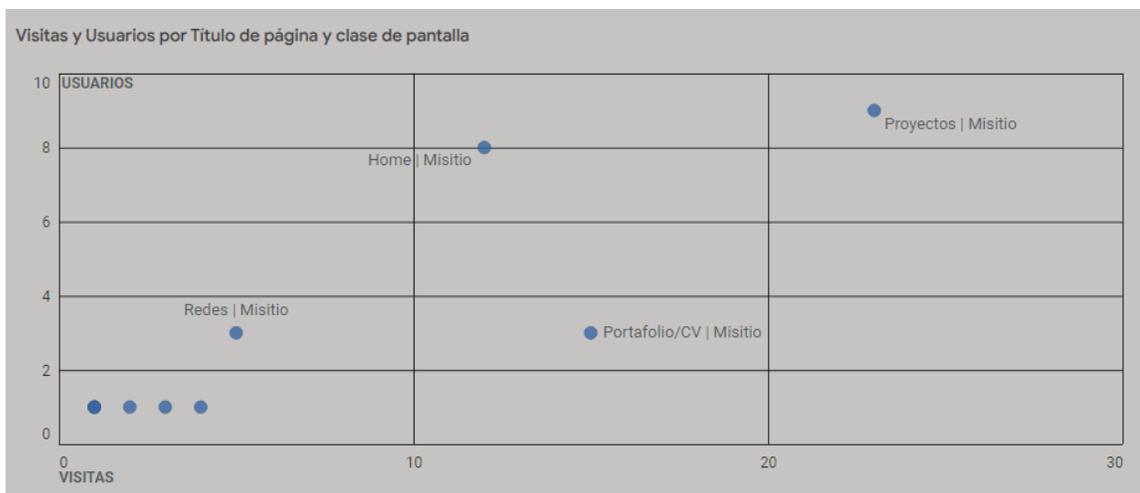


Figura 13. Gráfica de frecuencia por visitas de los usuarios.

Durante la recopilación de métricas a través de Google Analytics, también se tomó en cuenta el registro de acciones realizadas por los usuarios en cada sección de la página web. Estas acciones pueden incluir la descarga de archivos, la reproducción de videos, la interacción con formularios, entre otras. De esta manera, se pudo analizar el nivel de interacción de los usuarios con cada sección de la página y determinar cuáles son las secciones más atractivas para ellos

(ver Figura 13). Con base a estos resultados, se pueden implementar estrategias de mejora en las secciones menos populares y reforzar la presencia de las secciones más visitadas.

Título de página y clase de pantalla +	↓ Visitas	Usuarios	Usuarios nuevos	Visitas por usuario	Tiempo de interacción medio	Desplazamientos de usuarios únicos	Número de eventos Todos los eventos ▾	Conversiones Todos los eventos	Total de ingresos
Totales	67 100 % respecto al total	13 100 % respecto al total	11 100 % respecto al total	5,15 Media 0 %	0 min y 36 s Media 0 %	10 100 % respecto al total	115 100 % respecto al total	0,00	0,00 \$
1 Proyectos Misitio	23	9	5	2,56	0 min y 26 s	6	44	0,00	0,00 \$
2 Portafolio/CV Misitio	15	3	0	5,00	0 min y 15 s	1	17	0,00	0,00 \$
3 Home Misitio	12	8	6	1,50	0 min y 03 s	5	32	0,00	0,00 \$
4 Redes Misitio	5	3	0	1,67	0 min y 27 s	1	7	0,00	0,00 \$

Figura 14. Resumen por seccion de la pagina web

Otra información que se obtuvo fue que también se rastreó el origen de las visitas a la página web (ver Figura 14). Se encontró que la mayoría de las visitas provienen de redes sociales orgánicas, lo que indica que la estrategia de vincular la página web con las publicaciones en Instagram y crear su propio hashtag ha sido efectiva. Los links directos y la búsqueda orgánica también fueron fuentes significativas de visitas (ver Figura 15). La información recopilada de Google Analytics ha permitido identificar la efectividad de las diferentes estrategias de marketing digital utilizadas y ajustarlas en consecuencia para mejorar el tráfico de la página web.

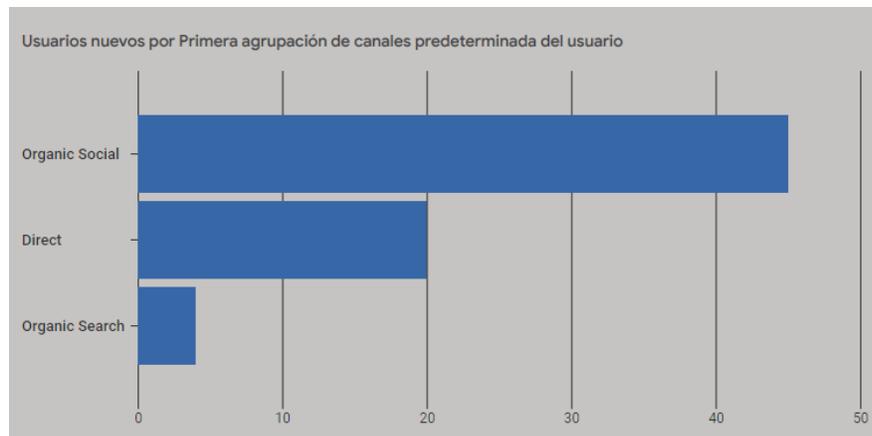


Figura 15. Canales frecuentes donde provienen las visistas.

La duración media de la visita de los usuarios en la página web es un factor importante a considerar al momento de evaluar la experiencia del usuario y la calidad del contenido.

Después de analizar los datos obtenidos de la página web Isabel Taracena a través de Google Analytics, se determinó que la media de visualización en la página web es de 1 minuto 12 segundos (ver Figura 16). Es por ello que se tomó la decisión de cambiar la sección más visitada

y ofrecer un contenido más desarrollado, con comentarios y reflexiones de la fotografía sobre su galería expuesta en la página web. Con esta acción se busca aumentar el tiempo de permanencia en la página web y mejorar la experiencia del usuario, ya que se ofrecerá un contenido más enriquecedor y valioso para los visitantes.

Primera agrupación de canales predeterminada del usuario +	+Usuarios nuevos	Sesiones con interacción	Porcentaje de interacciones	Sesiones con interacción por usuario	Tiempo de interacción medio	Número de eventos Todos los eventos +	Conversiones Todos los eventos	Total de ingresos
Totales	69 100 % respecto al total	44 100 % respecto al total	50 % Media 0 %	0,63 Media 0 %	0 min y 41 s Media 0 %	635 100 % respecto al total	0,00	0,00 \$
1 Organic Social	45	26	56,52 %	0,58	0 min y 33 s	343	0,00	0,00 \$
2 Direct	20	9	31,03 %	0,45	0 min y 23 s	143	0,00	0,00 \$
3 Organic Search	4	9	69,23 %	1,80	3 min y 02 s	149	0,00	0,00 \$

Figura 16. Resumen de los canales donde provienen los usuarios.

En conclusión, la herramienta de análisis web de Google Analytics ha sido de gran ayuda para obtener información valiosa sobre el comportamiento de los usuarios en la página web y tomar decisiones informadas para mejorar la experiencia del usuario y aumentar el engagement en la página web. Las métricas obtenidas han permitido identificar patrones de comportamiento y preferencias de los usuarios, lo que puede ayudar a mejorar la experiencia de usuario en la página web y reforzar la presencia de las secciones más visitadas. La información recopilada de Google Analytics también ha permitido evaluar la efectividad de las diferentes estrategias de marketing digital utilizadas y ajustarlas en consecuencia para mejorar el tráfico y el engagement en la página web. Es importante tener en cuenta que estas métricas deben ser consideradas junto con otros factores relevantes para la toma de decisiones informadas y eficaces. En definitiva, el análisis de datos es esencial para optimizar el rendimiento de la página web y ofrecer una experiencia de usuario excepcional.

6.1.1 Medición del engagement en las publicaciones en redes sociales

Se optó por usar Instagram como el principal canalizador de audiencia ya que era la red social donde la empresa (estudio Isabel Taracena) es más activa y principal medio de difusión para la página web, ya que es la red social más adecuada para el contenido generado y es donde se genera más actividad. Además, Instagram es la mejor opción para los fotógrafos, ya que tiene una enorme audiencia de más de mil millones de usuarios activos mensuales, de los cuales 226.3 millones se encuentran en Latinoamérica. (Galeano, 2022)

Esto significa que es una plataforma ideal para mostrar un portafolio de proyectos y llegar a una gran cantidad de personas. Instagram es principalmente visual, lo que la hace perfecta para

mostrar trabajos artísticos, y es muy fácil de usar ya que se pueden publicar fotos con solo unos pocos clics.

Para aumentar el alcance de las publicaciones, se cambió la privacidad de la cuenta de Instagram a una cuenta de creador de contenido, lo que ayuda a aumentar la visibilidad de las publicaciones. Además, se vinculó una sección de galería de Instagram en la página web, lo que permite aumentar el contenido disponible sin la necesidad de agregarlo manualmente.

Para complementar la recolección de interacciones no solo era necesario utilizar los datos de google analytics, además se usaron las publicaciones que se realizaron a lo largo del 2022 realizadas por Isabel Taracena, para la medición del engagement en las publicaciones en redes sociales es fundamental para comprender cómo interactúan los usuarios con el contenido. El engagement se refiere a la cantidad de interacciones que se generan en una publicación, tales como reacciones, comentarios, compartidos, etc. Estas interacciones no solo muestran la popularidad de una publicación, sino que también indican la relevancia del contenido y el interés por parte del público objetivo.

Para medir el engagement en las publicaciones, se pueden utilizar herramientas de análisis de redes sociales, como Hootsuite, Sprout Social e inclusive en parte Google Analytics. Estas herramientas permiten identificar el número de interacciones y la tasa de engagement de una publicación, así como también la segmentación de la audiencia que interactúa con el contenido, principalmente si están publicitados.

Es importante tener en cuenta que el engagement no solo se mide por la cantidad de interacciones, sino también por la calidad de estas. Por ejemplo, un comentario positivo de un usuario fiel puede tener un mayor valor que un simple me gusta de un usuario desconocido (ver Figura 17). Por lo tanto, es necesario analizar el tipo de interacción y la calidad de estas para considerar si las publicaciones tienen el impacto esperado.

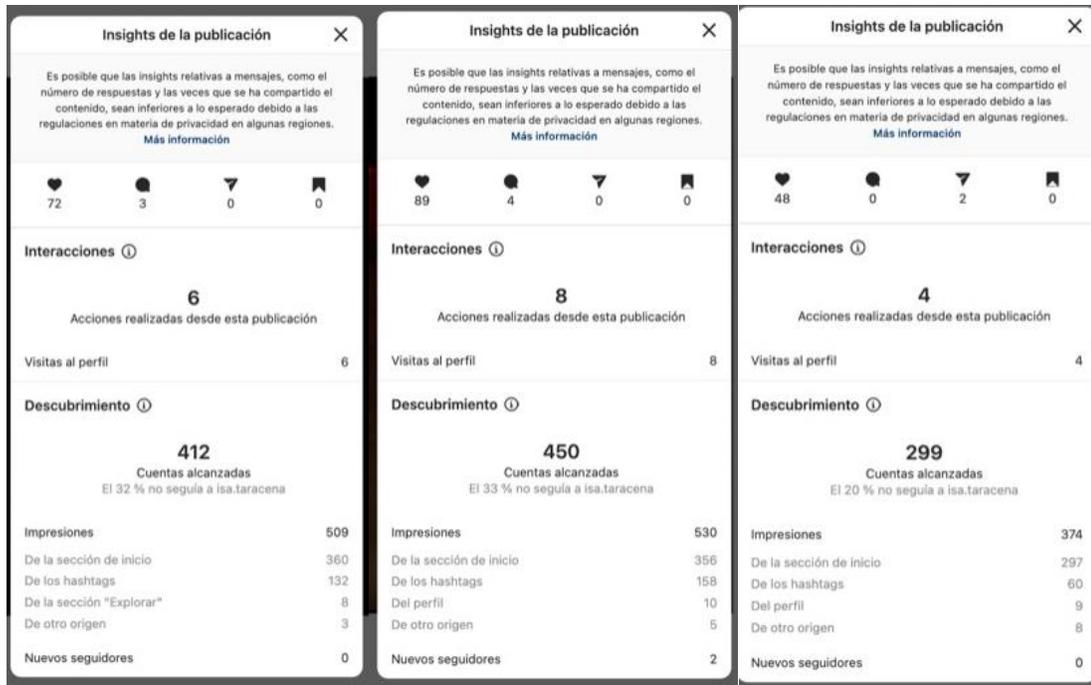


Figura 17. Resumen de las estadísticas generadas en instagram

Para sacar el indicador de engagement es una formula sencilla, $(\text{reacciones} + \text{comentarios}) / \text{seguidores totales} * 100$. Para el total de seguidores se consideró 1866 que se tenían en ese momento. El promedio de este indicador aplicado a las redes sociales varía según la plataforma y el sector en el que trabaje la marca (ver Tabla 2). Según los expertos en redes sociales, un buen engagement rate en Instagram estaría entre el 1% y el 5%, dependiendo del tamaño de su audiencia o el tipo de contenidos que publique. En Facebook e promedio es del 2%, esto considerando un promedio de seguidores grande.

Tabla 2. Calculo del engagement de las 8 publicaciones mas populares en instagram de Isabel Taracena.

ID	Tipo de fotografía	Reacciones	Comentarios	engagement
5	Reel	494	27	27.92068596
9	Autorretrato	367	34	21.48981779
57	Autorretrato	340	7	18.59592712
29	Autorretrato-Blanco y negro	233	5	12.7545552
56	Autorretrato-Blanco y negro	224	4	12.21864952
38	Autorretrato	217	7	12.00428725
51	Autorretrato	199	12	11.30760986
55	Autorretrato	190	10	10.71811361

En resumen, la medición del engagement en las publicaciones en redes sociales es esencial para comprender el comportamiento de la audiencia y mejorar la estrategia de marketing digital. Mediante el análisis de los usuarios, se puede identificar el contenido más relevante y atractivo para la audiencia, así como también identificar oportunidades de mejora en la estrategia de marketing.

6.2 Exploración y estudio de mercado aplicado por geomarketing

El geomarketing es una disciplina que combina técnicas de análisis geográfico y de marketing para entender mejor el comportamiento de los consumidores y tomar decisiones estratégicas en el ámbito comercial. El objetivo principal del geomarketing es maximizar el retorno de inversión (ROI) de una empresa, a través del uso de información geográfica y de la localización para comprender mejor el mercado y optimizar la toma de decisiones comerciales.

Entre las prácticas que se aplican en geomarketing se encuentran:

Segmentación de mercado: dividir por zonas para tener un mejor entendimiento del mercado por diferentes grupos según variables geográficas, sociodemográficas, económicas, etc., para personalizar la oferta comercial y adaptarse mejor a las necesidades de los consumidores o identificar las oportunidades y amenazas, y definir la estrategia comercial más adecuada.

Estudios de localización: Con el fin de realizar estudios de localización para identificar la mejor ubicación para un negocio, en función de factores como la densidad de población, la competencia, el tráfico, la accesibilidad, etc.

Análisis de la demanda: Para analizar la demanda de un producto o servicio en una zona geográfica determinada, para entender mejor el comportamiento de los consumidores y definir la estrategia comercial más adecuada.

Marketing de proximidad: Establecer plan de acción por proximidad, que consisten en enviar mensajes publicitarios o promociones a los consumidores en función de su ubicación geográfica, gracias al uso de tecnologías como el GPS o las redes Wi-Fi.

Estudios de competencia: para analizar la competencia en una zona determinada, para identificar sus fortalezas y debilidades, y definir la estrategia comercial más adecuada para destacar en el mercado.

Para que las empresas pueden tomar decisiones más efectivas y personalizadas para satisfacer las necesidades de sus clientes. Las técnicas de geomarketing, como la segmentación geográfica y la visualización de datos espaciales, son cada vez más importantes en el mundo del marketing digital y pueden ayudar a las empresas a obtener una ventaja competitiva en el mercado. (Elías Adolfo Albornoz Del Valle, 2020)

Durante la exploración realizada en el Directorio Estadístico Nacional de Unidades Económicas (DENUE), se pudo detectar que el tipo de industria al que pertenece la microempresa estudiada es el de "artistas, escritores y técnicos independientes" (Código DENUE 711510). Esta información es relevante para entender el contexto en el que se desenvuelve la empresa y conocer a su competencia directa. (Ver Figura 18)

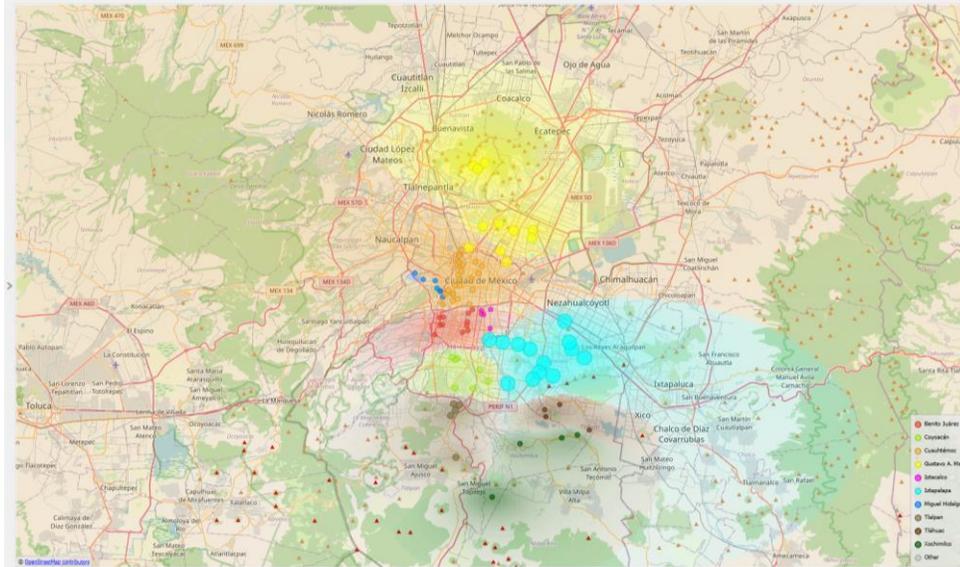


Figura 18. Registro de instalaciones código 711510 DENUE (artistas, escritores y técnicos independientes).

Además, se realizó un muestreo en el directorio de empresas de Google Maps, donde se tomó una muestra de 54 estudios fotográficos en la Ciudad de México. (Ver Figura 19) Esta muestra permitió tener un panorama más amplio del mercado del giro empresarial al que pertenece el estudio de fotografía Isabel Taracena, así como identificar algunas características comunes en cuanto a servicios ofrecidos y ubicación geográfica. Esta información es relevante para entender el entorno competitivo en el que se encuentra la microempresa y poder realizar comparaciones y análisis de mercado.

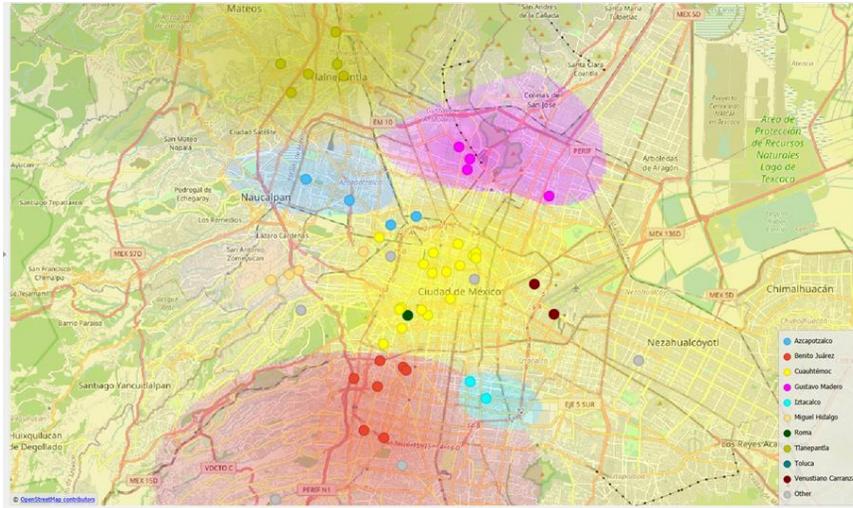


Figura 19. Ubicación de Estudios fotográficos presentes en el directorio de google maps.

Además, se investigó la densidad poblacional por delegación de la Ciudad de México. Esto permitió tener una idea más clara de las zonas con mayor concentración de habitantes, lo que puede ser útil para identificar posibles áreas de oportunidad en términos de mercado y para orientar las estrategias de promoción y publicidad de la empresa hacia las zonas con mayor potencial de clientes.

Cabe destacar que tanto la base de datos del DENUE como el directorio de Google Maps nos proporcionaron la información necesaria para poder estimar la posición geográfica de los estudios fotográficos en un mapa, esta información también puede implicar el uso del modelo knn para considerar los competidores más cercanos en las zonas más demandadas, gracias a la inclusión de datos de latitud y longitud. Además, para realizar el análisis de datos se utilizó la herramienta digital Orange, la cual forma parte de la distribución Anaconda y permitió el procesamiento y visualización de los datos de manera efectiva.

6.3 Selección automática de imágenes y correlación de engagement

Se decidió utilizar esta metodología como una forma de automatizar a través de machine learning una relación entre el tipo de publicación y su nivel de engagement, para la utilización de una red neuronal es conocida por su gran efectividad en la clasificación de imágenes y su capacidad para reconocer características específicas de las mismas. Utiliza una arquitectura compleja y

altamente optimizada que permite clasificar imágenes con alta precisión en una amplia variedad de tareas, como la identificación de objetos y la detección de rostros.

La arquitectura de Inception v3 se caracteriza por utilizar múltiples filtros de diferentes tamaños en cada capa, lo que le permite aprender características más complejas y abstractas en las imágenes. Además, utiliza técnicas de regularización para prevenir el sobreajuste y mejorar la generalización del modelo.

Inception v3 ha sido ampliamente utilizado en aplicaciones de visión artificial, como la clasificación de imágenes médicas, la identificación de objetos en tiempo real y la clasificación de imágenes en redes sociales. Además, el modelo ha sido pre entrenado en una gran base de datos de imágenes denominada ImageNet, lo que permite a los usuarios transferir el conocimiento aprendido en una tarea de clasificación de imágenes a otras tareas específicas con menor cantidad de datos de entrenamiento. (Google, 2023)

Recolectar todas las publicaciones realizadas por la microempresa de fotografía en Instagram durante el año 2022, incluyendo la cantidad de reacciones y comentarios que cada publicación recibió. Esta información se utilizó para generar una base de datos en formato con valores separados por comas (CSV).

Utilizar la distribución Anaconda para acceder a la herramienta de análisis de datos Orange.

Para clasificar las imágenes según su similitud, utilizamos el software Orange, que nos permitió realizar una serie de procesos para obtener resultados precisos y útiles.

El procedimiento se realizó de la siguiente manera (ver Figura 20):

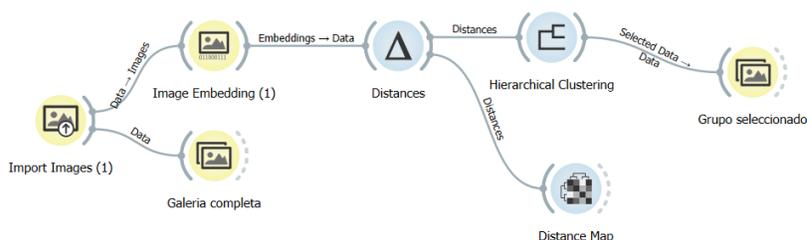


Figura 20. Proceso de análisis modelado en Orange.

1. Seleccionamos y cargamos la carpeta donde se encuentran las fotografías que deseamos clasificar. En este estudio, consideramos un total de 57 imágenes.

2. Conectamos las imágenes cargadas a un nodo llamado "Image Embedding". Este nodo procesa las imágenes según el algoritmo que seleccionamos, y en este caso, utilizamos Inception v3 por su efectividad al clasificar imágenes.
3. A este proceso, le conectamos otro nodo conocido como "Distances", que clasifica las filas de los atributos de las imágenes en una métrica euclidiana normalizada.
4. En este punto, se segmenta el proceso en dos:

Crear una matriz "Distance Map" para visualizar el proceso en general. (Figura 21)

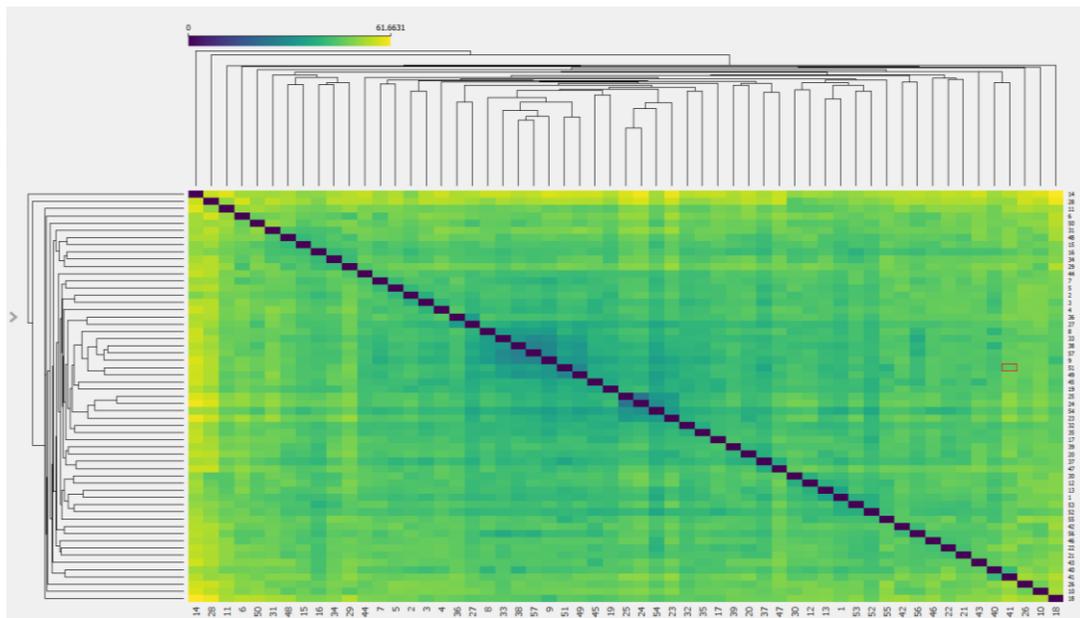


Figura 21. Matriz de Distance Map por cada fotografía.

Proceso llamado "Hierarchical Clustering", donde se muestra la segmentación realizada. Aquí, podemos seleccionar grupos específicos y añadir un último proceso que nos enseña el grupo de imágenes que conforman ese grupo seleccionado. (Figura 22)

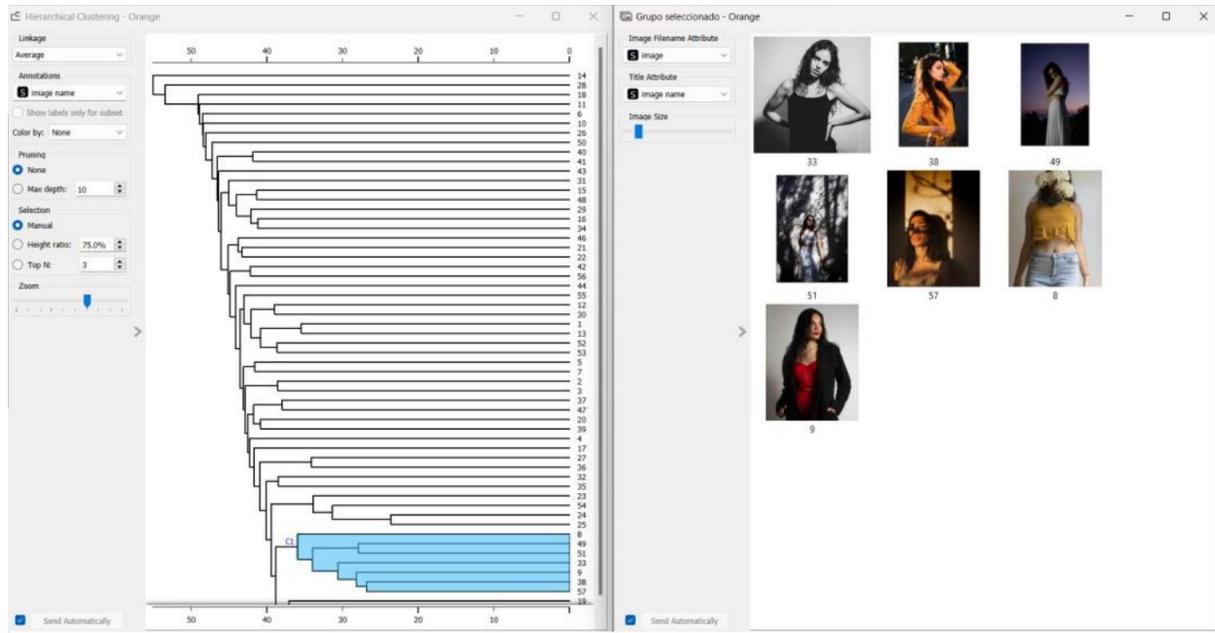


Figura 22. Selección de grupos y visualización de imagen seleccionadas.

Una vez que se obtuvieron los grupos de imágenes a través del proceso de clustering en Orange, se añadió un nuevo campo a la base de datos que correspondió al orden asignado por el modelo a cada imagen. Luego, se realizó una correlación de Pearson entre este nuevo dato y la cantidad de reacciones y comentarios que recibió cada imagen. De esta forma, se pudo determinar si el orden asignado por el modelo de clasificación de imágenes se correlacionaba con una mayor o menor interacción del público en redes sociales.

Para medir la correlación entre el orden que generó el modelo y el engagement se cargó la base de datos al software estadístico minitab y se eligió una correlación Pearson para la medición de esas dos variables (Figura 23Figura 24)

```

Ecuación de regresión
Tipo de fotografía
Autorretrato engagement = 12.69 - 0.0120 Orden en la clasificacion autom
Autorretrato-Blanco y negro engagement = 9.31 - 0.0120 Orden en la clasificacion autom
Autorretrato-Collage engagement = 5.79 - 0.0120 Orden en la clasificacion autom
Collage engagement = 4.52 - 0.0120 Orden en la clasificacion autom
paisajismo engagement = 4.09 - 0.0120 Orden en la clasificacion autom
Paisajismo engagement = 3.72 - 0.0120 Orden en la clasificacion autom
Paisajismo-Blanco y negro engagement = 4.42 - 0.0120 Orden en la clasificacion autom
Reel engagement = 13.93 - 0.0120 Orden en la clasificacion autom
Retrato engagement = 4.43 - 0.0120 Orden en la clasificacion autom
Retrato-Blanco y negro engagement = 4.46 - 0.0120 Orden en la clasificacion autom
Retrato.-Blanco y negro engagement = 4.95 - 0.0120 Orden en la clasificacion autom

Ajustes y diagnósticos para observaciones poco comunes

Obs engagement Ajuste Resid Resid
5 27.92 13.61 14.31 4.65 R
9 21.49 12.63 8.86 2.52 R
27 6.43 13.73 -7.30 -2.40 R
33 4.88 4.88 0.00 * X
34 6.43 13.44 -7.01 -2.31 R

Residuo grande R
X poco común X

```

Figura 23. Primera parte del reporte de correlaciones entre engagement, el orden automático y tipo de fotografías

Análisis de regresión: engagement vs. Orden en la clasificacion autom, Tipo de fotografia

Método

Codificación de predictores categóricos (1, 0)

Análisis de Varianza

Fuente	GL	SC Ajust.	MC Ajust.	Valor F	Valor p
Regresión	11	719.60	65.418	4.60	0.000
Orden en la clasificacion autom	1	1.13	1.131	0.08	0.779
Tipo de fotografia	10	658.89	65.889	4.64	0.000
Error	45	639.44	14.210		
Total	56	1359.04			

Resumen del modelo

S	R-cuad. (ajustado)	R-cuad. (pred)
3.76959	52.95%	41.45%

Coefficientes

Término	Coef	EE del coef.	Valor T	Valor p	VIF
Constante	12.69	1.41	8.97	0.000	
Orden en la clasificacion autom	-0.0120	0.0426	-0.28	0.779	1.97
Tipo de fotografia					
Autorretrato-Blanco y negro	-3.38	2.50	-1.35	0.183	2.01
Autorretrato-Collage	-6.90	2.65	-2.61	0.012	1.83
Collage	-8.17	2.25	-3.63	0.001	1.63
paisajismo	-8.60	2.98	-2.88	0.006	1.21
Paisajismo	-8.97	1.92	-4.67	0.000	1.78
Paisajismo-Blanco y negro	-8.27	2.27	-3.64	0.001	2.49
Reel	1.25	2.66	0.47	0.641	1.41
Retrato	-8.25	2.20	-3.75	0.000	2.34
Retrato-Blanco y negro	-8.23	2.32	-3.54	0.001	1.73
Retrato.-Blanco y negro	-7.74	4.00	-1.93	0.060	1.11

Figura 24. Segunda parte del reporte de correlaciones entre engagement, el orden automático y tipo de fotografías.

No hay una relación clara entre la agrupación automática y el éxito en el engagement de las publicaciones en Instagram. Es posible que otros factores como el contenido del texto o la hora

de publicación tengan una mayor influencia en la interacción de los usuarios con las publicaciones. En general, la agrupación automática de imágenes puede ser útil para la organización y clasificación de grandes cantidades de datos, pero no necesariamente se correlaciona directamente con el éxito en el engagement en redes sociales.

6.3.1 Conclusión de la clasificación automática del tipo de fotografía

En definitiva, el modelo tuvo una clasificación bastante acertada en ciertos grupos de fotografías, como por ejemplo las relacionadas con paisajes y la naturaleza, mientras que en otras tuvo dificultades para distinguir entre ellas, como en el caso de fotografías en forma de collage. La agrupación automática de imágenes a través de Google Inception v3 resultó ser una herramienta útil para clasificar automáticamente un conjunto de fotografías en diferentes categorías y obtener información valiosa sobre las características visuales que tienen en común las imágenes en cada grupo.

6.3.2 Conclusión de la correlación

El resultado de la correlación de Pearson entre el orden en la clasificación automática de imágenes y el engagement del contenido en Instagram, arrojó un coeficiente de correlación de -0.211 y un valor p de 0.115. (Figura 25) Esto indica una correlación negativa débil entre ambas variables, lo que sugiere que el orden en la clasificación automática no está significativamente relacionado con una mayor o menor interacción del público en redes sociales.

Correlación: Orden en la clasificación autom, engagement

Correlación de Pearson de Orden en la clasificación autom y engagement = -0.211
Valor p = 0.115

Figura 25. Resultado de la correlación Pearson y valor p.

Por lo tanto, podemos concluir que la agrupación automática de imágenes mediante el modelo de entrenamiento para clasificar imágenes no es una variable lo suficientemente sólida por sí sola para tener impacto significativo en el engagement del contenido en redes sociales. Sin embargo, es importante tener en cuenta que estos resultados son específicos para este caso en particular y podrían variar en función de otros factores como el tipo de contenido, el público objetivo, las Tendencias y la estrategia de marketing utilizada.

Al igual considerar las limitaciones de los modelos de aprendizaje automático y la necesidad de realizar una interpretación crítica de los resultados obtenidos, en el caso del marketing digital es un entorno en constante cambios guiados por tendencias y actualidad, por lo que se necesitan un constante seguimiento de las variables para tener una predicción del mercado digital más acertada.

6.3.2 Clasificación (Hierarchical Clustering) agrupadas por etiquetas y el orden automático

El tema que se abordó fue la clasificación de imágenes utilizando el algoritmo de agrupamiento jerárquico (Hierarchical Clustering) en función de las etiquetas asociadas a cada imagen (Figura 26). Además, se aplicó un ordenamiento automático basado en el nivel de engagement obtenido en cada publicación.

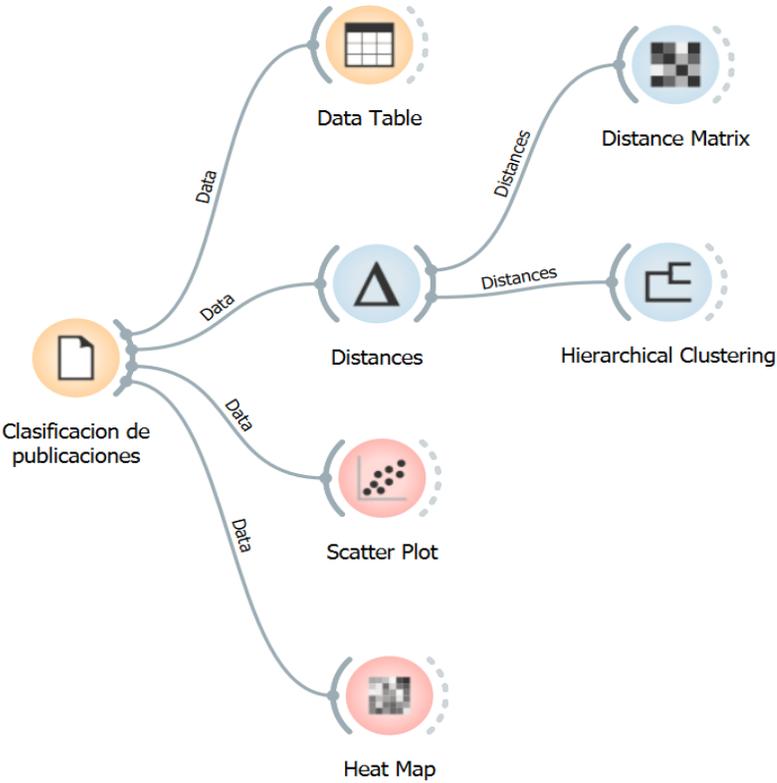


Figura 26. Flujo de datos para su clasificación hierarchical clustering.

Para llevar a cabo este proceso, se utilizó una base de datos de imágenes etiquetadas y se aplicó el algoritmo de agrupamiento jerárquico para clasificarlas en diferentes grupos. Luego, se utilizó una medida de similitud entre las imágenes en función de las etiquetas para construir el flujo de trabajo que representa la jerarquía de los grupos (Figura 27).

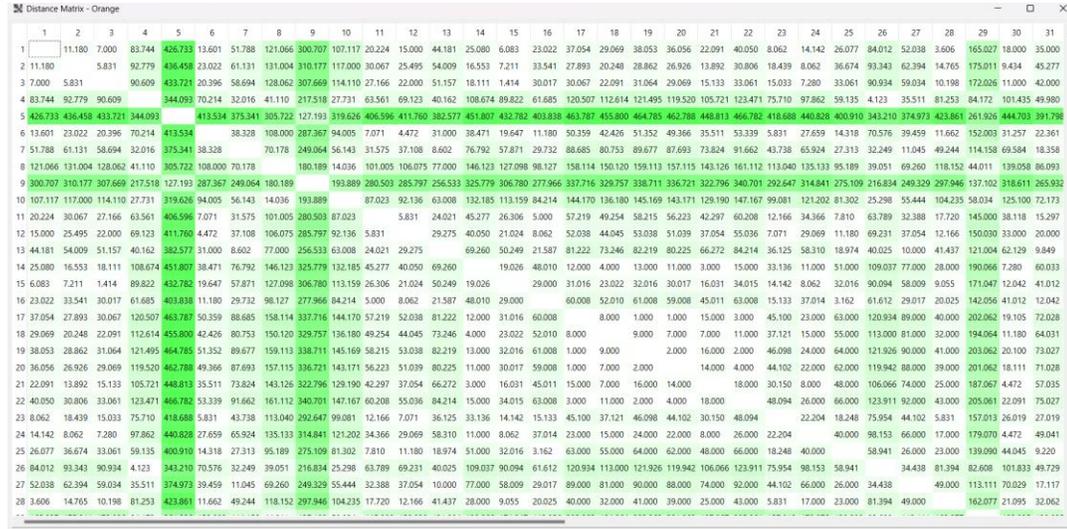


Figura 27. Visualización en forma de matriz de la distancia (diferencias) entre variables

Una vez que se clasificaron las imágenes, se aplicó un ordenamiento automático basado en el nivel de engagement obtenido en cada publicación. (Figura 28) De esta manera, se pudieron identificar los grupos de imágenes que generan mayor interacción y, por lo tanto, se espera que sean relevantes para el público.

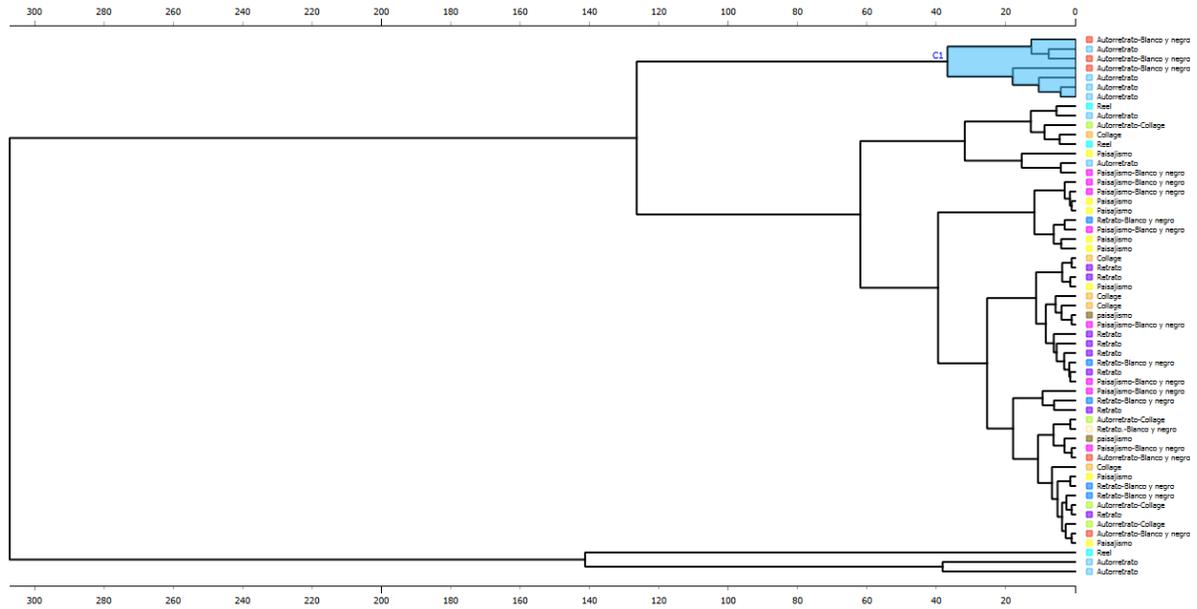


Figura 28. Visualización de grupos que se generaron por similitud.

El resultado de este proceso fue una clasificación de imágenes en diferentes grupos, ordenados automáticamente en función del nivel de engagement. Este enfoque puede resultar muy útil para los analistas de marketing y la publicidad, (Figura 29) ya que permite identificar rápidamente los tipos de imágenes que generan mayor interacción con el público y, por lo tanto, son más efectivas en la promoción de productos o servicios para ayudar a mejorar la estrategia de marketing y aumentar el éxito de las campañas publicitarias

	Tipo de fotografía	engagement	Orden en la clasificación automática	Cluster	Reacciones	Comentarios
1	Autorretrato	10.4502	9	C1	189	6
2	Autorretrato-Bl...	9.75348	52	C1	175	7
3	Autorretrato-Bl...	12.7546	43	C1	233	5
4	Autorretrato	12.0043	4	C1	217	7
5	Autorretrato	11.3076	7	C1	199	12
6	Autorretrato	10.7181	34	C1	190	10
7	Autorretrato-Bl...	12.2186	36	C1	224	4

Figura 29. Visualización del grupo seleccionado en las agrupaciones automáticas.

En conclusión, la clasificación de imágenes mediante el algoritmo de agrupamiento jerárquico y el ordenamiento automático por nivel de engagement puede ser una herramienta muy útil para los profesionales del marketing y la publicidad. Este enfoque permite identificar los grupos de imágenes más relevantes y efectivos para la promoción de productos o servicios, lo que puede

6.4 Evaluación de diferentes modelos de machine learning en predicción de engagement

Para evaluar la eficiencia de diferentes modelos de predicción del engagement según el tipo de fotografía, se cargaron los modelos en Orange, por la facilidad de ser plataforma de aprendizaje automático visual. Primero, se prepararon los datos de entrada a través de la recolección y selección de un conjunto de datos etiquetados que incluyó información sobre el tipo de fotografía y el nivel de engagement obtenido en cada publicación.

A continuación, se procedió a la construcción y entrenamiento de diferentes modelos de aprendizaje automático, incluyendo árboles de decisión, regresión logística y redes neuronales. Para la evaluación de los modelos, (Figura 30) se utilizaron diferentes técnicas de validación cruzada y se compararon las métricas de desempeño obtenidas, tales como la precisión y correlación cruzada. Finalmente, se seleccionó el modelo con mejor desempeño para realizar predicciones sobre el engagement de nuevas publicaciones según el tipo de fotografía.

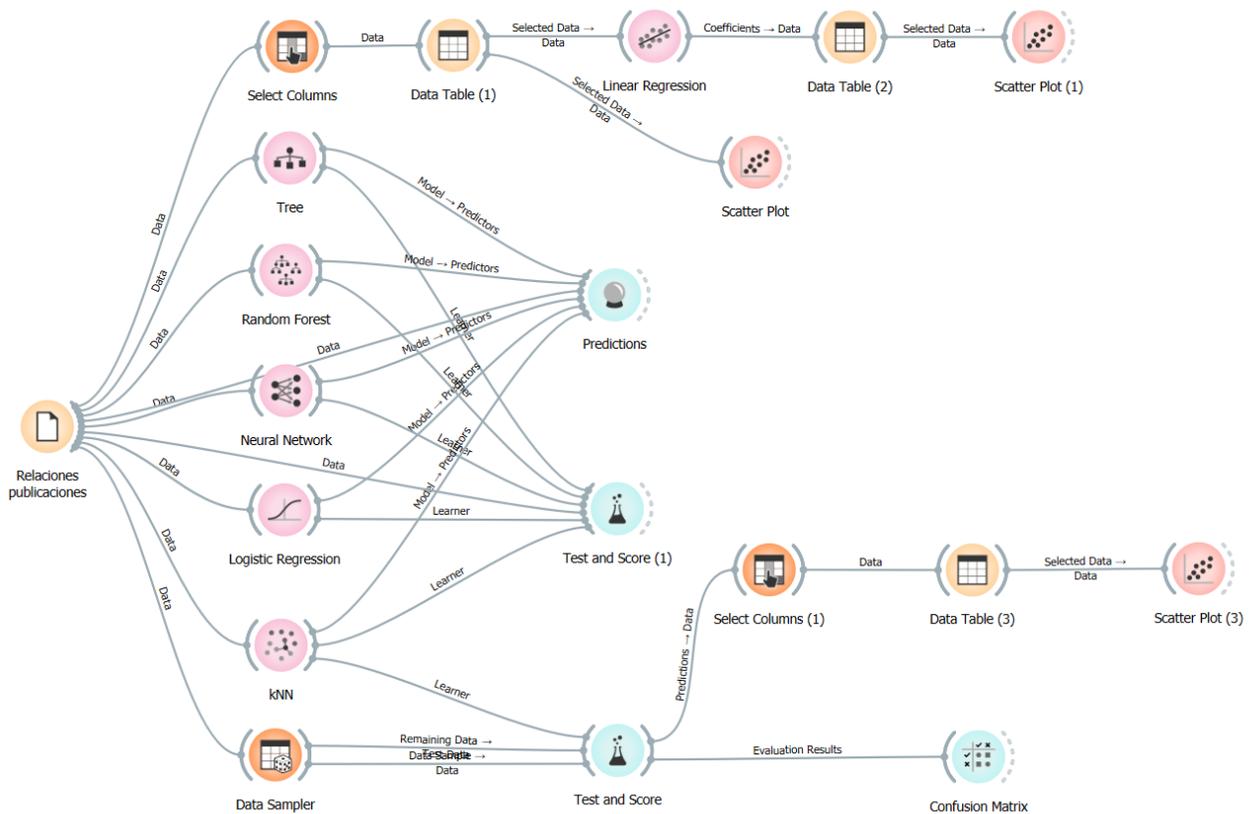


Figura 30. Flujo de datos para la evaluación y predicción de los múltiples modelos de machine learning.

Durante la evaluación de diferentes modelos de machine learning en la predicción de engagement, se utilizó un archivo que contenía una base de datos que pasó por diferentes nodos que tenían precargados los modelos de machine learning más comunes. Estos modelos, a su vez, se realizaron comparaciones en los nodos de predicciones y de testeo para determinar cuál de ellos ofrecía la mejor precisión en la predicción del engagement.

En cada corrida se evaluaron múltiples tipos de fotografía utilizando cada modelo de machine learning, lo que permitió obtener una visión más amplia y completa del rendimiento de cada modelo en diferentes contextos y situaciones. (Figura 31) A través de estos resultados, se pudo determinar qué modelos ofrecían la mejor precisión y qué tipos de fotografías eran más propensas a ser predichas con mayor precisión.

Este enfoque de evaluación sistemática y rigurosa de múltiples modelos de machine learning en la predicción de engagement permite identificar las fortalezas y debilidades de cada modelo y utilizar esta información para mejorar la precisión de la predicción en el futuro. Además, también

se utiliza para identificar patrones y tendencias en la base de datos que de otra manera podrían pasar desapercibidos.

	Tree	error	Logistic Regression	error	Neural Network	Tipo de fi
1	0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.67: 0.33: 0.00: 0.00 → Retrato	0.333	0.00: 0.00: 0.01: 0.02: 0.03: 0.03: 0.00: 0.87: 0.02: 0.00: 0.00 → Retrato	0.129	0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.99: 0.00: 0.00	Retrato
2	0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 1.00: 0.00: 0.00: 0.00 → Retrato	0.000	0.00: 0.00: 0.01: 0.02: 0.02: 0.02: 0.00: 0.90: 0.02: 0.00: 0.00 → Retrato	0.099	0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 1.00: 0.00: 0.00	Retrato
3	0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 1.00: 0.00: 0.00: 0.00 → Retrato	0.000	0.00: 0.00: 0.01: 0.02: 0.03: 0.03: 0.00: 0.88: 0.02: 0.00: 0.00 → Retrato	0.124	0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.99: 0.00: 0.00	Retrato
4	0.40: 0.00: 0.00: 0.00: 0.20: 0.20: 0.00: 0.00: 0.00: 0.00: 0.00 → Autorretrato	0.600	0.91: 0.00: 0.03: 0.01: 0.00: 0.00: 0.02: 0.02: 0.01: 0.00: 0.00 → Autorretrato	0.093	0.99: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00	Autorretr
5	0.67: 0.00: 0.00: 0.00: 0.00: 0.00: 0.33: 0.00: 0.00: 0.00: 0.00 → Autorretrato	0.667	0.09: 0.00: 0.00: 0.00: 0.00: 0.00: 0.91: 0.00: 0.00: 0.00: 0.00 → Reel	0.094	0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.99: 0.00: 0.00: 0.00: 0.00	Reel
6	0.00: 0.00: 0.33: 0.33: 0.00: 0.00: 0.00: 0.00: 0.33: 0.00: 0.00: 0.00 → Autorretrato-Collage	0.667	0.00: 0.00: 0.85: 0.02: 0.03: 0.03: 0.00: 0.04: 0.02: 0.01: 0.00 → Autorretrato-Collage	0.150	0.00: 0.00: 0.99: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00	Autorretr
7	0.00: 0.00: 0.20: 0.20: 0.00: 0.00: 0.20: 0.00: 0.20: 0.00: 0.00 → Autorretrato-Collage	0.800	0.02: 0.00: 0.87: 0.02: 0.01: 0.02: 0.01: 0.02: 0.01: 0.01: 0.00 → Autorretrato-Collage	0.127	0.00: 0.00: 0.99: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00	Autorretr
8	0.33: 0.67: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00 → Autorretrato-Blanco y ne...	0.667	0.92: 0.04: 0.00: 0.00: 0.00: 0.00: 0.03: 0.00: 0.00: 0.00: 0.00 → Autorretrato	0.076	1.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00	Autorretr
9	0.67: 0.00: 0.00: 0.00: 0.00: 0.00: 0.33: 0.00: 0.00: 0.00: 0.00 → Autorretrato	0.333	0.97: 0.00: 0.00: 0.00: 0.00: 0.00: 0.03: 0.00: 0.00: 0.00: 0.00 → Autorretrato	0.032	1.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00	Autorretr
10	0.75: 0.25: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00 → Autorretrato	0.750	0.06: 0.89: 0.00: 0.00: 0.00: 0.00: 0.04: 0.00: 0.00: 0.00: 0.00 → Autorretrato-Blanco y ne...	0.106	0.00: 0.99: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00	Autorretr
11	0.00: 0.00: 0.50: 0.50: 0.00: 0.00: 0.00: 0.00: 0.25: 0.00: 0.25: 0.00 → Autorretrato-Collage	0.500	0.01: 0.01: 0.85: 0.02: 0.03: 0.03: 0.00: 0.03: 0.02: 0.01: 0.00 → Autorretrato-Collage	0.151	0.00: 0.00: 0.99: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00	Autorretr
12	0.00: 0.00: 0.50: 0.50: 0.00: 0.00: 0.00: 0.00: 0.25: 0.00: 0.25: 0.00 → Autorretrato-Collage	0.500	0.01: 0.01: 0.83: 0.02: 0.03: 0.04: 0.00: 0.02: 0.02: 0.01: 0.00 → Autorretrato-Collage	0.165	0.00: 0.00: 0.99: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00	Autorretr
13	0.00: 0.00: 0.20: 0.20: 0.00: 0.00: 0.20: 0.00: 0.20: 0.00: 0.00 → Autorretrato-Collage	0.800	0.02: 0.01: 0.02: 0.84: 0.02: 0.03: 0.01: 0.02: 0.02: 0.01: 0.00 → Collage	0.160	0.00: 0.00: 0.00: 0.99: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00	Collage
14	0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.33: 0.00: 0.00: 0.33: 0.00: 0.33 → Paisajismo-Blanco y negro	0.667	0.00: 0.00: 0.01: 0.02: 0.04: 0.03: 0.00: 0.03: 0.85: 0.00: 0.01 → Retrato-Blanco y negro	0.147	0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.99: 0.00	Retrato-B
15	0.00: 0.00: 0.00: 0.00: 0.33: 0.67: 0.00: 0.00: 0.00: 0.00: 0.00 → Paisajismo-Blanco y negro	0.333	0.00: 0.00: 0.01: 0.02: 0.03: 0.88: 0.00: 0.03: 0.02: 0.00: 0.01 → Paisajismo-Blanco y negro	0.121	0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.99: 0.00: 0.00: 0.00: 0.00	Paisajism
16	0.00: 0.33: 0.00: 0.00: 0.00: 0.33: 0.00: 0.00: 0.00: 0.00: 0.33 → Autorretrato-Blanco y ne...	0.667	0.01: 0.02: 0.01: 0.01: 0.02: 0.87: 0.00: 0.01: 0.01: 0.01: 0.01 → Paisajismo-Blanco y negro	0.126	0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.99: 0.00: 0.00: 0.00: 0.00	Paisajism
17	0.00: 0.00: 0.00: 0.00: 0.60: 0.40: 0.00: 0.00: 0.00: 0.00: 0.00 → Paisajismo	0.400	0.00: 0.00: 0.00: 0.02: 0.89: 0.03: 0.00: 0.03: 0.02: 0.00: 0.01 → Paisajismo	0.107	0.00: 0.00: 0.00: 0.00: 0.00: 0.99: 0.00: 0.00: 0.00: 0.00: 0.00	Paisajism
18	0.00: 0.00: 0.00: 0.00: 0.60: 0.40: 0.00: 0.00: 0.00: 0.00: 0.00 → Paisajismo	0.400	0.00: 0.00: 0.01: 0.02: 0.89: 0.03: 0.00: 0.03: 0.02: 0.00: 0.01 → Paisajismo	0.109	0.00: 0.00: 0.00: 0.00: 1.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00	Paisajism
19	0.00: 0.00: 0.00: 0.00: 0.60: 0.40: 0.00: 0.00: 0.00: 0.00: 0.00 → Paisajismo	0.400	0.00: 0.00: 0.00: 0.02: 0.89: 0.03: 0.00: 0.03: 0.02: 0.00: 0.01 → Paisajismo	0.107	0.00: 0.00: 0.00: 0.00: 0.99: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00	Paisajism
20	0.00: 0.00: 0.00: 0.00: 0.60: 0.40: 0.00: 0.00: 0.00: 0.00: 0.00 → Paisajismo	0.600	0.00: 0.00: 0.01: 0.02: 0.04: 0.87: 0.00: 0.03: 0.02: 0.00: 0.01 → Paisajismo-Blanco y negro	0.128	0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.99: 0.00: 0.00: 0.00: 0.00	Paisajism
21	0.00: 0.00: 0.00: 0.00: 0.33: 0.00: 0.00: 0.00: 0.33: 0.00: 0.33 → Paisajismo-Blanco y negro	0.667	0.00: 0.00: 0.01: 0.02: 0.04: 0.88: 0.00: 0.03: 0.02: 0.00: 0.01 → Paisajismo-Blanco y negro	0.125	0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.99: 0.00: 0.00: 0.00: 0.00	Paisajism
22	0.00: 0.00: 0.00: 0.00: 0.60: 0.40: 0.00: 0.00: 0.00: 0.00: 0.00 → Paisajismo	0.600	0.00: 0.00: 0.00: 0.02: 0.04: 0.87: 0.00: 0.03: 0.02: 0.00: 0.01 → Paisajismo-Blanco y negro	0.129	0.00: 0.00: 0.00: 0.00: 0.00: 0.99: 0.00: 0.00: 0.00: 0.00: 0.00	Paisajism
23	0.00: 0.25: 0.00: 0.00: 0.50: 0.00: 0.00: 0.00: 0.25: 0.00: 0.00 → Paisajismo	0.500	0.00: 0.01: 0.01: 0.02: 0.87: 0.03: 0.00: 0.03: 0.02: 0.00: 0.00 → Paisajismo	0.126	0.01: 0.00: 0.00: 0.00: 0.98: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00	Paisajism
24	0.00: 0.00: 0.00: 0.00: 0.00: 0.33: 0.00: 0.00: 0.33: 0.00: 0.33 → Paisajismo-Blanco y negro	0.667	0.00: 0.00: 0.01: 0.02: 0.04: 0.03: 0.00: 0.03: 0.02: 0.00: 0.85 → paisajismo	0.154	0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00: 0.00	Paisajism

Model	AUC	CA	F1	Precision	Recall
Tree	0.934	0.526	0.472	0.458	0.526
Logistic Regression	1.000	1.000	1.000	1.000	1.000
Neural Network	1.000	1.000	1.000	1.000	1.000
kNN	0.880	0.386	0.321	0.280	0.386
Random Forest	1.000	0.982	0.981	0.984	0.982

Figura 31. Pruebas por cada modelo en cada variable y matriz de resultados con sus diferencias.

La matriz comparativa entre los modelos nos permite evaluar su desempeño en términos de diferentes indicadores, como el AUC (área bajo la curva ROC), la precisión, la exhaustividad (recall) y la puntuación F1. Estos indicadores nos permiten determinar cuál modelo es el más adecuado para predecir el engagement de los usuarios con las diferentes fotografías.

La matriz comparativa (Figura 32) muestra el desempeño de cada modelo para cada tipo de fotografía. Al comparar los resultados, podemos identificar cuál modelo se desempeña mejor para cada tipo de fotografía. Además, podemos identificar los modelos que tienen un desempeño consistente en todos los tipos de fotografía evaluados.

Es importante tener en cuenta que la elección del modelo adecuado depende de los objetivos y requisitos específicos del proyecto. Por lo tanto, es importante evaluar cuidadosamente los resultados de la matriz comparativa y seleccionar el modelo que mejor se adapte a las necesidades del proyecto.

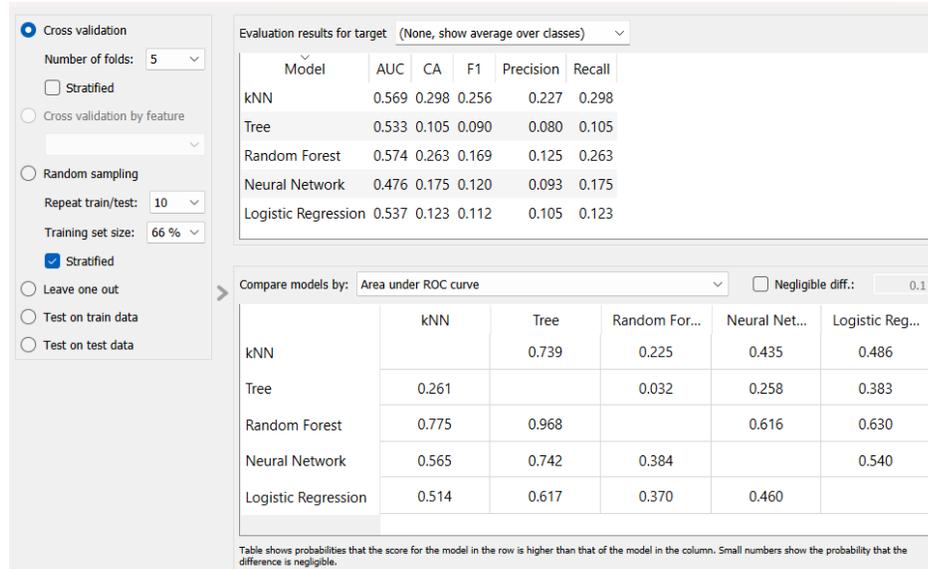


Figura 32. Revisión a detalle de las diferencias entre modelos.

En base a los resultados obtenidos por los modelos de machine learning evaluados, se puede concluir que ninguno modelo sobresale entre los demás modelos en la predicción de engagement en fotografías. El modelo con el mejor desempeño fue Random Forest con un AUC de 0.574, pero aun así no logró superar un nivel satisfactorio de predicción. Además, los valores de CA, F1, Precision y Recall para todos los modelos fueron bastante bajos, lo que indica que hay mucho espacio para mejorar la capacidad de predicción de estos modelos. Los resultados pueden estar influenciados por la cantidad de información disponible en la base de datos.

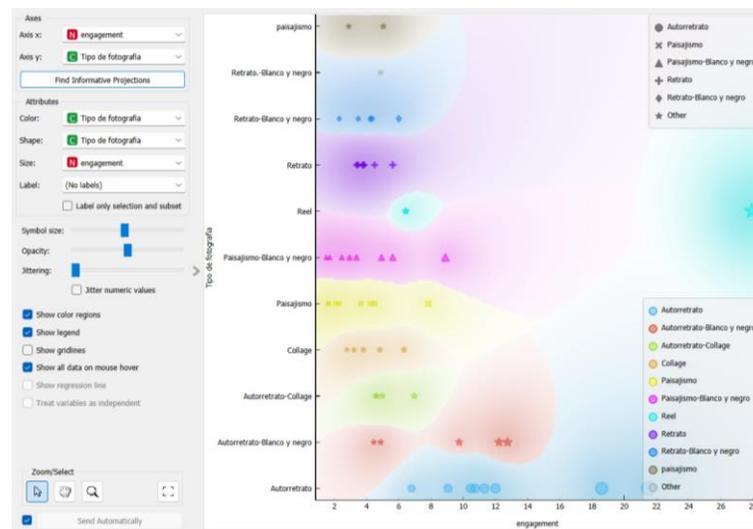


Figura 33. Grafica de segmentación por tipo de publicación.

6.5 Análisis de la página web con los datos de google analytics

El objetivo de esta sección de la metodología es realizar un análisis exhaustivo de la página web utilizando los datos obtenidos a través de Google Analytics. Para ello, se utilizarán diferentes herramientas y técnicas para obtener información relevante sobre el comportamiento de los usuarios en el sitio web (Figura 35).

En primer lugar, se procederá a la configuración de Google Analytics, estableciendo los objetivos y las conversiones que se desean medir en la página web. A continuación, se analizará la información proporcionada por Google Analytics, como el tráfico de la página web, el comportamiento de los usuarios, los canales de adquisición, las fuentes de tráfico y el rendimiento de la página web en diferentes dispositivos.

Una vez recopilada y organizada la información, se procederá a la identificación de patrones y tendencias en el comportamiento de los usuarios. Se utilizarán técnicas de análisis exploratorio de datos para visualizar y resumir los datos obtenidos. También se utilizarán técnicas de análisis de cohortes para comprender mejor el comportamiento de los usuarios a lo largo del tiempo y su relación con los objetivos y conversiones establecidos.

Además, se llevará a cabo un análisis comparativo de diferentes páginas de la página web, identificando aquellas que tienen un mayor rendimiento en términos de engagement y conversiones. Se analizará el contenido y la estructura de estas páginas para identificar las mejores prácticas y replicarlas en otras páginas de la página web.

Finalmente, se utilizarán técnicas de segmentación para dividir a los usuarios en diferentes grupos en función de su comportamiento en la página web. Se analizarán los diferentes grupos de usuarios para comprender mejor sus necesidades y preferencias, y así poder personalizar la experiencia en la página web y mejorar el engagement y las conversiones.

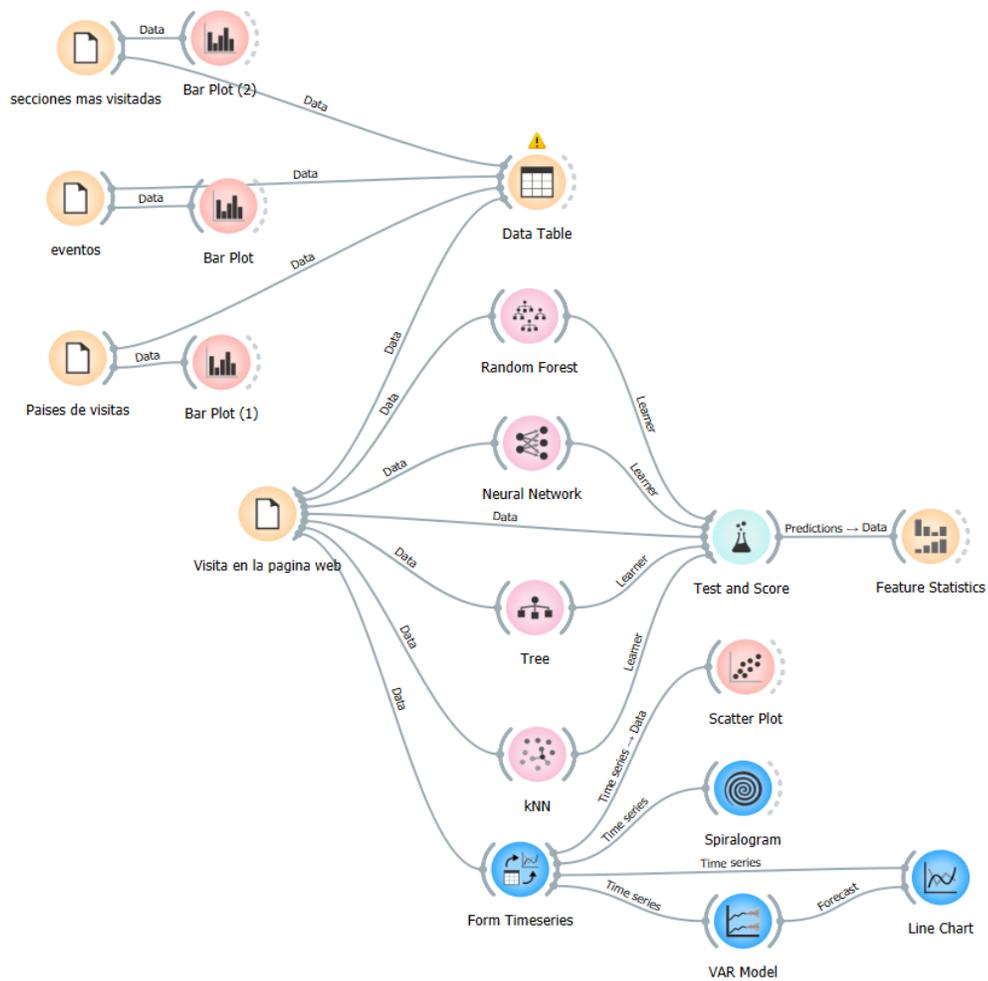


Figura 35. Flujo de datos para el análisis de adquisición e interacciones de los usuarios.

En resumen, el análisis de la página web con los datos de Google Analytics permitirá obtener información valiosa sobre el comportamiento de los usuarios en la página web, identificar patrones y tendencias, comparar el rendimiento de diferentes páginas, y segmentar a los usuarios para personalizar la experiencia y mejorar las conversiones.

6.5.1 Análisis por series de tiempo (Adquisición de usuarios y tiempo de interacción)

Para medir las interacciones de usuarios a lo largo del tiempo se optó por realizar un análisis por series de tiempo, que es considerada como una técnica estadística que nos permite estudiar la evolución de una variable a lo largo del tiempo. (En este caso será por la adquisición de usuarios nuevos y tiempo de interacción promedio) dentro del contexto del estudio de la página web, el análisis por series de tiempo es una herramienta fundamental para comprender cómo evoluciona la adquisición de usuarios y el tiempo de interacción en la página web.

Para llevar a cabo el análisis por series de tiempo, se recolectaron datos a lo largo de un período de tiempo determinado utilizando herramientas de analítica web, en este caso el uso de Google Analytics (Figura 36).

A continuación, se utilizan técnicas de análisis estadístico para identificar patrones y tendencias en los datos a lo largo del tiempo. Se pueden utilizar diferentes herramientas para realizar el análisis por series de tiempo, como ARIMA (AutoRegressive Integrated Moving Average) y modelos de suavizado exponencial.

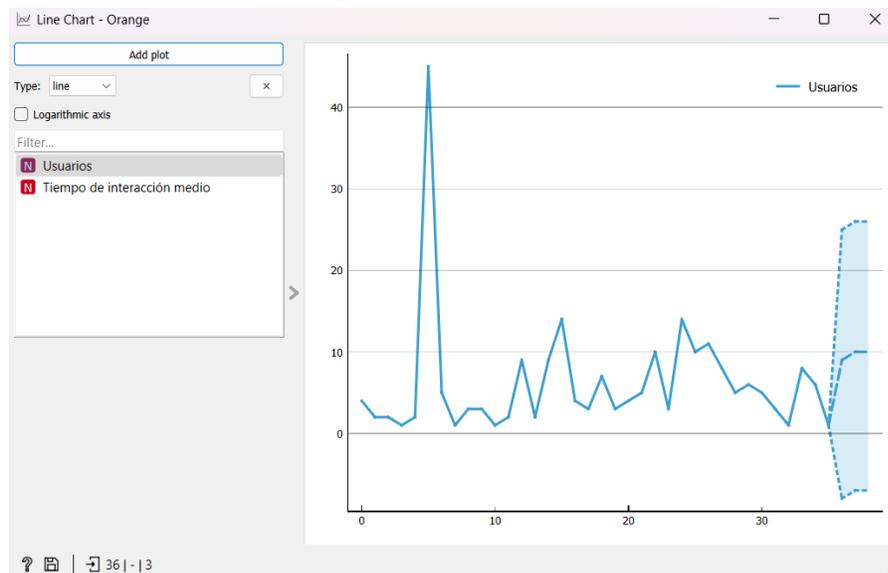


Figura 36. Gráfica se series de tiempo en intervalos de semanas sobre usuarios adquiridos en 2022.

Se puede observar en la gráfica que la actividad presenta picos constantes en la serie de tiempo. Cabe destacar que se registró un pico de visualizaciones en la página web que alcanzó un máximo de 45 usuarios, el cual coincide con la publicación realizada en Instagram que dirigió tráfico hacia la página web.

Una vez que se han identificado patrones y tendencias en los datos, se pueden realizar pronósticos y predicciones sobre la evolución futura de la adquisición de usuarios y el tiempo de interacción en la página web. Estos pronósticos y predicciones son útiles para la toma de decisiones estratégicas, ya que permiten identificar posibles problemas y oportunidades a largo plazo.

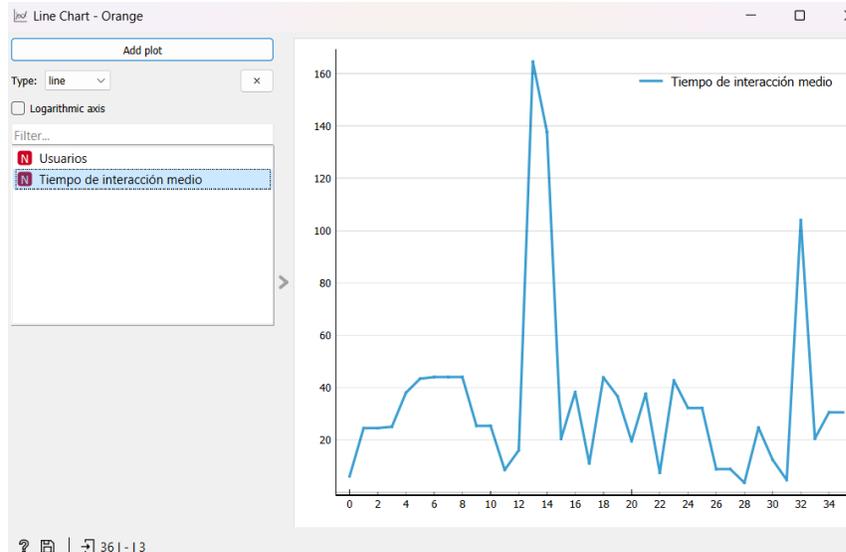


Figura 37. Grafica se series de tiempo en intervalos de semanas sobre tiempo de interacción realizado en 2022.

En el análisis de la serie de tiempo que representa el tiempo que los usuarios (Figura 37) pasan en la página web a lo largo del tiempo, se observa que se registraron dos picos notables. El primero, el más grande, muestra un tiempo promedio de 160 segundos en la página web, lo cual indica que los usuarios pasaron un tiempo considerable navegando por el sitio. El segundo pico, aunque no tan alto como el primero, también muestra un tiempo promedio de más de 100 segundos, lo que indica que los usuarios se mantuvieron comprometidos con el contenido de la página web durante un período de tiempo significativo. Estos picos podrían explicarse por la presencia de contenido de alta calidad y relevante que mantuvo a los usuarios involucrados, pero en el momento donde no había publicaciones tenían caídas de interacciones.

En resumen, el análisis por series de tiempo es una técnica estadística esencial para comprender la evolución de la adquisición de usuarios y el tiempo de interacción en una página web. A través de esta técnica, se pueden identificar patrones y tendencias en los datos, realizar pronósticos y predicciones, y tomar decisiones estratégicas basadas en datos objetivos y precisos.

6.5.2 Evaluación de predicciones de múltiples modelos de machine learning

La evaluación de predicciones de múltiples modelos de machine learning es una técnica que nos permite identificar el modelo que mejor se ajusta a nuestros datos. En este caso, los modelos utilizados son kNN, Tree, Random Forest y Neural Network.

Para llevar a cabo esta evaluación (Figura 38), se utilizará una técnica conocida como validación cruzada, que consiste en dividir los datos en un conjunto de entrenamiento y un conjunto de prueba. El conjunto de entrenamiento se utilizará para entrenar el modelo, mientras que el conjunto de prueba se utilizará para evaluar el rendimiento del modelo.

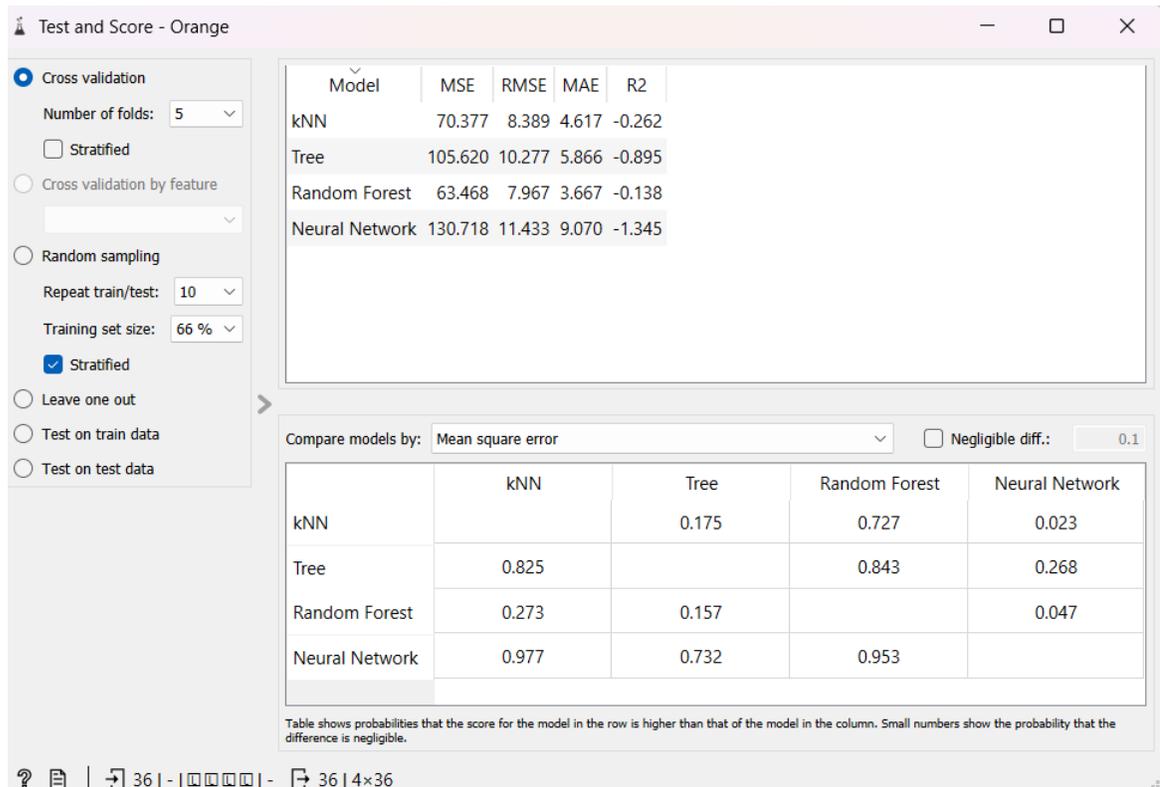


Figura 38. Resumen de diferencias entre modelos de machine learning por los datos de google analytics.

Una vez que se han entrenado los modelos, se procederá a evaluar su desempeño utilizando diferentes métricas, como la precisión, la sensibilidad y la especificidad. Además, se comparará el rendimiento de cada modelo para determinar cuál ofrece la mejor precisión.

También se realizará una evaluación de los resultados de cada modelo utilizando diferentes técnicas, como la matriz de confusión y la curva ROC. Estas técnicas permiten identificar las fortalezas y debilidades de cada modelo en términos de su capacidad para clasificar correctamente los datos.

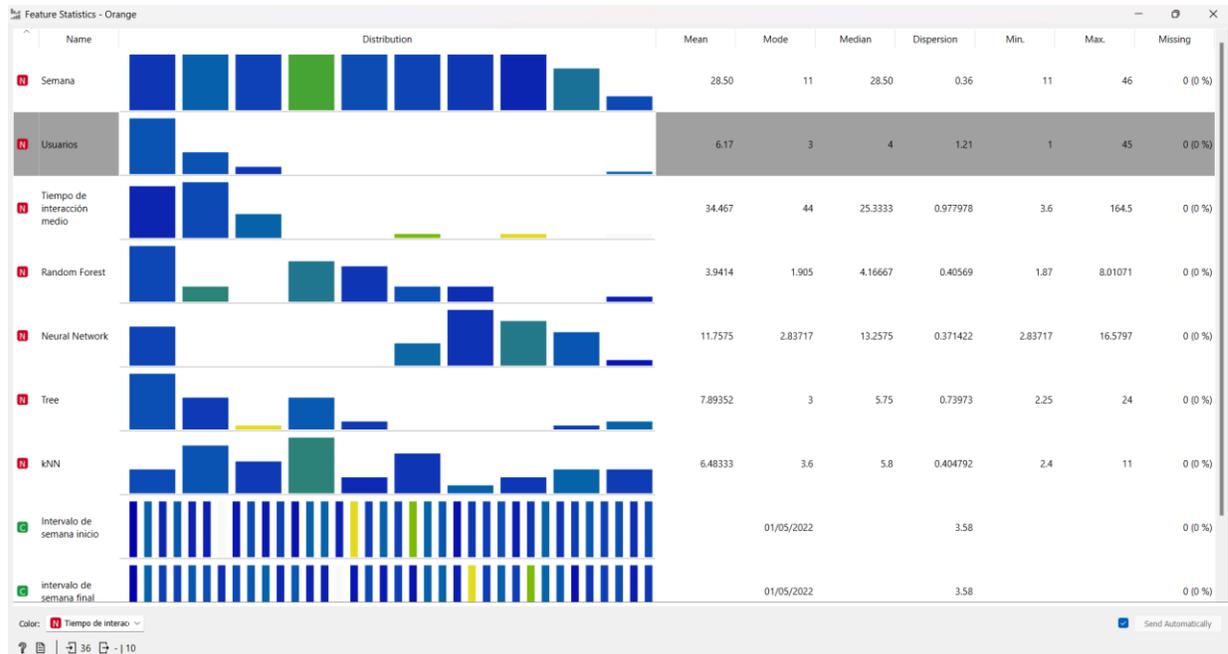


Figura 39. Resumen gráfico de las variables generadas por los modelos y sus promedios.

En esta sección, (Figura 39) se presentan tanto un resumen como las generalidades de los datos que se generaron en Google Analytics, así como su representación en histogramas. También se incluye la representación numérica que se obtuvo a través de los modelos de machine learning, los cuales se basaron en los datos generados.

Para el uso de los diferentes modelos de machine learning, se seleccionará el modelo que ofrezca el mejor rendimiento para utilizarlo en la predicción de nuevos datos. Además, se discutirán las implicaciones y limitaciones de los modelos de machine learning utilizados en este estudio, y se propondrán posibles mejoras para futuras investigaciones en el campo.

6.5.3 Histograma de acciones más realizadas dentro de la página web, secciones más visitadas y países que visualizaron la página web

Continuando con otros datos de google analytics se consideró un análisis en forma de histograma de acciones más realizadas dentro de la página web (Figura 40) permitirá conocer las actividades que los usuarios realizan con mayor frecuencia en la página web, lo que a su vez permitirá identificar los aspectos de la página web que resultan más relevantes para los usuarios. Para este análisis se utilizarán los datos de Google Analytics que solo tienen dos variables una categórica y otra numérica.

En el histograma anterior se destaca el área de proyectos como la sección más popular, seguida por una galería de fotografías de una cafetería y por último, la sección que describe el perfil de la fotógrafa en sí. Este resultado era de esperar, considerando el tipo de contenido que ofrece la página web.

Finalmente, se evaluará la ubicación geográfica de los usuarios (Figura 42) que visualizan la página web para conocer la audiencia internacional de la página web. Para este análisis se utilizarán los datos de la sección de audiencia de la página web, podemos observar que el público principal es mexicano, sin embargo también se visualiza un mercado potencial estadounidense.

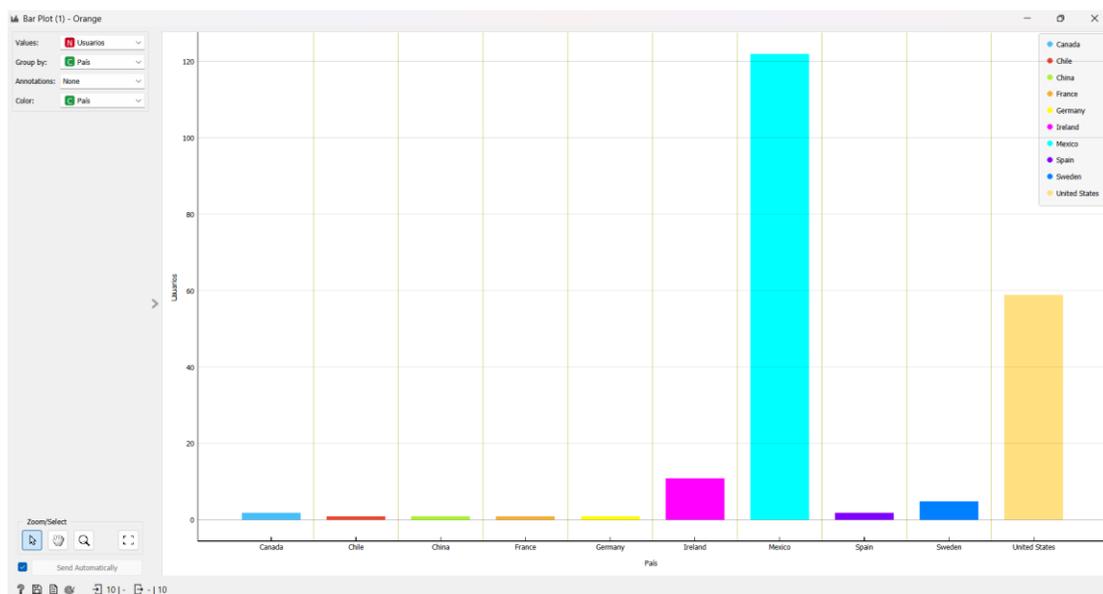


Figura 42. Histograma de los países donde fue visualizada la página web.

Todos estos análisis permiten conocer los patrones de uso del usuario y plantear estrategias al respecto, así como también los aspectos que no resultan tan relevantes para ellos para considerar hacerlos más atractivos. De esta manera, se podrán identificar oportunidades de mejora en la página web y se podrá diseñar una estrategia para aumentar el tráfico y la interacción de los usuarios en el contenido.

6.6 Análisis de texto en SKAGS y hashtags

Por último recurso por analizar utilizando un apartado llamado text mining aplicado sobre un análisis de texto en SKAGS y hashtags es una herramienta importante en la búsqueda y comprensión de cómo los usuarios interactúan con una marca o producto en línea. Para llevar a cabo este análisis, se puede utilizar una variedad de herramientas estadísticas, incluyendo el análisis de frecuencia de palabras clave (Figura 43) y la identificación de patrones de comportamiento de los usuarios (análisis semántico).



Figura 43. Nube de palabras basado en los hashtags utilizados, de las más generales (palabras grandes) a las más específicas (palabras pequeñas) como etiqueta.

En este estudio, se pretende analizar la frecuencia de palabras clave y hashtags en los anuncios de SKAGS (Single Keyword Ad Groups) de una marca específica. Para ello, se recolectarán datos de las campañas publicitarias de la marca en varias plataformas, como Google Ads y Facebook Ads (Figura 44).



Figura 44. Aplicación del modelo de tópicos por cada palabra ingresada de los hashtags.

Una vez recopilados los datos, se utilizarán técnicas de minería de texto para analizar la frecuencia de palabras clave y hashtags en los anuncios de SKAGS (Figura 45). También se analizará la relación entre los términos clave y los hashtags más utilizados y la interacción de los usuarios con ellos.

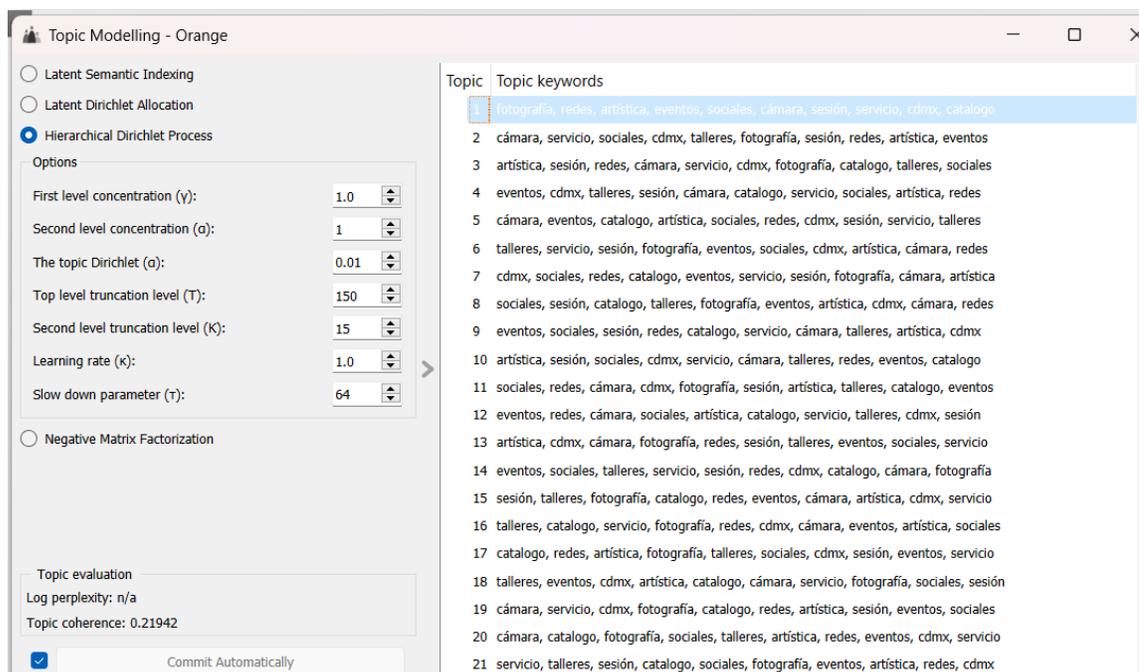


Figura 45. Lista de términos desplegados por SKAGS.

Además, se utilizarán herramientas de análisis de sentimiento para determinar el tono general de los mensajes en los anuncios de SKAGS y en los hashtags utilizados en las publicaciones en línea (Figura 46). Esto ayudará a comprender cómo los usuarios perciben la marca y cómo la marca puede mejorar su comunicación y estrategias de marketing en línea.

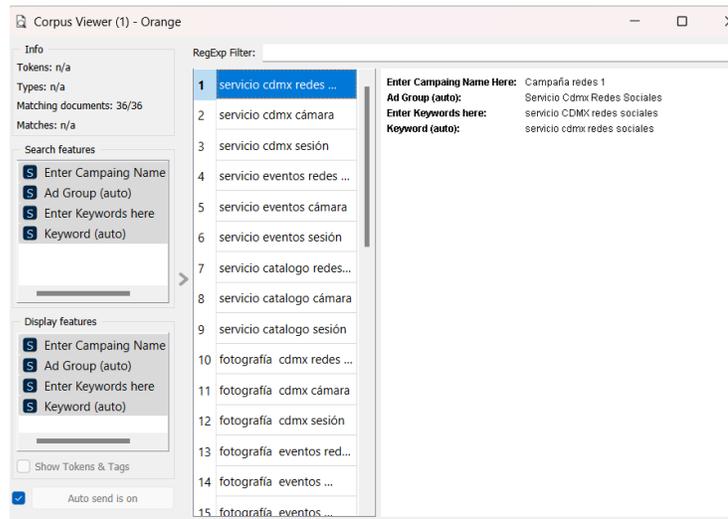


Figura 46. Visualización de metadatos por combinación de SKAGS.

En resumen, el análisis de texto en SKAGS y hashtags es una herramienta valiosa en la evaluación de la efectividad de las campañas publicitarias en línea y en la comprensión de cómo los usuarios interactúan con una marca en línea (Figura 47). A través de técnicas de minería de texto y análisis de sentimiento, se puede obtener una comprensión más profunda de los patrones de comportamiento de los usuarios y cómo mejorar la estrategia de marketing en línea de la marca.

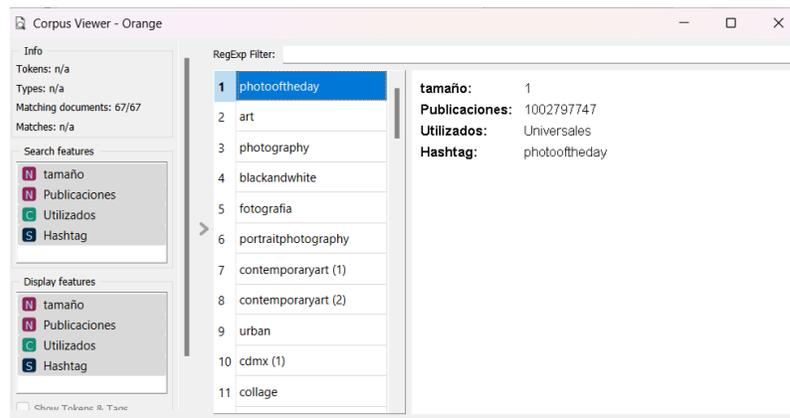


Figura 47. Visualización de metadatos por hashtags.

En conclusión, el análisis de texto en SKAGS y hashtags es una metodología útil para la investigación de la interacción de los usuarios con una marca o producto en línea. La recopilación de datos y su análisis a través de técnicas de minería de texto y relevancia de estos términos

que permite obtener una comprensión más profunda de los patrones de comportamiento de los usuarios y su percepción de la marca.

Esta información es valiosa para mejorar las estrategias de marketing en línea y la comunicación con los usuarios, con el objetivo de lograr una mayor efectividad en las campañas publicitarias y una mejor relación con los clientes. En definitiva, el análisis de texto en SKAGS y hashtags es una herramienta clave para el éxito de cualquier marca o negocio en línea.

6.6.1 Análisis con Google trends

Para complementar el uso de data mining hacemos uso de Google Trends ya que resulta ser una herramienta muy útil para medir el impacto de las palabras claves destacadas en una campaña de redes sociales. Específicamente, para analizar el rendimiento de las palabras clave seleccionadas en los hashtags y SKAGS con los datos que tiene Google. Al ingresar estas palabras en las tendencias del motor de búsqueda, es posible analizar su rendimiento a lo largo del tiempo y en diferentes regiones geográficas. De esta manera, se puede determinar la relevancia y popularidad de las palabras clave en la audiencia objetivo (Figura 48) y ajustar la estrategia de marketing en consecuencia.



Figura 48. Términos seleccionados entre hashtags y SKAGS en Google trends.

En la metodología, se procedió a utilizar 5 términos en Google Trends, seleccionados entre hashtags y skags los cuales fueron seleccionados por ser los más generales y comúnmente utilizados en las búsquedas. Los términos utilizados fueron "cdmx", "bellas artes", "contemporary art", "estudio fotográfico" y "cursos de fotografía". Después de analizar los resultados obtenidos, se encontró que el término "cdmx" es el más solicitado por un amplio margen (Figura 50). Si se elimina este término, se puede observar que "bellas artes" es el segundo término más buscado (Figura 49), mientras que los otros 3 términos se mantienen constantes aunque no son tan

populares como los primeros dos. Este análisis fue realizado para determinar la relevancia y el impacto de estas palabras clave en las búsquedas del motor de Google.

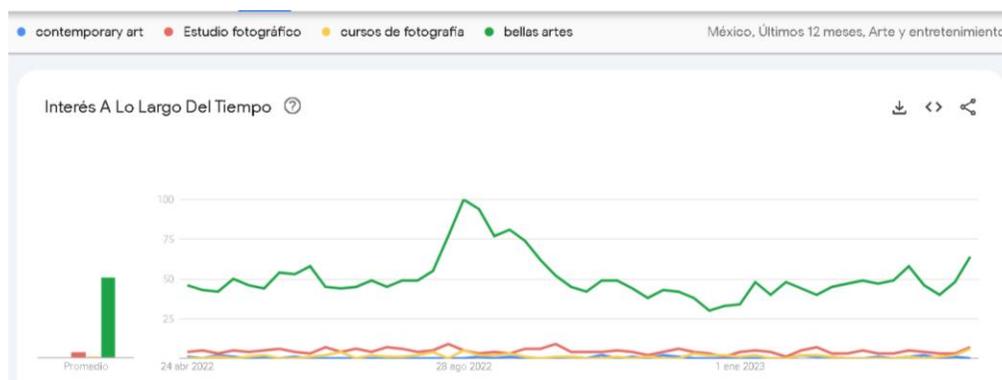


Figura 49. Visualización de series de tiempo sobre los términos ingresados (sin la etiqueta cdmx)

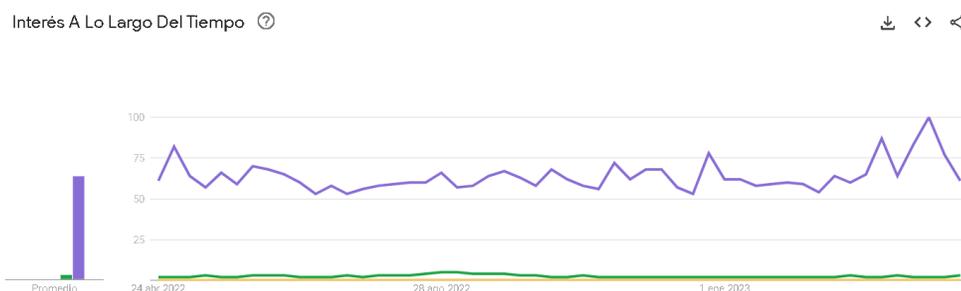


Figura 50. Visualización de series de tiempo sobre los términos ingresados (con la etiqueta cdmx)

Podemos mencionar que, en cuanto a los términos "cursos de fotografía" y "estudio fotográfico", los cuales están basados en los SKAGS de la página web, se observó que "estudio fotográfico" (Figura 52) es un término más solicitado en todo el territorio nacional, mientras que "cursos de fotografía" tiene mayor presencia en algunas zonas, como Monterrey y Puebla (Figura 51), además de la Ciudad de México. Estos resultados pueden proporcionar un nuevo enfoque de mercado para esas ciudades y así adaptar la estrategia de marketing de la página web en

consecuencia. Es importante destacar que estos resultados fueron obtenidos gracias al uso de Google Trends y la correlación con los SKAGS utilizados en la página web.

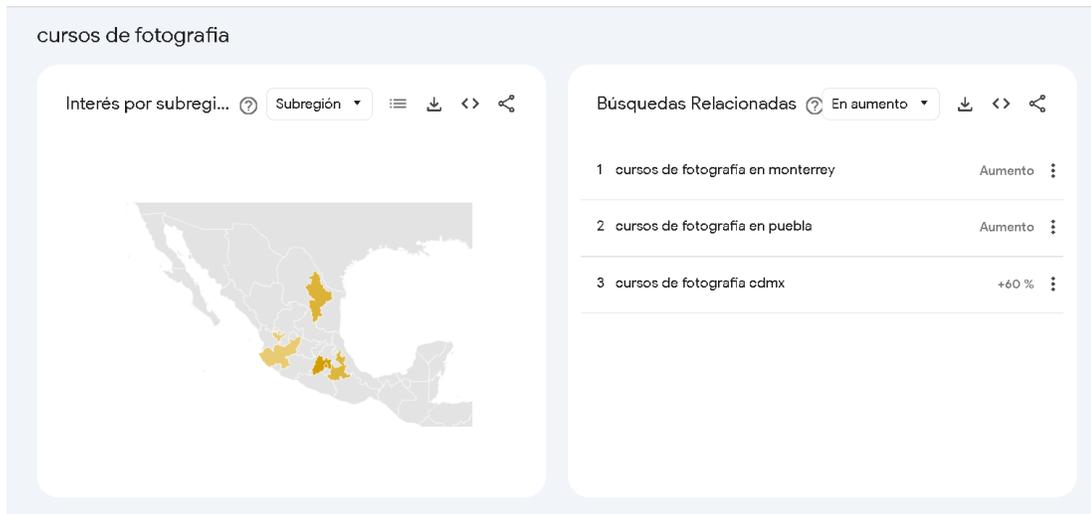


Figura 51. Búsqueda a nivel nacional basado en SKAGS más buscados. (Primera parte)

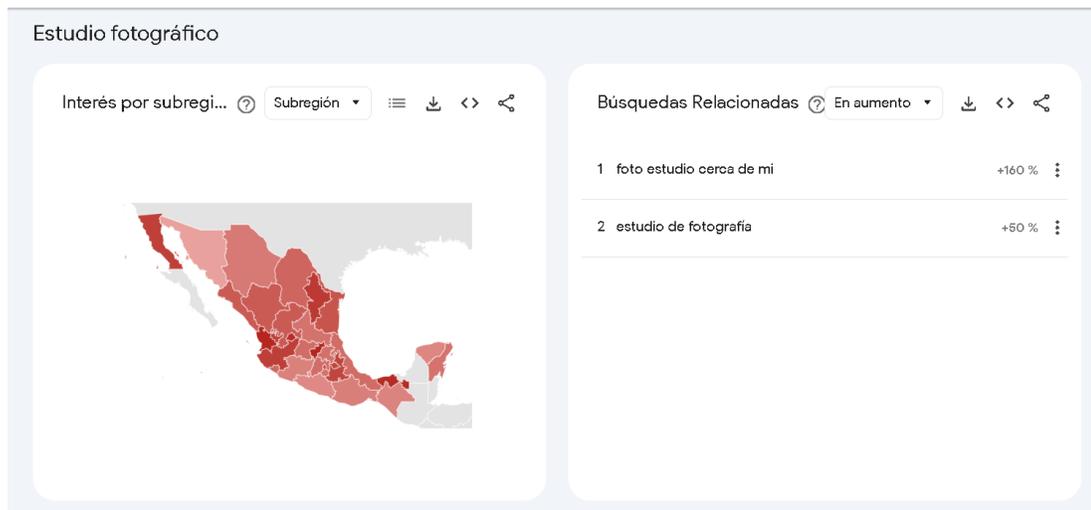


Figura 52. Búsqueda a nivel nacional basado en SKAGS más buscados. (Segunda parte)

En conclusión, el uso de herramientas como Google Trends nos permite analizar y comprender el impacto de las palabras clave en la búsqueda de contenido en línea. En este caso, se utilizó una combinación de términos generales y específicos relacionados con el contenido de la página web de fotografía.

Los resultados mostraron que, aunque algunos términos fueron más solicitados que otros, la combinación de todos ellos proporcionó información valiosa sobre el interés del público en relación con el contenido de la página. Además, se descubrieron patrones interesantes sobre el uso de ciertos términos en diferentes regiones geográficas, lo que podría proporcionar un enfoque de mercado más específico para ciertas áreas.

En resumen, el uso de herramientas como Google Trends puede ayudar a los propietarios de sitios web a comprender mejor a su audiencia y optimizar su estrategia de marketing en línea.

6.7 Estrategias consideradas e implementaciones realizadas en la campaña de marketing

Para la campaña de marketing del estudio Isabel Taracena, se consideraron varias estrategias con el objetivo de maximizar el impacto y la visibilidad de la marca. En primer lugar, se realizó lo que ya se tenía para trabajar sobre eso (página web y redes sociales).

Para posicionar una página web a través de las redes sociales, se debe considerar el alcance de la audiencia, el engagement y la reputación. El alcance de la audiencia se refiere a la cantidad de personas que ven el contenido de la página web. El engagement se refiere a la cantidad de personas que interactúan con el contenido de la página web. La reputación se refiere a la opinión que tienen las personas sobre el contenido y optimización de la página web.

Otra estrategia implementada fue la optimización del contenido de la página web, tanto para mejorar el posicionamiento en buscadores como para brindar una mejor experiencia de usuario. Se realizó un análisis de las palabras clave más relevantes para el negocio y se trabajó en la optimización de los textos y las imágenes en la página web.

En resumen, la campaña de marketing de la microempresa consideró diversas estrategias, desde el posicionamiento en redes sociales hasta la publicidad en buscadores, se trabajó en la optimización del contenido de la página web y se realizó un seguimiento y análisis de los resultados utilizando herramientas de análisis de datos y técnicas de Machine Learning. Todo esto con el objetivo de maximizar el impacto y la visibilidad de la marca y mejorar la efectividad de la campaña de marketing.

6.7.1 Estrategia AIDA

En el ámbito del marketing, el embudo de ventas (Ver Figura 53. Representación de un embudo de ventas es una herramienta utilizada para visualizar el proceso que sigue un cliente potencial desde el momento en que tiene un primer contacto con la marca hasta el momento en que realiza una compra. Por otro lado, el modelo AIDA (ver Figura 54) es una estrategia que se utiliza para guiar a los consumidores a través del proceso de toma de decisiones. En este subtema, se planteará cómo el embudo de ventas y el modelo AIDA tienen similitudes para poder utilizarlos juntos y como fue aplicado durante la campaña de marketing.

El embudo de ventas y el modelo AIDA tienen similitudes en cuanto a la estructura y el propósito. Ambos están diseñados para guiar a los consumidores a través del proceso de toma de decisiones y convertirlos en clientes. El modelo AIDA se compone de cuatro etapas: atención, interés, deseo y acción, mientras que el embudo de ventas se compone de varias etapas, que pueden variar dependiendo de la empresa y el sector en el que se opera. Sin embargo, en ambos casos, la idea es que los consumidores son participes en las diferentes etapas antes de realizar una compra.

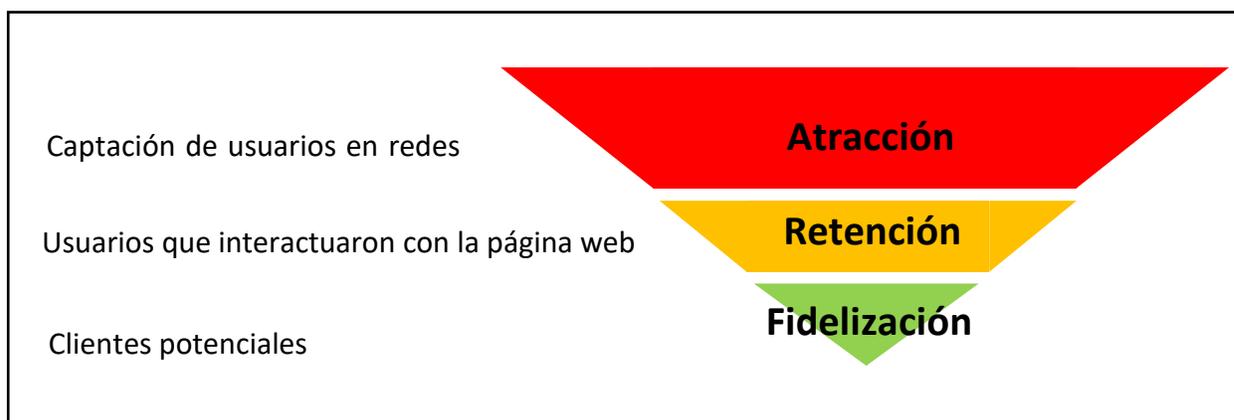


Figura 53. Representación de un embudo de ventas

La primera etapa en el modelo AIDA es la atención, donde se busca captar la atención del consumidor y hacer que se interese en el producto o servicio. En el embudo de ventas, esto se traduce en la fase de concienciación, donde se busca crear conciencia sobre la marca o el producto. La siguiente etapa en ambos casos es el interés, donde se busca retener la atención del consumidor y destacar las características y beneficios del producto. En el embudo de ventas,

esto se traduce en la fase de interés, donde se busca generar dar a conocer el producto y sus características.

La tercera etapa en el modelo AIDA es el deseo, donde se busca generar una necesidad percibida en el consumidor para adquirir el producto o servicio. En el embudo de ventas, esto se traduce en la fase de decisión, donde se busca que el consumidor tome la decisión de comprar el producto. Por último, la cuarta etapa del modelo AIDA es la acción, donde se busca que el consumidor realice la compra. En el embudo de ventas, esto se traduce en la fase de fidelización, donde se busca que el consumidor adquiera el producto y se convierta en un cliente.

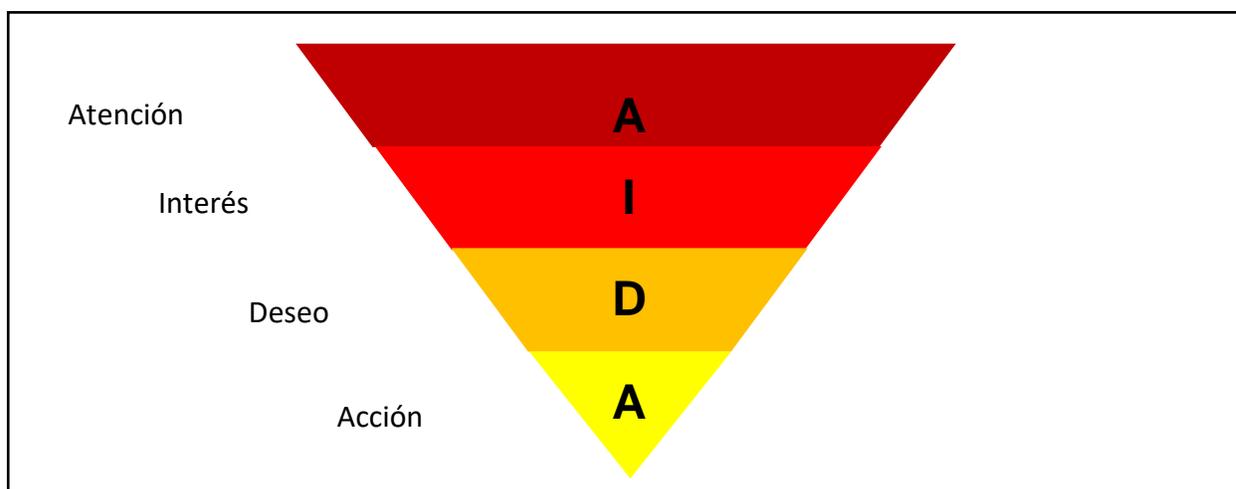


Figura 54. Modelo AIDA

Tanto en el embudo de ventas y el modelo AIDA son herramientas importantes en el proceso de marketing y ventas. Al utilizar estas dos herramientas juntas, se puede maximizar el impacto de una campaña de marketing al asegurarse de la combinación de ambas herramientas es esencial para el éxito de cualquier campaña de marketing

6.7.2 Atracción de público a la página web

Dentro de la metodología empleada en esta parte de la tesis se basa en el modelo AIDA, el cual se enfoca en las etapas de atracción, interés, deseo y acción del cliente potencial. Para la fase de atracción, se propone generar un posicionamiento en Instagram a través de la creación de un hashtag propio y relevante que tenga correlación con otros hashtags populares dentro del nicho de mercado. Además, se recomienda utilizar los reels, que es el formato de contenido más visto

en la plataforma, para aumentar el alcance de las visitas y lograr captar la atención del público objetivo.

En cuanto a la creación del hashtag (en este caso #isataracena), se sugiere realizar una investigación previa sobre los términos más buscados por la audiencia para determinar cuál puede ser el más adecuado para el negocio. Una vez seleccionado, se recomienda utilizarlo de manera constante en todas las publicaciones, así como animar a los seguidores y clientes a utilizarlo en sus propias publicaciones relacionadas con la marca.

Por otro lado, los reels son una herramienta muy útil para la generación de contenido atractivo y entretenido. Se sugiere utilizarlos para presentar productos o servicios de manera original y creativa, aprovechando la capacidad de la plataforma para incluir música y efectos visuales. Asimismo, es importante que los reels estén alineados con la estrategia de contenido general de la marca y que sean coherentes con la imagen que se desea proyectar.

En resumen, la metodología propuesta se basa en la utilización del modelo AIDA para guiar la estrategia de marketing en Instagram, centrándose en la fase de atracción mediante la creación de un hashtag propio (#isataracena) y relevante, y la utilización de los reels para generar contenido atractivo y aumentar el alcance de las publicaciones.

6.7.3 Retención de atención al contenido

Este punto de la metodología se enfoca en la retención de la atención del usuario al contenido, una tarea importante dentro de la estrategia de marketing digital. Para lograr esto, se llevaron a cabo varias acciones. (Figura 55) En primer lugar, se reestructuró el apartado más visitado de la página web, con el objetivo de enriquecerlo con más contenidos y hacerlo más atractivo para los usuarios.

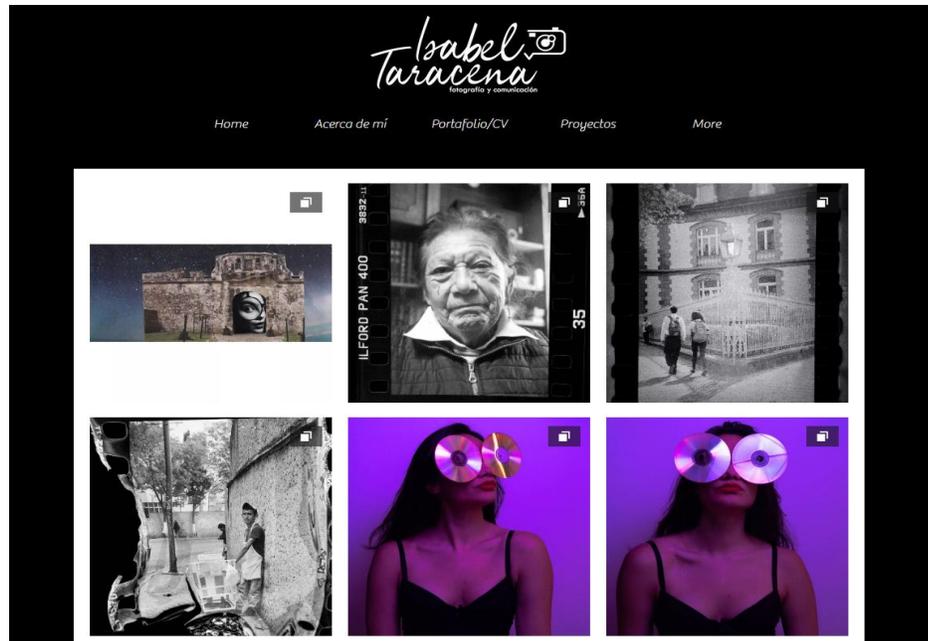


Figura 55. Sección de la página más visitada en el análisis.

La autora se encargó de la creación del contenido, que fue estructurado de tal manera que los usuarios puedan navegar por él de manera intuitiva. Para esto, se diseñó un menú de navegación del contenido a lado derecho de la página web, lo que facilita la búsqueda de contenido que le interesa al usuario.



Figura 56. Reestructuración del contenido de la página web.

Además, se incluyó una galería digital en la página web, en la que cada imagen está acompañada por una redacción de los pensamientos que tiene la autora al respecto. Esto permite al usuario conectar de una manera más profunda con el contenido, lo que aumenta la retención de su atención (ver Figura 56 y Figura 57).

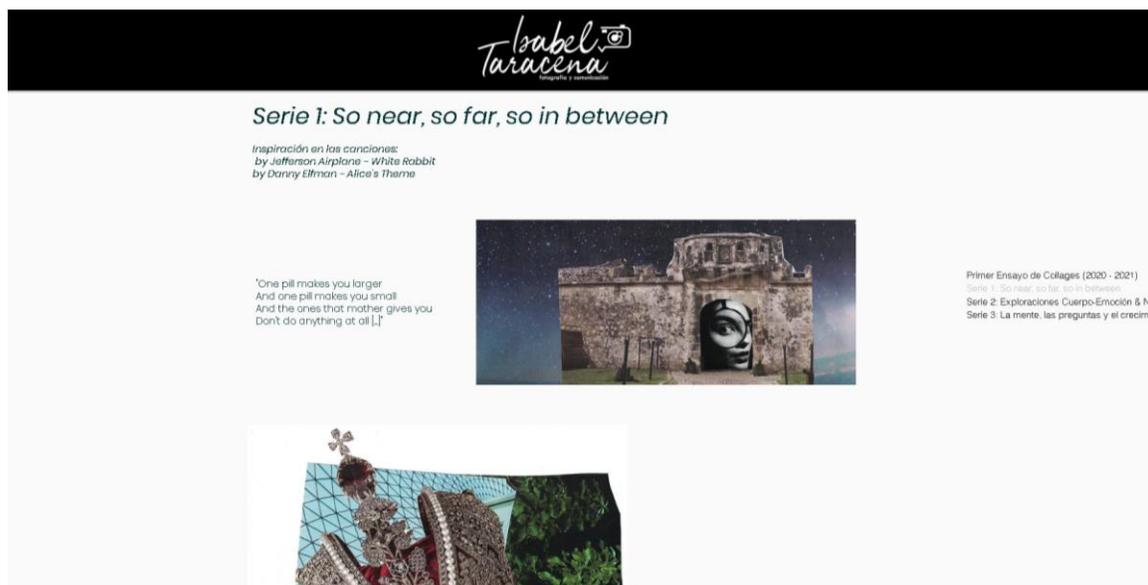


Figura 57. Segunda parte de la reestructuración de la página web.

Otra de las estrategias utilizadas para retener la atención del usuario fue la creación de contenido multimedia. Se incluyeron imágenes y videos en el contenido de la página web, con el objetivo de hacerlo más visualmente atractivo y lograr que los usuarios se involucren más en el contenido.

Para concluir, el objetivo de este subtema fue lograr la retención de la atención del usuario al contenido, para lo cual se llevaron a cabo varias acciones, como la reestructuración del apartado más visitado de la página web, la inclusión de una galería digital y la creación de contenido multimedia. Todo esto con el fin de lograr una experiencia más atractiva e interactiva para el usuario.

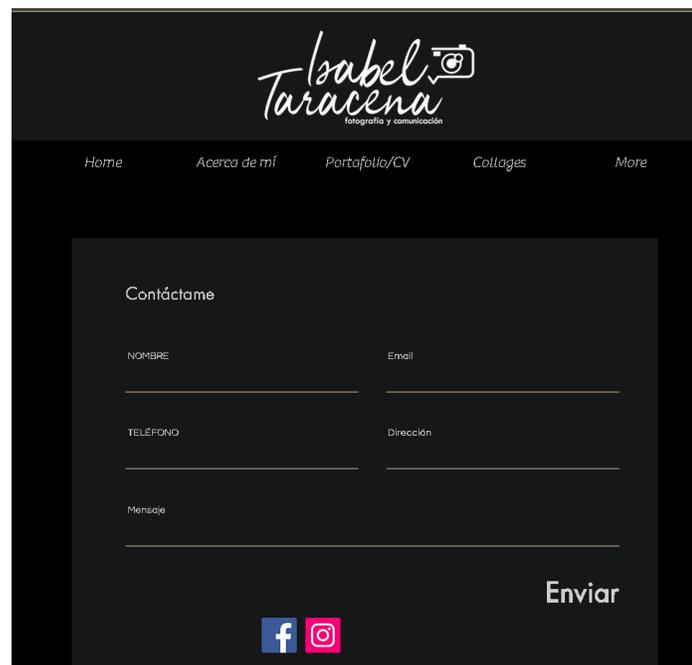
6.7.4 Llamado a la acción

La última parte de la metodología se enfoca en el llamado a la acción, que es fundamental para convertir las visitas en ventas. En este caso, el objetivo es lograr que los usuarios se pongan en contacto con la empresa para solicitar más información o hacer una compra en caso de que se

trate de una ecommerce. Para lograrlo, se optó por crear un anuncio en forma de reel en Instagram que muestre de manera clara y concisa los servicios que se ofrecen y que invite al usuario a tomar acción.

El anuncio en reel es una herramienta muy efectiva para dar a conocer los servicios de la empresa, ya que este tipo de contenido es muy popular en Instagram y genera un alto nivel de engagement. El video debe ser breve y conciso, mostrando de manera creativa los beneficios que ofrece la empresa y llamando a la acción al final del mismo.

Además, se diseñó una página de contacto en la web (Figura 58) que sea fácil de usar y permita al usuario enviar sus consultas o solicitar información adicional de manera rápida y sencilla. Para lograrlo, se incluyó un formulario de contacto y un chat en línea que permita a los usuarios interactuar en tiempo real con un agente de la empresa.



The image shows a dark-themed contact page for 'Isabel Taracena' (fotografía y comunicación). At the top, there is a navigation menu with links: Home, Acerca de mí, Portafolio/CV, Collages, and More. Below the menu is a contact form titled 'Contáctame'. The form contains five input fields: 'NOMBRE', 'Email', 'TELÉFONO', 'Dirección', and 'Mensaje'. At the bottom right of the form is a button labeled 'Enviar'. Below the form are icons for Facebook and Instagram.

Figura 58 Apartado de contacto en la pagina web

Para la metodología propuesta se enfoca en el modelo AIDA, donde se busca atraer la atención del usuario con un posicionamiento en Instagram y contenido atractivo en la página web, retener su atención con una galería digital y un menú de navegación del contenido, y finalmente llamar a la acción con un anuncio en forma de reel en Instagram y una página de contacto fácil de usar.

Con esta estrategia, se espera incrementar el tráfico a la página web y convertir las visitas en ventas efectivas.

En conclusión, la metodología utilizada para el desarrollo de las estrategias de marketing se basó en el modelo AIDA (consistió en los capítulos 6.7.2 al 6.7.4; Atracción del público objetivo, retención de atención y llamado a la acción). En la fase de atracción, se generó un posicionamiento en Instagram mediante la creación de un hashtag propio y su vinculación con los hashtags más populares, así como la implementación de reels para incrementar el alcance de las visitas. Para la retención de la atención del usuario al contenido, se reestructuró el apartado más visitado de la página web, enriqueciéndolo con más contenidos y diseñando un menú de navegación del contenido a la derecha de la página web. Además, se incluyó una galería digital en la página web, en la cual cada imagen está acompañada por una redacción de los pensamientos de la autora al respecto.

Finalmente, para el llamado a la acción, se optó por un anuncio en forma de reel para dar a conocer los servicios y establecer un contacto entre la empresa y el cliente. Se sugiere evaluar la efectividad de esta estrategia y, en caso de ser necesario, realizar ajustes para mejorar su impacto en la conversión de ventas.

6.8 Resumen y comprobación de hipótesis

Hipótesis Central

Con la implementación de técnicas de machine learning y estrategias de marketing en redes sociales y sitio web se obtendrá un impacto positivo en el aumento del tráfico del sitio web <https://www.isabeltaracena.com/> y mejorará el posicionamiento de la microempresa de fotografía (Estudio Isabel Taracena) en el mercado objetivo.

Comprobación

Tráfico del Sitio Web: Se utiliza herramientas como Google Analytics para medir el tráfico actual del sitio web. Identifica métricas clave como el número de visitantes, páginas vistas, y el tiempo promedio de permanencia. (Figura 12. Cantidad de visitantes por sección de la página web)

Posicionamiento en Redes Sociales: La presencia actual en redes sociales utilizando métricas como seguidores, interacciones, y alcance de publicaciones. (Tabla 2. Cálculo del engagement de las 8 publicaciones más populares en Instagram de Isabel Taracena.)

Aportación de las técnicas de machine learning: A través del análisis comparativo de métricas generadas por diversos modelos de Machine Learning (ver Figura 39. Resumen gráfico de las variables generadas por los modelos y sus promedios.), se evidencia que, aunque las variaciones son mínimas, los resultados convergen hacia el mismo patrón: un mayor atractivo hacia el contenido específico a la galería de fotografías.

Incremento de visitas en la página web: Este hallazgo se corrobora al examinar las secciones más visitadas, como se ilustra a través de una serie de tiempo (Figura 36. Gráfica de series de tiempo en intervalos de semanas sobre usuarios adquiridos en 2022.). Así, la estrategia de vincular la página web con las publicaciones en redes sociales emerge como un factor clave, potencialmente generando picos de actividad en la página web que se reflejan de manera palpable en el análisis.

La comprobación de la hipótesis central revela un incremento considerable de usuarios para el Estudio Isabel Taracena. Las estrategias de machine learning y marketing generaron un aumento significativo en el tráfico del sitio web, evidenciado por métricas clave en Google Analytics. El análisis del posicionamiento en redes sociales respaldó ese incremento de visitas, destacando un aumento en seguidores, interacciones y alcance de publicaciones.

En la implementación de estas técnicas demostró un impacto positivo, validando la hipótesis central al mejorar el tráfico del sitio web y consolidar el posicionamiento del Estudio Isabel Taracena en su mercado objetivo definido en sus redes sociales.

CAPÍTULO VII: CONCLUSIONES Y RECOMENDACIONES

En la metodología presentada en esta tesis aborda varios temas clave en el análisis de datos en el ámbito del marketing digital. La recolección y análisis de datos se ha convertido en una disciplina fundamental para la toma de decisiones basadas en hechos en lugar de suposiciones. El estudio se enfoca en diferentes técnicas y metodologías de análisis de datos, como el uso de métricas, geomarketing, selección automática de imágenes, modelos de machine learning y análisis de datos de Google Analytics a través de series de tiempo.

La investigación en estos temas permitirá a la microempresa a mejorar su presencia en línea y aumentar audiencia a través de las estrategias comerciales y una mayor eficiencia en la toma de decisiones. En resumen, la metodología presentada en esta tesis ofrece una valiosa contribución al conocimiento de la disciplina del marketing digital y tiene el potencial de ser utilizada por empresas que buscan mejorar su posicionamiento.

7.1 Conclusiones generales de los resultados

Se decidió implementar una campaña de posicionamiento en paralelo con el análisis de datos. Para solucionar este problema sobre la falta de datos por analizar, se optó por una estrategia SEO backlinks a través de las redes sociales de la empresa, lo que permitió redirigir a los usuarios a la página web y realizar un seguimiento del alcance de las publicaciones. Esta estrategia permitió canalizar visitas a la página web y mejorar la cantidad y calidad de los datos obtenidos para el análisis.

Utilizar Instagram como principal canalizador de audiencia y medio de difusión para la página web (En la página 127, Figura 36, En la serie de tiempo los picos de usuarios en la página web cada que se realizaba una publicación en Instagram, por lo se logró el objetivo de aumentar las visitas y visibilidad de la página web), ya que es la red social más apropiada para el contenido generado y es el medio donde más actividad se genera.

Se concluyó que el cambio de privacidad a una cuenta asignada como creadora de contenido y la vinculación de un apartado de galería de Instagram en la página web incrementó significativamente el alcance de las publicaciones y el engagement con el contenido. Además, se logró mejorar la visualización de la página web a través de la opinión positiva de los usuarios sobre el contenido y la optimización del sitio.

7.2 Recomendaciones generales

Con base a las conclusiones obtenidas, se recomienda a Isabel Taracena seguir trabajando en la generación de contenido de calidad para Instagram y la página web, así como también continuar con la interacción con los seguidores y la implementación de estrategias de marketing digital. Además, se sugiere explorar otras redes sociales para complementar el posicionamiento de la página web (Facebook e Pinterest) y evaluar su impacto en el alcance de la audiencia, el

engagement y la reputación. También se sugiere considerar la utilización de herramientas de análisis de datos para medir y evaluar el impacto de las estrategias implementadas en la tesis.

7.3 Conclusión sobre el proceso realizado

El uso de algoritmos de Machine Learning puede mejorar significativamente la efectividad de las campañas publicitarias en redes sociales de la microempresa. Esto se debe a que el análisis de datos con técnicas de Machine Learning permite una mejor comprensión del comportamiento del consumidor, lo que puede llevar a una mayor segmentación y personalización de los anuncios, lo que aumenta la probabilidad de conversión. Además, el análisis de datos también permite la identificación de las mejores horas para publicar, lo que puede llevar a una reducción de costos a campañas de marketing más efectivas.

Las herramientas de Machine Learning más efectivas para analizar el comportamiento de los consumidores en línea y obtener información relevante para la microempresa son aquellas que permiten la segmentación y el análisis de los datos. Las herramientas de análisis de datos en tiempo real pueden ser muy útiles para obtener información en tiempo real sobre los patrones de comportamiento de los consumidores. Además, las herramientas de segmentación permiten la creación de perfiles más precisos de los clientes, lo que a su vez permite la personalización de los anuncios.

Las estrategias específicas de marketing digital que pueden usarse para aumentar el alcance y la visibilidad de la microempresa en su mercado objetivo incluyen el SEO, la publicidad en redes sociales y la creación de contenido de calidad. Estas estrategias pueden mejorarse aún más con técnicas de Machine Learning, que permiten una mejor segmentación y personalización de los anuncios, así como la identificación de las palabras clave y temas más relevantes para un público objetivo.

Los datos más relevantes que deben ser recopilados para mejorar el rendimiento de la publicidad en línea de la microempresa incluyen el comportamiento del usuario, el rendimiento de los anuncios, las conversiones y la información demográfica del usuario. Estos pueden ser analizados y utilizados eficazmente en un plan de marketing.

Los factores clave que influyen en la experiencia de usuario en línea son la velocidad de carga del sitio web, la facilidad de navegación y la relevancia del contenido. El uso de técnicas de Machine Learning puede ayudar a priorizar elementos para mejorar la experiencia de usuario en

línea a través de la personalización y la segmentación de los anuncios, lo que aumenta la relevancia del contenido.

Los canales de publicidad en línea más efectivos para alcanzar a la audiencia objetivo de la microempresa son aquellos que se utilizan con mayor frecuencia por los clientes potenciales. Las herramientas de Machine Learning son comúnmente utilizadas para determinar la eficacia de los diferentes canales de publicidad y para identificar los canales que generan el mayor retorno de inversión (en este caso en particular resulta más eficiente realizar publicidad pagada por redes sociales que por el buscador).

Los tipos específicos de contenido más efectivos para atraer a un mayor número de visitantes y aumentar el tráfico en línea de la microempresa son aquellos que son relevantes, informativos y atractivos. El análisis de datos con técnicas de Machine Learning puede ser muy útil para determinar qué tipo de contenido es más popular entre la audiencia objetivo (tipo de fotografías y constante interacciones con el público).

Identificar los factores clave que influyen en el éxito de la publicidad en línea de la microempresa en su mercado objetivo. Para ello, se pueden utilizar técnicas de análisis de datos con Machine Learning, como la regresión y la clasificación, para identificar los factores más relevantes y optimizar la publicidad en línea de la microempresa.

CAPÍTULO VIII: PROPUESTA

Basándonos en los resultados y análisis detallados a lo largo de esta investigación, se pueden considerar diversas acciones con el fin de desarrollar un plan de marketing en línea efectivo y coherente. En primer lugar, establecer un calendario de publicaciones por parte de Isabel Taracena, que asegure una presencia constante en las plataformas de redes sociales. Como se ha identificado en el capítulo 6.5.1 Análisis por series de tiempo (Adquisición de usuarios y tiempo de interacción) donde los picos de más actividad en la página son cuando se realiza una publicación en redes sociales, además de considerar palabras claves apropiadas entre palabras generales y específicas (de lo general a lo particular como por ejemplo pasar de fotografía a paisajismo urbano). Posicionando un hashtag propio (#isataracena) para distinguirse como marca y generar mayor alcance e interacción en las redes sociales. Mantener una presencia activa es esencial para aumentar la interacción y el alcance. Por lo que es crucial priorizar la calidad del contenido compartido.

Adicionalmente, se recomienda asignar un presupuesto mensual destinado al público objetivo establecido en la investigación. Esta estrategia se alinea con el análisis presentado en el capítulo 6.2 Exploración y estudio de mercado aplicado por geomarketing, el cual focaliza las zonas con mayor posibilidad en el mercado de ciudad de México, puede generar mayores oportunidades de conversión. Este presupuesto puede ser utilizado para promocionar publicaciones, llevar a cabo campañas publicitarias y emprender otras acciones de marketing que permitan alcanzar una audiencia más amplia. Cabe destacar que esta estrategia está dirigida principalmente a clientes interesados en la fotografía artística y servicios de eventos, como se detalla en el capítulo 5.3 Muestra.

En conjunto, las acciones antes mencionadas se traducirán en un plan de marketing en línea efectivo. Este plan está diseñado específicamente para mejorar la presencia en línea de la microempresa, incrementar el engagement de los clientes y generar oportunidades de conversión. Las estrategias propuestas son respaldadas por los hallazgos de esta investigación y están vinculadas con el problema y el objetivo establecidos desde el inicio de este estudio.

8.1 Propuesta por objetivos específicos

1. Evaluar el impacto del uso de algoritmos de Machine Learning en las campañas publicitarias de redes sociales de "Estudio Isabel Taracena" para aumentar el número de conversiones y reducir los costos asociados, mejorando el retorno de la inversión (ROI).

En la clasificación automática de fotografías (6.3 Selección automática de imágenes y correlación de engagement) resultó tener no tener una correlación tan directa al momento de ser implementada, sin embargo al segmentarlas por grupos (presentado en la Figura 33. Grafica de segmentación por tipo de publicación. Del capítulo 6.4 Evaluación de diferentes modelos de machine learning en predicción del engagement) se puede destacar que las publicaciones tipo reel aun sin ser patrocinadas y se realizaron pocas a comparación de otro tipo de contenido publicado en redes sociales tienen un alcance mayor, por lo que invertir en este tipo de publicación sería la opción más rentable al invertir en publicidad en redes, focalizando en resaltar la sección de proyectos y portafolio, ya que representan áreas de alto interés según las métricas de Google buscando el aumento de visualizaciones y, por ende, en la mejora del retorno de la inversión.

2. Identificar las herramientas de Machine Learning más adecuadas para analizar el comportamiento de los consumidores en línea y obtener información relevante que permita optimizar las estrategias de marketing digital de "Estudio Isabel Taracena".

Durante la evaluación de los modelos de machine learning (Figura 32. Revisión a detalle de las diferencias entre modelos.) Se identifica que el modelo Tree es el que mejor se adapta a la dispersión de datos referente a la medición de engagement respecto al tipo de publicación.

3. Diseñar estrategias específicas de marketing digital basadas en técnicas de Machine Learning para aumentar el alcance y la visibilidad de "Estudio Isabel Taracena" en su mercado objetivo, maximizando el impacto y la interacción con el público.

Las estrategias específicas de marketing digital planteadas (6.7 Estrategias consideradas e implementaciones realizadas en la campaña de marketing), están planteadas para tener el máximo alcance y visibilidad del trabajo que realiza Isabel Taracena y enfocar las secciones de proyectos y portafolio en redes sociales ya que es la mejor estrategia que puede tener por el giro empresarial en el participa.

4. Analizar cómo la implementación de técnicas de Machine Learning influye en la retención de usuarios y el engagement en línea de "Estudio Isabel Taracena", identificando los factores clave que contribuyen a una mayor interacción con los clientes.

Para el análisis de retención de usuarios (6.5 Análisis de la página web con los datos de google analytics) se recolectaron datos a través de Google analytics y se observaron las diferentes secciones que tiene la página web (Figura 41. Histograma de las secciones más visitadas en la página web.). Por lo que se mide retención de usuarios en el contenido, especialmente aquellos interesados en la sección de proyectos y portafolio. Identificar los factores clave, como la correlación de engagement, que contribuyen a una mayor interacción con estos clientes.

5. Mejorar la experiencia del usuario en línea de "Estudio Isabel Taracena" mediante la aplicación de técnicas de Machine Learning, optimizando la navegación, el diseño y el contenido para aumentar la satisfacción y el tiempo de permanencia en la página web y redes sociales.

Para tratar este objetivo, se trabajó sobre la experiencia del usuario en línea focalizando en la navegación y diseño de las secciones de proyectos y portafolio, (Figura 56. Reestructuración

del contenido de la página web.) Por lo que se realizaron cambios para obtener una interfaz optimizada en la presentación de contenido y aumentar la satisfacción y tiempo de permanencia del usuario.

6. Identificar qué tipos específicos de contenido generan un mayor tráfico y engagement en línea para "Estudio Isabel Taracena" mediante el análisis de datos con técnicas de Machine Learning, con el fin de optimizar la estrategia de contenidos.

Para la identificación de publicaciones a través de técnicas de Machine Learning, (Tabla 2. Calculo del engagement de las 8 publicaciones mas populares en instagram de Isabel Taracena.) Se clasificaron tipos específicos de contenido generan un mayor tráfico y engagement en las secciones de proyectos y portafolio. Utilizar esta información para optimizar la estrategia de contenidos, por el contenido más popular que son autorretratos en formato de reel.

7. Definir de manera más objetiva y precisa al cliente objetivo de "Estudio Isabel Taracena" a través de técnicas de Machine Learning, segmentando y analizando datos demográficos y patrones de comportamiento, para adaptar estrategias de marketing de manera efectiva.

Definir al cliente objetivo de "Estudio Isabel Taracena" (6.2 Exploración y estudio de mercado aplicado por geomarketing), considerando datos demográficos y patrones de comportamiento específicos de Google analytics, donde se recomienda usar etiquetas o hashtags de las locaciones para tener un mejor posicionamiento cuando los usuarios busquen contenido relacionado del lugar.

8. Identificar las herramientas de Machine Learning adecuadas para aplicar en ventas mediante el análisis de las principales herramientas open source.

Para la identificación de herramientas de Machine Learning (5.5 Instrumentos de análisis de datos) se evaluaron varias plataformas open source, sin embargo Orange destaca por la facilidad de ingresar, limpiar y procesar datos. Además de que tiene modelos de machine learning ya implementados, por lo que fue la herramienta principal utilizada en este trabajo.

9. Evaluar las plataformas que permiten crear una página web con funciones de marketing digital y la posibilidad de aplicar técnicas de Machine Learning como administrador.

Para la plataforma que la página web el estudio ya contaba con la página en GoDaddy. Por lo que vincular esa herramienta con Google analytics fue accesible, y dentro del presupuesto de la empresa, el pago de la plataforma ya incluye el dominio, el hosting y el uso de esas herramientas de análisis para marketing digital.

10. Determinar las métricas actuales de la página web seleccionada y calcular su media de visitas antes de aplicar las técnicas de Machine Learning.

Las mediciones de la página web (6.1 Métricas de la página web (Estudio fotográfico Isabel Taracena) durante la investigación) se recolectaron a través de google analytics y se concluyó que eran muy pocos datos (una media de 3 visitas) para hacer un análisis basado solamente en la página web, por lo que también se realizó un análisis de redes sociales para impulsar las visualizaciones en la página web donde el análisis de datos fue más relevante por la interacción que tiene la plataforma con sus usuarios.

11. Seleccionar y aplicar las técnicas de Machine Learning más apropiadas para analizar las métricas y el historial de visitas de la página web seleccionada.

En el tema (6.5 Análisis de la página web con los datos de google analytics) donde el método de análisis que se empleo fue una serie de tiempo (Figura 36. Grafica se series de tiempo en intervalos de semanas sobre usuarios adquiridos en 2022.) Donde se refleja los picos de interacciones en la página web y retención de los usuarios en la página web.

12. Comparar las métricas y el historial de visitas con las métricas obtenidas mediante las técnicas de Machine Learning y evaluar los resultados para comprobar la hipótesis planteada.

Dentro del análisis comparativo (Figura 39. Resumen gráfico de las variables generadas por los modelos y sus promedios.) Se observan las métricas comparadas con diferentes modelos de Machine Learning. Aunque las variaciones sean mínimas al momento de evaluar los resultados, señalan los mismos resultados al contenido más atractivo de la página web (Figura 41. Histograma de las secciones más visitadas en la página web.). Por lo que la estrategia de vincular la página web con las publicaciones en redes sociales se pudo tener picos de actividad en la página web, donde se ven reflejados dentro del análisis.

8.2 Explicación

Se presentan explicaciones a las conclusiones generales y recomendaciones derivadas de la investigación realizada en la tesis. La relevancia de las herramientas digitales de marketing y del análisis acompañado con un estudio profundo de los datos con técnicas de Machine Learning para una mejor segmentación y personalización de los anuncios, aumentando la probabilidad de conversión. Además, el análisis de datos permite la identificación de las mejores horas para publicar, lo que puede llevar a una reducción de costos y a campañas de marketing más efectivas.

En este punto se destaca la importancia principal de Instagram como principal canalizador de audiencia y medio de difusión para la página web. La investigación realizada permitió concluir que el cambio de privacidad a una cuenta asignada como creadora de contenido y la vinculación de un apartado de galería de Instagram en la página web incrementó significativamente el alcance de las publicaciones y el engagement con el contenido. Además, se logró mejorar la reputación de la página web a través de la opinión positiva de los usuarios sobre el contenido y la optimización del sitio.

En cuanto a las recomendaciones, se sugiere seguir trabajando en la generación de contenido de calidad para Instagram y la página web, así como también continuar con la interacción con los seguidores y la implementación de estrategias de marketing digital. Además, se sugiere explorar otras redes sociales para complementar el posicionamiento de la página web y evaluar su impacto en el alcance de la audiencia, el engagement y la reputación. También se sugiere considerar la utilización de herramientas de análisis de datos para medir y evaluar el impacto de las estrategias implementadas en la tesis.

En resumen, las conclusiones y recomendaciones presentadas resaltan la importancia de la metodología propuesta, así como la relevancia de Instagram como canalizador de audiencia y medio de difusión para la página web. Además, se destaca la importancia en la efectividad de las campañas publicitarias en redes sociales de la microempresa.

Las recomendaciones propuestas buscan continuar mejorando la generación de contenido de calidad y la implementación de estrategias de marketing digital, así como explorar otras redes sociales y considerar la utilización de herramientas de análisis de datos para medir y evaluar el impacto de las estrategias implementadas.

8.3 Fechas

En esta sección, se presenta el cronograma detallado sobre la realización de este trabajo. El cronograma se divide en varias etapas y actividades que abarcan desde la investigación inicial hasta la conclusión del proyecto.

Preparativos

- Recolección de datos en google analytics: 1 año (2022)
- Recopilación de información para el marco teórico: (2021-2023)
- Recolección de datos en Instagram: 1 año (2022)
- Fase de atracción en Instagram: 1 mes (Febrero 2023)
- Mejora del apartado más visitado de la página web: 2 semanas (Agosto 2022)
- Diseño y publicación de anuncio en forma de reel: 1 semana (Febrero 2023)
- Implementación de la campaña de marketing: 1 mes (Abril 2023)

Desarrollo del contenido de la tesis: (2021-2023)

Introducción (Agosto 2021)

- Presentación del tema
- Justificación de la investigación
- Objetivos de la tesis
- Metodología utilizada
- Marco teórico

Revisión de literatura relevante (Agosto 2021- Enero 2023)

- Fundamentos teóricos relacionados con el tema de investigación
- Recopilación de técnicas de machine learning

Generación de posicionamiento en Instagram (Noviembre 2022)

- Creación de un hashtag propio relacionado con hashtags populares
- Utilización de Reels para aumentar el alcance de las visitas
- Retención de la atención del usuario al contenido

Reestructuración del apartado más visitado de la página web (Mayo 2022)

- Enriquecimiento del contenido existente
- Creación de una galería digital con redacciones relacionadas con las imágenes

Recapitulación de los hallazgos y resultados obtenidos (Octubre 2022 - Marzo 2023)

- Evaluación de la efectividad de las estrategias de marketing utilizadas
- Implicaciones y recomendaciones para futuras investigaciones o aplicaciones prácticas
- Elaboración de conclusiones

Últimos cambios y revisión de tesis (Agosto-Diciembre 2023)

8.4 Presupuesto

El presupuesto destinado para la implementación de las estrategias de marketing, se realizó un pago de \$1300 pesos mexicanos para adquirir el dominio web en GoDaddy. Este pago incluyó el dominio, hospedaje y vinculación con herramientas de análisis digital como Google Analytics, además de un abono de \$300 pesos para realizar algunas pruebas de publicaciones patrocinadas. Cabe destacar que los demás gastos de producción y contenido corrieron por cuenta de la propietaria de la página web.

Es importante señalar que, aunque se realizaron algunos gastos adicionales, estos fueron asumidos de manera independiente y no se contabilizan en el presupuesto mencionado anteriormente.

8.5 Cierre de la propuesta

Para cerrar la propuesta presentada, se busca mejorar la implementación de un plan de marketing estructurado y la optimización del contenido a publicar en redes sociales. Al establecer un calendario de publicaciones y diseñar contenido adecuado para cada publicación, se puede aumentar la interacción y el engagement con el público objetivo. Asimismo, al designar un presupuesto mensual y seguir vinculando las publicaciones con la página web de la empresa, se puede consolidar la presencia en línea y mejorar la conversión de visitantes en clientes potenciales.

Es importante destacar la relevancia de la investigación de mercado previa y el análisis de los datos recopilados para identificar el público objetivo y las tendencias de comportamiento de los consumidores en las redes sociales. A partir de esta información, se pueden tomar decisiones

informadas y adaptar la estrategia de marketing de la empresa a las necesidades y preferencias del público.

En definitiva, al seguir estas recomendaciones se espera obtener mejores resultados en términos de interacción y conversión, lo que se traducirá en un aumento de la visibilidad y el reconocimiento de la marca en línea, así como en un incremento en las ventas y el crecimiento de la empresa en general.

8.6 Reflexiones finales

Durante el desarrollo de esta tesis, se tuvo la oportunidad de trabajar en tres áreas principales que son la base de este trabajo: estadística, marketing y machine learning. Aunque se pudo estudiar sus fundamentos, se descubrió que al trabajar con ellas se encontraban múltiples variables que afectan los resultados, desde los datos de entrada relevantes hasta la implementación de modelos de aprendizaje automático con sus diversas funciones y aplicaciones en las estrategias de marketing.

El uso de estas herramientas en el mundo real es un constante trabajo de pruebas en campo, ya que los resultados pueden ser muy fluctuantes. Lo que funciona hoy puede no funcionar mañana, por lo que solo se puede hacer especulaciones al respecto.

A lo largo de este trabajo, se ha aprendido que la combinación de estas áreas es esencial para la toma de decisiones informadas y eficaces. La interacción entre ellas se visualizan mejores resultados, sino también desarrollar nuevas soluciones y enfoques para los próximos desafíos sociales, económicos y tecnológicos.

BIBLIOGRAFÍA

- Ana Prieto, M. M. (2004). Sistema de información en las organizaciones: Una alternativa para mejorar la productividad gerencial en las pequeñas y medianas empresas. *Revista de ciencias sociales (RCS) No. 2*, 16.
- Antonio J. Serrano, E. S. (2009-2010). *Redes Neuronales Artificiales*. Valencia: Open Course Ware.
- Asociación de internet. (2020). *Estudio sobre las implicaciones jurídicas del uso de cómputo en la nube por el sector público mexicano: Recopilación de estándares y buenas prácticas de fuente internacional*. CDMX.
- AWS, A. (27 de 11 de 2021). *Amazon web services*. Obtenido de Amazon web services: https://aws.amazon.com/es/what-is-aws/?nc2=h_ql_le_int
- Azure, M. (28 de 11 de 2021). *Azure*. Obtenido de Azure: https://azure.microsoft.com/es-mx/free/search/?&ef_id=Cj0KCQiA15yNBhDTARIsAGnwe0UrmgMS-1UU-RbYo6KgmqR9G8hu1MGaTu-GTxo4dxImxSNRREdyUN4aAm5zEALw_wcB:G:s&OCID=AID2200215_SEM_Cj0KCQiA15yNBhDTARIsAGnwe0UrmgMS-1UU-RbYo6KgmqR9G8hu1MGaTu-GTxo4dxImxSNRREdyUN4aAm5z
- Barnes, J. (2015). *Azure Machine Learning Microsoft Azure Essentials*. Microsoft Press.
- Barviera, T. (2017). Técnicas para el análisis del sentimiento en Twitter: Aprendizaje Automático Supervisado y SentiStrength. *Dígitos*, 33-50.
- BI, M. P. (28 de 11 de 2021). *Power BI*. Obtenido de Power BI: <https://powerbi.microsoft.com/es-es/getting-started-with-power-bi/>
- Calancha, Z. (20 de 10 de 2021). *Niefar*. Obtenido de Niefar: <https://niefcz.files.wordpress.com/2011/07/breve-aproximacion-a-la-tecnica-de-arbol-de-decisiones.pdf>

Catalunya vanguardista. (28 de 12 de 2022). *Catalunya vanguardista*. Obtenido de Un algoritmo predice el comportamiento de los usuarios digitales: <https://www.catalunyavanguardista.com/un-algoritmo-predice-el-comportamiento-de-los-usuarios-digitales/>

Chacon, R. F. (2017). *Bosque aleatorios como extensión de los árboles de clasificación con los programas R y Python*. Lima: Instituto Nacional de Estadística e informática.

Chicaiza, A. G. (2020). Technologies in artificial intelligence for Marketing: a review of the literature. *Pro Sciences*, 36-47.

Cuadrado, D. A. (2019). *Utilización del machine learning en la industria 4.0*. Valladolid: Escuela de ingenierías industriales.

Cutler Adele, C. D. (2012). En Ensemble machine learning. *Springer*, 157-175.

Daniel Santos, L. D. (2020). Algoritmos de rastreo de movimiento utilizando técnicas de inteligencia artificial y machine learning. *Información tecnológica*, 23-38.

Edwards Antony, L. L. (1995). A method for cluster analysis. *Biometrics*, 365-375.

Elías Adolfo Albornoz Del Valle, F. J. (2020). Geomarketing: Desde una visión comercial a una aplicación social en contextos metropolitanos. *Revista de Geografía Norte Grande*, 143-167.

Fanjie Kong, T. F. (2023). *Neural Insights of digital Marketin content design*. Long beach California: cs.LG.

Franco, S. M. (2020). *Análisis y propuesta para el aprovechamiento de herramientas de machine learning en el sector del marketing digiral en Netbangers*. Bogotá: FCEA.

- Galeano, S. (28 de Enero de 2022). *Marketing 4 ecommerce*. Obtenido de Marketing 4 ecommerce: <https://marketing4ecommerce.net/cuales-redes-sociales-con-mas-usuarios-mundo-ranking/>
- García, N. F. (2017). Fake news: una oportunidad para la alfabetización mediática. *Nueva Sociedad*, 66-77.
- George Box. (2011). *Bayesian inference in statistical analysis*. John Wiley & Sons.
- Géron, A. (2017). *Hands-on machine learning with scikit-learn & tensorflow: Concepts, tools, and techniques to build intelligent systems*. California: O'Reilly Media.
- Huber, H. L. (2022). *How causal machine learning can leverage marketing strategies*. Fribourg: University of Fribourg. Dept. of Economics.
- Ibarra, F. (04 de 11 de 2021). *Twitter*. Obtenido de Twitter: <https://twitter.com/faustoibarra?lang=es>
- IBM. (20 de 10 de 2021). *IBM documentación*. Obtenido de <https://www.ibm.com/docs/es/spss-modeler/SaaS?topic=networks-neural-model>
- INEGI. (2021). *En México hay 84.1 millones de usuarios de internet y 88.2 millones de usuarios de teléfonos celulares: ENDUTIH 2020*. CDMX.
- Inegi. (06 de 10 de 2021). *Inegi*. Obtenido de Inegi: <https://www.inegi.org.mx/temas/vabcoel/>
- Jiang, L. (2020). *Support Vector Machine (SVM) Machine Strategy Analysis Method Based on Time Series*. Anhui china: The Electrochemical Society.
- Jonathan Benchimol, S. K. (2021). *Text mining methodologies with R: an application to central Bank Texts*. Richmond, VA-EU: scholar harvard.

- Karina Bricio Samaniego, J. C. (2018). *El marketing digital como herramienta en el desempeño laboral en el entorno ecuatoriano: estudio de caso de los egresados de la universidad de Guayaquil*. Guayaquil Ecuador: Revista científica de la universidad de cienfuegos.
- Kaushik, A. (2011). *Analítica web 2.0 El arte de analizar resultados y la ciencia de centrarse en el cliente*. Barcelona: Gestión 2000.
- Laguna, C. (2020). *Correlación y regresión lineal*. Instituto Aragónes de ciencias de la salud, diplomado en salud pública.
- Larranaga, P. I. (1997). Tema 8. redes neuronales. *Redes Neuronales. U. del P. Vasco*, 12-17.
- López, V. H. (2 de Septiembre de 2021). *Retailers.mx*. Obtenido de <https://retailers.mx/google-mexico-retail-crecera-40-en-los-proximos-cinco-anos/>
- Ma. Jiménez, V. G. (2000). La predicción del rendimiento académico: regresion lineal versus regresion logística. *Psicothema*, 248-525.
- Maisueche Cuadrado, A. (2019). *Utilización del Machine Learning en la industria 4.0*. Valladolid: Universidad de valladolid, Escuela de ingenierías industriales .
- Mathworks. (20 de 10 de 2021). *Mathworks*. Obtenido de Mathworks: <https://la.mathworks.com/discovery/linear-regression.html>
- Matich, D. J. (2001). *Redes Neuronales: Conceptos Básicos y Aplicaciones*. Regional Rosario: Universidad Tecnológica Nacional.
- Meyer, D. F. (2015). Support vector machines. *The interface to libsvm in package*, 28.
- MIT Technology Review. (2017). *Aprendizaje automático: La nueva base de prueba para la ventaja competitiva*. Massachusetts-EU: MIT y Google cloud.

- Moreno, C. G. (2006). *Algoritmos de aprendizaje: knn & kmeans. Inteligencia en Redes de Comunicación*. Madrid: Universidad Carlos III de Madrid.
- Muhammad Ridwan Andi Purnomo, A. A. (2020). *Effective Marketing Strategy Determination Based on Customers Clustering Using Machine Learning Technique*. Islam indonesia: Journal of physics: conference series.
- Naranjo, S. (2021). Tendencias 2021; Hacia dónde va el ecommerce y en qué invertir. *Forbes*.
- Norvig, S. R. (2009). Artificial intelligence - a modern approach. *Cambridge core*, 78-79.
- Niguez, R. (2 de Septiembre de 2021). *Forbes México*. Obtenido de <https://www.forbes.com.mx/negocios-ecommerce-mexico-crecera-226-2025-las-tiendas-desapareceran/#:~:text=Sin%20embargo%2C%20el%20comercio%20electr%C3%B3nico,2019%20y%208%25%20del%202020>.
- Orlando Troisi, G. M. (2019). *Growth hacking: Insights on data-driven decision-making from three firms*. Roma: Elsevier.
- Pareja, G. (2021). *SEO Aumenta tus ventas con google: Aprende a optimizar tu sitio web en buscadores*. Buenos aires: Gonzalo Pascual.
- pwc. (15 de 8 de 2022). *pwc*. Obtenido de Cómo potenciar la función fiscal con inteligencia artificial: <https://www.pwc.com/mx/es/funcionfiscaldigital/20190516-potenciar-ffd-ia.html>
- Reuters. (1 de Febrero de 2022). *El economista*. Obtenido de El economista: <https://www.eleconomista.com.mx/mercados/Alphabet-obtuvo-ingresos-de-75300-millones-de-dolares-en-el-cuarto-trimestre-de-2021-supera-expectativas-20220201-0065.html>

Saura, J. R. (2020). *Using data science in digital marketing: Framework, methods, and performance metrics*. Madrid España: Journal of innovation & knowledge.

Serrano Sandoval, L. J. (2018). Algoritmos de aprendizaje automático para análisis y predicción de datos. *Revista Tecnológica; no. 11.*, 109-114.

Silverman, B. W. (2018). *Density estimation for statistics and data analysis*. Boca Raton: Routledge.

statista. (04 de 11 de 2021). *statista*. Obtenido de statista: <https://es.statista.com/estadisticas/600210/principales-marcas-tecnologicas-en-el-mundo/>

statista. (04 de 11 de 2021). *statista*. Obtenido de statista: <https://es.statista.com/estadisticas/634462/cuota-de-mercado-mundial-de-los-motores-de-busqueda/>

statista. (6 de Febrero de 2022). *Statista*. Obtenido de <https://es.statista.com/estadisticas/1073677/usuarios-internet-pais-america-latina/>

T., I. (3 de 12 de 2021). *isabeltaracena.com*. Obtenido de <https://www.isabeltaracena.com/>

Tello, J. C. (2007). *Reconocimiento de patrones y el aprendizaje no supervisado*. Madrid: Universidad de Alcalá.

Thomas H. Davenport, A. G. (10 de Marzo de 2023). *Harvard Business Review*. Obtenido de <https://hbr.org/2021/07/how-to-design-an-ai-marketing-strategy>

Thomas, M. (3 de Abril de 2023). *Built in*. Obtenido de <https://builtin.com/artificial-intelligence/machine-learning-marketing>

Unir. (18 de 07 de 2023). *Unir la universidad en internet*. Obtenido de Unir la universidad en internet: <https://www.unir.net/marketing-comunicacion/revista/machine-learning-marketing/>

While, T. (2012). *Hadoop the definitive guide*. O'Reilly.

ANEXOS

En el mismo orden que se mencionaron en el texto

Los anexos recogen información adicional que no es esencial para la comprensión del texto, aunque resulten de interés. Incluyen en su título la palabra Anexo, escrito en mayúscula, seguido de número o letra y de un título explicativo, si el autor lo entiende necesario.

Semana	Usuarios
0	1
1	1
2	2
3	45
4	5
5	1
6	3
7	3
8	1
9	2
10	9
11	2
12	9

Tabla 3 Usuarios de la pagina web Isabel Taracena por semana, en la semana 3 donde se realizo la primera publicacion en redes sociales con el enlace a la pagina web.

Nombre	Dirección	coordenadas	pagina web
Vioo	Tonalá 97, Roma Norte, C.P.0670	19.41678345298456, -99.16245765905	vioo.mx/servi
Klik Fotografía	R'ó Becerra 168, San Pedro de los	19.39413715364172, -99.18113598388	https://www.
Galeana Visual Fotografía	Viad. R'ó Becerra 168, San Pedro	19.394146008681446, -99.1812365667	galeanavisual.
FOTO ESTUDIO G.CANTON	MERCADO MEDELLIN LOC 467-46	19.41046668441244, -99.16270134390	https://foto-e
Foto Studio 93 CDMX	Uxmal 173-local B-1, Narvarte Po	19.395099426762705, -99.1546541022	NA
FOTO VARGAS	Cerrada De, Monte L'bano 16, Lor	19.418373072242385, -99.2151759887	NA
A POLI FOTO ESTUDIO	Av Instituto Politécnico Nacional	19.488861515322483, -99.1335847427	NA
Foto Estudio Salazar	C. 1 #69, Valent'n G—mez Far'as,	19.417021668211984, -99.0972560151	https://www.
Foto Estudio Solano	Luz Savi—on 702, Col del Valle Nte	19.39270737413485, -99.16561955500	https://www.
Foto Estudio IPN	Av Instituto Politécnico Nacional	19.494678129568694, -99.1324785039	NA
Foto Estudio Reforma	Mercado 10, Guerrero, Cuauhtz'm	19.449846302862152, -99.1388286776	NA
Foto estudio Prisma	Pto. Tampico 173, Casas Alem'n,	19.475884774759972, -99.0805392583	NA
Estudio Fotografico Alori	Cuernavaca 126, Colonia Condesa	19.411143138426013, -99.1769311979	NA
Foto Estudio Oropeza	C. Membrillo 180-Local 3A, Nueva	19.46493061616882, -99.16947171851	NA
Rodr'guez Sandoval Jonathan	C. Membrillo 195, Nueva Santa M	19.46544145496747, -99.16927055277	NA
Foto Estudio Williams	Av. Cuauhtz'moc 51, La Romita, R	19.423256254105443, -99.1541048537	NA
Foto Estudio "Nikon"	Av. M'xico 119, Raul Romero, 57	19.39899948625465, -99.04143601448	NA
Foto Estudio Herrera	Miguel Lerdo de Tejada 347, Tezo	19.485843043805662, -99.1972306844	NA
Foto Estudio Maricela	06300, Mercado 10, Guerrero, Cu	19.449679389146567, -99.1389104834	NA
Fotos pasaporte	Av. Ricardo Flores Mag—n 237, G	19.4500598513662, -99.139023955036	NA
Foto Ruiz	Mercado 10, Guerrero, Cuauhtz'm	19.44985246168424, -99.13879596729	NA
Foto&Video Danubio	Calle de mercado #10local_B, Gue	19.44986211990158, -99.13899363191	https://fotovi
Fotos Inn (Foto Estudio)	Av. de los Maestros 45-local B, Ag	19.446026624340565, -99.1669614803	https://fotosi
One Shooting Fotografía Profesional	Santa Mar'a La Ribera 119, Sta M	19.44723630163462, -99.15717721019	NA
Foto Luna	Calz. Gral. Mariano Escobedo 71-	19.451089545050667, -99.1821075243	NA
Foto Estudio Luz	Calle de Tacuba #87 Despacho 24	19.43511181416949, -99.13426364835	NA
Foto Latino	Av. Insurgentes Sur 160, Roma Nt	19.422854905077063, -99.1630062868	NA
Foto Estudio Machain	Av. Ricardo Flores Mag—n 432, St	19.45505435774918, -99.15958662005	NA
Foto Shop Estudio	Av Arcos de Bel'zn 59 C, Colonia C	19.42742472036428, -99.14670235438	NA
ESTUDIO FOTOGRAFICO KAPELLMANN	Calle Aztecas 15, Centro de Azcap	19.479935543022723, -99.1871846347	NA
Foto Estudio Arte	Lerdo 31 local f sobre Magnolia, C	19.44246843290195, -99.14278999785	NA
DASBILD PHOTOGRAPHY	Av Instituto Politécnico Nacional	19.496764647297415, -99.1321618543	NA
Estudio 2957 Estudio fotogr'fico profesional	Adolfo Prieto 1758, Del Valle, Ben	19.365584386857783, -99.1741448353	NA
Estudio F22	Agust'n Gonz'lez de Coss'ó 206, C	19.391785467133893, -99.1693888199	NA
Foto Estudio Lucy	C. Viveros de la Hacienda 6, Habit	19.529299747237285, -99.2195035254	foto-estudio-l
Foto Estudio Gaby	Av. del Trabajo 25-L-D, San Andre	19.540332539622813, -99.2212966897	NA
Foto Estudio Gregor	Jos' Mar'a Morelos 51, San Javier	19.538638167003715, -99.1914618400	http://www.f
Estudio de fotograf'ya y Video IMAGINARTE	C. Cerro de la Malinche 209, Hab	19.547414808351824, -99.2196078360	NA
Tiny Studio Photography	Av. de los Reyes 12-s—tano, Resic	19.553177440181873, -99.2087177592	http://www.t
Foto Estudio Torresal	Calle Cuauhtz'moc, Av R'ó Lerma	19.542608530817798, -99.1902808912	NA
FOTO ESTUDIO SAN FRANCISCO	C. Dr. Jos' Mar'a V'rtiz 189, Doct	19.418294619974656, -99.1478961310	https://sites.g
FOTO ESTUDIO BROADWAY	Calz. de la Viga 1249, Reforma lzt	19.377485107937375, -99.1228529869	NA
FOTEC Fotograf'ya y video profesional	Business Center Black, Av. Popoca	19.364314809322217, -99.1611330568	https://www.
Estudio fotogr'fico Olimpia	C. Ote. 160 231, Moctezuma 2da	19.43611433796511, -99.09579321452	https://estudi
Foto Estudio Maricela	06300, Mercado 10, Guerrero, Cu	19.449840909117285, -99.1388356299	NA
ARTIFICE FOTOGRAFIA	FELIX U.GOMEZ #7, Int: LOC. B, Gu	19.449725833010227, -99.1386451931	NA
Fotograf'ya Masao Itami	Av Miguel Hidalgo 296 ByC, Del Ce	19.349534126686702, -99.1611333786	https://www.
MACROPHOTO - fotograf'as	Santiago Sur, Iztacalco, 08800 Ciu	19.38774582914972, -99.12747857637	NA
Figallo Comunicaciones	Hamburgo 108 / 403, 06600 Ju're	19.427303205525433, -99.1625882311	https://figallo
Glitch Atom	Centro, 50000 Toluca de Lerdo, N	19.290123279196823, -99.6550276310	https://glitch
Kiwis Estudio	Corregidora 24-B, San Jer—nimo l	19.326911604664367, -99.2279750611	http://www.ig
Grupo Pixart	Presa Valsequillo 8B, Irrigaci—n, N	19.442361940196477, -99.2136777411	https://www.
FOTOESTUDIO CAMACHO	Calz Legaria 848-D, Col. Irrigaci—	19.442450464302606, -99.2150027523	NA
FOTO ESTUDIO PICAZO	Calz Legaria 827, Col. Irrigaci—n,	19.443217162570882, -99.2141224858	NA
Foto Estudio Clark	Calz Legaria 753, Col. Irrigaci—n,	19.4434142957, -99.21412345624	NA

Tabla 4 Directorio de empresas de fotografías y sus ubicaciones en ciudad de México que se emplea para el estudio de geomarketing.