



EDUCACIÓN

SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO NACIONAL DE MÉXICO

Instituto Tecnológico de Orizaba

“2020, Año de Leona Vicario, Benemérita Madre de la Patria”

DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN

OPCIÓN I.- TESIS

TRABAJO PROFESIONAL

“APLICACIÓN DE LA MINERÍA DE OPINIONES
PARA LA GENERACIÓN DE ESTRATEGIAS DE
MARKETING DIGITAL DE UNA EMPRESA
DE SERVICIOS”.

QUE PARA OBTENER EL GRADO DE:
MAESTRA EN INGENIERÍA
ADMINISTRATIVA

PRESENTA:

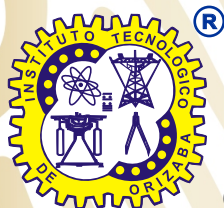
Lic. Diana Josheline Lozada Dario

DIRECTOR DE TESIS:

Dr. Eduardo Roldán Reyes

CODIRECTOR DE TESIS:

M.A.E. Maricela Gallardo Córdova



ORIZABA, VERACRUZ, MÉXICO.

FEBRERO 2020



EDUCACIÓN

SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO NACIONAL DE MÉXICO

Instituto Tecnológico de Orizaba

"2020, Año de Leona Vicario, Benemérita Madre de la Patria"

FECHA: 10/02/2020
DEPENDENCIA: POSGRADO
ASUNTO: Autorización de Impresión
OPCIÓN: I

C. DIANA JOSHELIN LOZADA DARIO
CANDIDATO A GRADO DE MAESTRO EN:
INGENIERIA ADMINISTRATIVA

De acuerdo con el Reglamento de Titulación vigente de los Centros de Enseñanza Técnica Superior, dependiente de la Dirección General de Institutos Tecnológicos de la Secretaría de Educación Pública y habiendo cumplido con todas las indicaciones que la Comisión Revisora le hizo respecto a su Trabajo Profesional titulado:

"APLICACION DE LA MINERIA DE OPINIONES PARA LA GENERACION DE ESTRATEGIAS DE MARKETING DIGITAL DE UNA EMPRESA DE SERVICIOS".

Comunico a Usted que este Departamento concede su autorización para que proceda a la impresión del mismo.

A T E N T A M E N T E


MARIO LEONCIO ARRIOJA RODRIGUEZ
JEFE DE LA DIV. DE ESTUDIOS DE POSGRADO

Avenida Oriente 9 Núm. 852, Colonia Emiliano Zapata, C.P. 94320 Orizaba, Veracruz, México

Tel. 01 (272) 7 24 40 96, Fax. 01 (272) 7 25 17 28 e-mail: orizaba@itorizaba.edu.mx

www.orizaba.tecnm.mx





EDUCACIÓN

SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO NACIONAL DE MÉXICO

Instituto Tecnológico de Orizaba

"2020, Año de Leona Vicario, Benemérita Madre de la Patria"

FECHA : 27/01/2020

ASUNTO: Revisión de Trabajo Escrito

C. MARIO LEONCIO ARRIJOA RODRIGUEZ
JEFE DE LA DIVISION DE ESTUDIOS
DE POSGRADO E INVESTIGACION.
P R E S E N T E

Los que suscriben, miembros del jurado, han realizado la revisión de la Tesis del (la) C. :

DIANA JOSHELINE LOZADA DARIO

la cual lleva el título de:

"APLICACION DE LA MINERIA DE OPINIONES PARA LA GENERACION DE ESTRATEGIAS DE MARKETING DIGITAL DE UNA EMPRESA DE SERVICIOS".

Y concluyen que se acepta.

A T E N T A M E N T E

PRESIDENTE : DR. EDUARDO ROLDAN REYES

SECRETARIO : M.A.E. MARICELA GALLARDO CORDOVA

VOCAL : DR. GUILLERMO CORTES ROBLES

VOCAL SUP. : M.C. NURIA ORTEGA PETTERSON

FIRMA

FIRMA

FIRMA

FIRMA

EGRESADO(A) DE LA MAESTRIA EN INGENIERIA ADMINISTRATIVA

OPCION: I Tesis



Agradecimientos

A Dios, por todas sus bendiciones.

A mis padres, han sido un apoyo fundamental en cada paso que doy. Mi pilar, mi sustento. Son ejemplo de amor y vida. ¡Tienen doctorado en ser padres!

A mis hermanas, Marilyn, Vicky y Jessy, sigan luchando por sus sueños.

Al CONACYT y PNPCC por aceptarme en su programa y permitirme realizar el sueño de estudiar un posgrado.

Al Instituto Tecnológico de México y al Instituto Tecnológico de Orizaba por permitirme enfrentarme a un nuevo reto, por profesionalizarme y obtener nuevos conocimientos.

Al Dr. Eduardo Roldán Reyes, por su apoyo, guía, atención y enseñanza.

Al Mtra. Maricela Gallardo, por su interés en formar profesionales de calidad, por la motivación.

Al Dr. Vijay Mago por recibirme en su "DataLab" y por estar dispuesto a compartir sus conocimientos en Lakehead University, Canadá.

A todos y cada uno de mis profesores que me enseñaron tanto. ¡Mil Gracias!: Mtra. Edna, Mtra. Gaby, Mtra. Zepeda, Mtra. Nuri, Mtro. Marcos, Dr. Memo y Dr. Aguirre.

Itai y Vero, gracias por su amistad incondicional en esta aventura. ¡Los quiero mucho!

Índice

Resumen	I
Abstract	II
Capítulo 1: Generalidades	1
1.1. Introducción	1
1.2. Posicionamiento de la tesis	1
1.3. Planteamiento del problema	4
1.4. Objetivo general	7
1.4.1. Objetivos específicos	7
1.6. Propuesta de solución A priori	10
1.7 Metodología de la investigación	11
1.8. Organización de la Tesis	15
1.9 Conclusión	16
Capítulo 2: Marketing Digital	17
2.1 Introducción	17
2.2 Principios del Marketing Digital	19
2.2.1 El Marketing e Internet	19
2.2.2 Marketing Digital	21
2.2.3 Medios propios, ganados y pagados	22
2.2.4 Panorama y tendencia de las plataformas del Marketing Digital	23
2.3 Aplicación del Marketing Digital	26
2.4 Beneficios del Marketing Digital	28
2.5 E-business: Modelos de negocios digitales	29
2.6 Plan de Marketing Digital	35
2.6.1 El entorno del Marketing Digital	37
2.6.2 Análisis de demanda	38
2.6.3 Análisis del competidor	40
2.6.4 Fijación de objetivos	40
2.6.5 Auditoria del plan de Marketing Digital	41
2.6.6 Construcción de marca	42
2.6.7 Marketing Mix	43
2.6.8 Modelo para la planificación de Marketing Digital: SOSTAC ®	46
2.7 Conclusión	50

Capítulo 3: Minería de la Web Social	52
3.1 Introducción	52
3.2 Minería de la Web	53
3.2.1 Contenido web y estructura de la Minería Web	55
3.2.2 Minería de Uso Web (Análisis Web)	56
3.2.3 Métricas de Análisis Web	59
3.2.5 Fuentes de tráfico	60
3.2.6 Perfiles del internauta	61
3.3 Minería de Texto	62
3.3.1 Análisis de texto y Minería de Texto conceptos y definiciones	62
3.3.2 Procesamiento del Lenguaje Natural	65
3. 3. 3 Minería de Opinión	67
3.4 Proceso de Minería de Opinión	69
3.5 Conclusión	85
Capítulo 4: Aplicación de un modelo de Minería de Web Social utilizando Marketing Digital y Hashtags	87
4.1 Introducción	87
4.2 Plan de Marketing Digital	88
4.2.1 Plan Integral de Estrategias para el cumplimiento de metas institucionales del Grupo Educativo Univo	88
4.3 Plan de Marketing Digital	90
4.3.1 Análisis de situación	90
4.3.1.1 Análisis de competencia	91
4.3.1.2 Análisis FODA	95
4.3.1.3 Objetivos estratégicos	96
4.3.1.4 Plan de Marketing Digital: estrategias y tácticas	97
4.3.1.5 Aplicación de la Minería de Web Social y Minería de Opinión en plataformas digitales	102
4.4 Aplicación de la Minería de Opiniones	105
4.4.1 Modelo para la Minería de Opiniones	105
4.5 Conclusión	127
Capítulo 5: Conclusiones Generales	129
5.1 Introducción	129
5.2 Conclusiones generales	129

5.3 Aportaciones	133
5.4 Trabajos futuros	134
5.5 Vinculación	135
Anexos	139
Anexo 1.- Diagrama de Minería de Opinión para analizar los comentarios.	139
Anexo 2.- Modelo de Minería de Opinión en RapidMiner Studio 9.1	140
Anexo 3.- Modelo de Minería de Opinión en RapidMiner Studio 9.1	141
Anexo 4.- Extracto de la base de datos de opiniones y su polaridad	142
Referencias bibliográficas	145

Índice de tablas

Tabla 1.- Comparación del uso de etapas con diferentes autores sobre la Minería de Datos, Textos y Opiniones.....	11
Tabla 2.- Modelos del negocio enfocados al e-business (Chaffey & Ellis-Chadwick, 2014).....	33
Tabla 3.- Ejemplo de sentimientos positivos y negativos	75
Tabla 4.- Selección de polaridad de términos.....	77
Tabla 5.- Análisis Grupo Educativo Univo.	91
Tabla 6.- Análisis de competencia directa de Grupo Educativo Univo.	92
Tabla 7.-Análisis de competencia indirecta de Grupo Educativo Univo.	93
Tabla 8.- Asignación de importancia	99
Tabla 9.- Escala de clasificación del 1 al 9 conforme a atributos.	99
Tabla 10.- Consenso de clasificación conforme a variables.	99
Tabla 11.- Normalización de valores	100
Tabla 12.- Comprobación	100
Tabla 13.- Plan de Marketing Digital de Grupo Educativo Univo.....	101
Tabla 14.- Calendario de las acciones y actividades que se realizaron.....	103
Tabla 15.-Polaridades obtenidas.....	117
Tabla 16.- Matriz de confusión SVM	119
Tabla 17- Matriz de confusión Naïve Bayes.	120
Tabla 18.- Matriz de confusión Árbol de Decisión.	120
Tabla 19.- Matriz de confusión Regresión Logística.....	120
Tabla 20.- Matriz de confusión Deep Learning	121
Tabla 21.- Tabla comparativa de rendimiento de algoritmos.....	121
Tabla 22.- Tabla de aspectos positivos.....	122
Tabla 23.- Gráfica de followers	132

Índice de ilustraciones

Ilustración 1.- Relación gráfica entre Marketing Digital, Minería Web y Minería de Opiniones	3
Ilustración 2.- Metodología del proyecto.....	13
Ilustración 3.- Marketing Mix (Jerome McCarthy ,1960)	20
Ilustración 4.-Intersección de los tres principales tipos de los medios en línea (Chaffey & Smith, 2013)	22
Ilustración 5.- Embudo del conversión, antes de la era digital. Funnel tradicional de Marketing (Giuffra, 2019)	25
Ilustración 6.- Embudo de conversión en era digital. Funnel del Marketing Digital. (Giuffra, 2019)	25
Ilustración 7.- Intercambio en línea según el autor (Combe, 2006)	35
Ilustración 8.- Modelo del plan de Marketing Digital según (Combe, 2006)	37
Ilustración 9.- 7 P's del Mix del Marketing Digital según (Chaffey & Smith, 2013)	44
Ilustración 10.- SOSTAC®: etapas del modelo. SOSTAC® es una marca registrada de PR Smith (www.prsmith.org).....	47
Ilustración 11.-Taxonomía de la Minería Web (Ramesh, Dursun, & Turban, 2015)	55
Ilustración 12.- Minería de Texto y áreas de aplicación relacionadas.	63
Ilustración 13.-Procesamiento del lenguaje natural en el procesamiento de texto como paso en el proceso de la Minería de Opinión	71
Ilustración 14.-Ejemplo de Case Conversation.....	72
Ilustración 15.- Ejemplo de Stopwords	72
Ilustración 16.-Aplicación de SentiWordNet en Minería de Opinión.....	76
Ilustración 17.- Programa RapidMiner Studio, versión 9.1	81
Ilustración 18.- Operadores en RapidMiner, versión 9.	82
Ilustración 19.-Procedimiento de análisis de datos en RapidMiner	85
Ilustración 20.- Integración e interacción del plan Integral de estrategias, plan de comunicación, plan de Marketing y plan de Marketing digita de Grupo Educativo Univo.....	90
Ilustración 21.-Análisis F.O.D.A del Marketing Digital de Grupo Educativo Univo.	95
Ilustración 22.- Análisis C.A.M.E conforme a objetivos estratégicos del plan de Marketing Digital de Grupo Educativo Univo.	96
Ilustración 23.- Clasificación de objetivos conforme a beneficios de Marketing Digital de Grupo Educativo Univo.	97
Ilustración 24.- Diagrama de Minería de Opinión para analizar los comentarios.	105
Ilustración 25.- Rastreador de Etiquetas de Texto: Social Searcher	106
Ilustración 26.-Los vectores son separados por el hiper-plano maximizando la separación entre las clases.	108
Ilustración 27.- Formula y ejemplo del comportamiento del algoritmo Naïves-Bayes.	109
Ilustración 28.- Ejemplo del comportamiento de algoritmo Árbol de Decisión.	110
Ilustración 29.- Función Sigmoides	111
Ilustración 30.- Modelo de Minería de Opinión en RapidMiner Studio 9.1.....	113
Ilustración 31.- Operador Read Excel.....	114
Ilustración 32.- Operador Set Role	114
Ilustración 33- Operador Nominal to Text	115

Ilustración 34.- Operador Process Documents	115
Ilustración 35.- Operadores: "Tokenize, Transform Case, Replace Tokens, Filter Stopwords" en el modelo de Minería de Opinión.	116
Ilustración 36.- Operador Cross Validation y Apply model.	116
Ilustración 37.- Operadores SVM, Apply Model y Performance en la etapa de capacitación.	117
Ilustración 38.- Operadores Selec Atributes y Write Excel.	117
Ilustración 39.- Desempeño de los algoritmos aplicados en el modelo de Minería de Opinión. ...	118
Ilustración 40.- Procedimiento en RapidMiner para crear una Nube de Palabras.....	123
Ilustración 41.- Operador Set Role	123
Ilustración 42.- Operador Select Attributes	124
Ilustración 43.- Operador Nominal to Text	124
Ilustración 44.- Operador Process Documents	124
Ilustración 45.- Operadores: "Tokenize, Transform Case, Replace Tokens, Filter Stopwords" en el modelo de Nube de Palabras.	125
Ilustración 46. - Operador WordList to Data.....	126
Ilustración 47.- Operador Execute R	126
Ilustración 48.- Nube de palabras de aspectos positivos.	127
Ilustración 49.- Polaridad de comentarios	130
Ilustración 50.- Número de seguidores enero 2018	131
Ilustración 51.- Número de likes en 2019	132

Resumen

El mercado actual no es el mismo que el de hace algunos años, mucho menos las organizaciones y por supuesto, los clientes. Para la gran mayoría, el Internet ha revolucionado el proceso de diferentes empresas y por ende, el consumo. Hoy en día, las organizaciones utilizan los medios digitales como una ventaja competitiva que les permite alcanzar nuevos mercados, encontrar socios estratégicos, proveedores y materia prima, así como entender a su consumidor y/o cliente y por consecuencia, realizar una planeación estratégica mercadológica enfocada a las necesidades de los mismos.

Por otra parte, es un hecho que los medios digitales generan diariamente muchos datos, estos, en la mayoría de los casos, contienen información valiosa que si pasa por desapercibida o no se procesa correctamente se podrían perder elementales indicadores digitales. La importancia de esta información recae en que puede aportar una gran ventaja a la hora de planificar estrategias puesto que, pueden encontrarse muchas y diferentes áreas de oportunidad.

En los últimos años, se han desarrollado muchas estrategias para extraer estos datos de plataformas digitales como las que ofrece Internet, la Minería de Opiniones es una opción con la que se puede trabajar dicha área ya que, esta es una herramienta orientada al análisis de textos donde se busca determinar la polaridad de las opiniones expresadas por el emisor hacia un determinado asunto. Las ventajas que ofrece el uso de esta herramienta en los medios digitales permite clasificar los datos en tiempo real de campañas de publicidad *online* y también que puede contribuir en la planeación de estrategias del Marketing Digital.

La Minería de Opiniones, utilizada como herramienta de extracción de información se piensa aplicar en una campaña de Marketing Digital a Grupo Educativo Univo con el objetivo de generar estrategias *online* fundamentadas en estadísticas y datos.

Abstract

Market is not the same from some years ago, much less the organizations and of course, the customers. For the vast majority, the Internet has revolutionized the process of different companies and therefore, consumption. Nowadays, organizations use digital media as a competitive advantage that allows companies to reach new markets, find strategic partners, suppliers and raw materials. As well as, understand their consumer and / or customer much better. Having said that, carry out in their own planning market strategies like the ones that are more focused on their company needs.

On the other hand, it is well known that digital media generate a lot of data every day, thus, in most cases, contain valuable information and that if it goes unnoticed or if it is correctly processed, elementary digital indicators could be lost. The importance of this information lies in the fact that it can provide a great advantage when planning strategies for many different areas of opportunity can be found.

Recently, many strategies have been developed to extract this data from digital platforms such as those offered by the Internet. Opinion Mining is an option for making a data analysis in this digital area. This is a tool oriented to the analysis of texts where it seeks to determine the polarity of the opinions expressed by the issuer towards a certain issue. The advantages offered by using it in digital media are that: it allows to classify the data in real time of *online* advertising campaigns and also, that it be contributed in the planning of digital Marketing strategies.

Mining Opinion as a Text Mining tool is thought to be applied in a digital Marketing campaign in "Grupo Educativo Univo" with the aim of generating *online* strategies based on data and statistics.

Capítulo 1: Generalidades

1.1. Introducción

En este capítulo se presenta la base sobre la cual se estructura el proyecto de tesis, en primer lugar, se define el posicionamiento de la misma, en este punto se especifica el lugar correspondiente del presente proyecto con respecto a los trabajos con aplicaciones similares, de igual forma sirve para delimitar el alcance del proyecto y diferenciarlo del estado del arte en el que está inmersa.

También se expone el planteamiento del problema, en él se describe la situación de una empresa de servicios, la cual se ha interesado en integrarse al mundo ampliando sus canales de comunicación mediante su sitio web y redes sociales. Sin embargo, a pesar de que la misma tiene un departamento de Marketing y comunicación, los resultados obtenidos no han sido óptimos; ya que existe poco tráfico e interacción en las redes sociales, todo esto como consecuencia de la falta de planificación de estrategias Marketing Digital para generar compromiso de interacción entre los usuarios (*engagement*) sus plataformas digitales.

Más adelante, se plantea tanto el objetivo general como los objetivos específicos, posteriormente se encuentra la justificación de la tesis en donde se definen las razones principales para desarrollar el proyecto.

De igual forma, se propone una posible solución, aquí se detalla la manera en que se busca dar respuesta a la situación problemática, también se define a grandes rasgos la idea de cómo se propone generar estrategias de Marketing Digital a partir de la aplicación de Minería de Opiniones a los clientes internautas de la organización.

Por último, se expone la organización de la tesis, en este punto se detallan los aspectos contenidos en cada uno de los capítulos que componen el presente proyecto de tesis.

1.2. Posicionamiento de la tesis

La tesis se clasifica en la línea de investigación de Administración del Desarrollo Empresarial del programa de Maestría en Ingeniería Administrativa ya que, se pretende detectar las áreas de oportunidad y mejora del Marketing Digital del Grupo Educativo

Univo a través del aprovechamiento de sus recursos y herramientas digitales por medio de una herramienta de Minería de Textos como lo es la Minería de Opiniones; el proyecto se centra en la detección de polaridad de los comentarios emitidos por los usuarios y/o seguidores de las redes sociales de Grupo Educativo Univo por medio de un rastreador de hashtag. Esto con la finalidad de extraer datos implícitos de los comentarios para procesarlos y posteriormente, convertirlos en conocimiento útil para detectar áreas de oportunidad conforme a la toma de decisiones de Marketing Digital de la institución, como se muestra en figura 3.0.

Es importante mencionar que esta investigación de tesis ya se ha trabajado anteriormente en la Maestría de Ingeniería Administrativa pero enfocado a la Administración Estratégica. A continuación, se describirán los trabajos de investigación realizados anteriormente en dónde se combinó de planeación estratégica con la Minería de Datos, la primera investigación que se realizó fue de (Vásquez Rojas, 2014) denominada “Desarrollo de un Modelo de Afinidad Utilizando Minería de Datos para el Análisis del Entorno y Generación de Proyectos Estratégicos”, en donde se propuso utilizar un sistema de categorización mediante un diagrama de afinidad utilizando técnicas de Minería de Textos e implementado un aplicación que fue desarrollada para el sistema Android. Este desarrollo fue incorporado en la segunda etapa del modelo SIGMIL® y su implementación fue probado con éxito en la de determinación de los planes estratégicos de una empresa de distribución de alimentos congelados.

Posteriormente, se realizó una investigación enfocada en la planeación estratégica pero utilizando la herramienta de Minería de Opiniones (Conde Sánchez, 2017) designada con el siguiente título: “Aplicación de la Minería de Opiniones para la extracción del conocimiento significativo durante la creación de los planes estratégicos de las PyMEs” en donde se buscó determinar la polaridad de opiniones expresadas para integrarlas en la priorización de proyectos estratégicos, esta investigación se aplicó en una empresa de construcción en la ciudad de Tehuacán obteniendo como resultado polaridades negativas hacia más de la mitad de proyectos.

Conforme a lo anterior, se puede observar que la Minería de Opiniones es una herramienta versátil que puede ser utilizada en diferentes áreas, como por ejemplo, la Administración Estratégica. Este proyecto de investigación tiene la intención de utilizar la Minería de Opiniones para la generación de estrategias de Marketing Digital en una

empresa de servicios, a partir de conocer la polaridad de opiniones de los usuarios de una campaña *online*.



Ilustración 1.- Relación gráfica entre Marketing Digital, Minería Web y Minería de Opiniones

1.3. Planteamiento del problema

El Internet ha cambiado el mundo de una forma drástica ya que, no solo modificó la forma en que se comunica la sociedad, sino también cómo aprende, se informa y se relaciona. El contenido que se encuentra en la Web crece a una velocidad considerable y cada vez se hace más difícil analizarlo con métodos tradicionales puesto que, la cantidad de información es abundante, mucho más compleja y el número de usuarios se incrementa vertiginosamente. Como cifra referente a este aumento, un estudio elaborado por la Instituto Mexicano de Estadística registró que, del 2013 al 2017, hubo un aumento del 11.5% en el número de usuarios de Internet en México, es decir se alcanzó una cifra de 79.1 millones de personas (INEGI, 2017).

Este aumento de usuarios en Internet se debe a diferentes factores, entre los cuales, se destacan: la facilidad del acceso al Internet y a las Tecnologías de Información, así como a la disminución del costo por uso y al incremento del soporte de tráfico de los sistemas de comunicación. Dichos factores, han logrado cambiar el contexto del Internet, puesto que, éste pasó de ser un medio informativo a uno de interacción comercial, en donde los usuarios se comunican a través de una enorme plataforma globalizada, creando contenido y teniendo una fuerte influencia sobre el éxito de las organizaciones. Bajo este contexto, diversas disciplinas han surgido como resultado del auge del Internet, entre las cuales destaca el Marketing Digital, el cual busca aprovechar los recursos digitales para incrementar las estrategias de crecimiento comercial en las empresas.

Si bien, muchas organizaciones a nivel internacional han aprovechado la capacidad y alcance que ofrece el Marketing Digital a través de sus potentes herramientas, las cuales son de fácil acceso y de bajo costo, un estudio realizado por la Asociación de Internet. MX señala que en México, el sector de servicios ha aprovechado muy poco el uso de tales herramientas ya que, hasta hace algunos años, los únicos sectores de servicios que se dedicaban a explotar las estrategias de Marketing Digital y comercio digital eran las aerolíneas y los servicios financieros (Asociación de Internet. MX, 2014).

Recientemente, dentro del área de servicios, diferentes sectores han demostrado interés en la incorporación de Tecnologías de Información para mejorar sus procesos. Por ejemplo, en el 2017 un estudio enfocado en el comercio electrónico en México señaló que, los procesos de compra han cambiado puesto que los resultados demostraron que, de una muestra de 1,019 personas, el 33% alguna vez ha comprado un transporte

terrestre por medio de una aplicación en su dispositivo móvil y 48% hicieron una compra referente a descargas digitales como por ejemplo: películas, libros, música (Asociación de Internet. MX, 2017). Las anteriores estadísticas demuestran que diferentes organizaciones enfocadas a servicios (desde aerolíneas hasta plataformas de entretenimiento) han adaptado sus estrategias comerciales al mundo digital.

Sin ser de menor importancia, el área educativa también ha tenido un sobresaliente interés en mejorar su proceso de comercialización y comunicación por medio de las Tecnologías de Información. Por ejemplo, un estudio sobre Marketing Digital y Social Media (Asociación de Internet. MX, 2016), identifica qué tipo de publicidad es la que se encuentra con mayor presencia en redes sociales siendo el sector educativo, el que presentó un mayor porcentaje: 21%, con respecto a otros sectores; por otra parte, el mismo estudio revela que 9 de cada 10 encuestados, de una muestra representativa del usuario promedio, utilizan el Internet para buscar ofertas educativas; de estos, 62% buscan en los sitios de las universidades; 19% en sitios especializados y 15% en redes sociales. Los datos anteriores, revelan que las instituciones de educación se han integrado a *social media*, principalmente la educación superior de iniciativa privada. Si bien este fenómeno (i.e. el uso de recursos en línea) ha tenido un desarrollo en las principales universidades del país, principalmente de las grandes ciudades pero por otra parte, en algunos estados se observa un cierto rezago.

Particularmente, en la zona centro del estado de Veracruz, existe un número importante de universidades de iniciativa privada (aproximadamente 16 en la región Córdoba-Orizaba), las cuales han comenzado a realizar esfuerzos para incorporar nuevas tecnologías de información como parte de sus estrategias de crecimiento. Tal es el caso del Grupo Educativo Univo, la cual se ha interesado en integrarse al mundo digital desde el año 2010, ampliando sus canales de comunicación mediante su sitio web y redes sociales. Sin embargo, a pesar de que el grupo educativo tiene un departamento especializado en dicha materia, su situación actual demuestra una falta de explotación de los recursos y herramientas en línea, mismos que podrían incrementar el alcance de su oferta académica. A pesar de que se han hecho esfuerzos en el uso de redes sociales (Facebook, Twitter e Instagram) estas no han logrado tener el impacto esperado, ya que sus publicaciones no logran un tráfico suficiente y proporcional comparado al número de seguidores que poseen. Por ejemplo, en Facebook la página oficial de Grupo Educativo

Univo cuenta con una comunidad de 8,150 usuarios mismos que, representan cerca del 10% de la población Joven y Adulta de Orizaba, Veracruz.

En las publicaciones del 2018, de su página oficial de Facebook del grupo en su bimestre octubre-noviembre, se tiene un promedio de 16 interacciones por publicación de diferentes campañas, donde se puede observar que la interacción está cerca del 0.19% de su total de usuarios. Relevante a lo anterior, es necesario mencionar que una de las políticas internas del grupo educativo es la de privilegiar la publicidad digital por encima de la tradicional (radio, televisión, periódicos, material publicitario, entre otros) y se están destinado diferentes recursos para las estrategias de mercadotecnia y comunicación enfocadas en este sector y que, a pesar de estos esfuerzos como por ejemplo, realizar constantemente publicaciones y pagar por publicidad y alcance del sitio, no se tiene la respuesta deseada por sus seguidores.

En contraste de lo anterior, Grupo Educativo Univo desea conocer el perfil de los usuarios que siguen su página ya que, al conocer diferentes aspectos de estos como por ejemplo, quiénes son, sus hábitos digitales y que piensan de su publicidad y/o campañas en línea podría ayudar al enfoque que se tiene de estrategias de mercadotecnia y comunicación.

1.4. Objetivo general

Implementar un modelo de Minería de Opiniones para identificar la polaridad de opiniones de los usuarios sobre las campañas digitales del Grupo Educativo Univo y contribuir con la toma de decisiones conforme a la planificación de estrategias de mercadotecnia y comunicación digital.

1.4.1. Objetivos específicos

- Generar estrategias de comunicación basadas en el Marketing Digital que permitan incrementar el tráfico de usuarios y fomentar la interacción en las campañas digitales en las redes sociales del Grupo Educativo Univo.
- Reunir los datos obtenidos con relación a las opiniones de los usuarios en las publicaciones realizadas de las plataformas correspondientes para estructurar la base de datos.
- Aplicar la Minería de Opiniones para el análisis de las opiniones que permita analizar y determinar la polaridad de los comentarios expresados por los usuarios mediante el software RapidMiner.
- Desarrollar estrategias de Marketing Digital conforme a los resultados obtenidos que permitan la planificación de futuras campañas *online* que realice el grupo educativo.

1.5. Justificación

Al haber delimitado el problema que pretende resolver esta tesis el cual, es identificar diferentes aspectos de los consumidores y/o clientes conforme al impacto y difusión del Marketing Digital que tiene las campañas *online* del Grupo Educativo Univo a través de sus redes sociales como Facebook, Twitter e Instagram. Lo anterior, no sólo es importante para este grupo educativo, sino para muchas organizaciones puesto que, se ha demostrado en los últimos años que los medios digitales tienen un gran peso conforme a la planeación de estrategias de mercadotecnia de las compañías ya que, la sociedad ha cambiado sus hábitos de consumo y también de comunicación en los últimos años.

Estadísticas muestran que el mexicano promedio pasa aproximadamente 8 horas en Internet, siendo éste el medio de entretenimiento que más utiliza en el día a comparación de la televisión, misma que cuenta con un consumo aproximado de 3 horas y la radio con un consumo de una hora por día (Asociación de Internet. MX, 2018).

Con este mismo enfoque, las personas han migrado de los medios tradicionales como lo es la radio y el periódico a dispositivos con acceso a Internet como lo son los celulares, tabletas, computadoras y Smart TV's. El IFT (Instituto Federal de Telecomunicaciones) señala que el consumo de información por diferentes medios, entre las edades 18 y 29 años se distribuye de la siguiente manera: los medios digitales llevan la delantera con 50%, la televisión con un porcentaje 27% y el radio con una presencia del 14%. Por otra parte, conforme a confiabilidad y credibilidad de los medios audiovisuales, el 28% de los usuarios confía más en lo que ve en Internet, 25% en lo que ve en televisión y 13% confía en lo que escucha en la radio. Bajo estos indicadores se puede concluir que los medios digitales tiene un mayor impacto, penetración y credibilidad entre la población.

Tomando en cuenta lo anterior, es esencial que las organizaciones aprovechen los recursos digitales ya que estos, están siendo utilizados como un indicador de competitividad dentro de las mismas y para estas es de suma importancia estar inmersas en el mundo digital aunque, el "ser digital" es más que ser una empresa en Internet, sino aprovechar al máximo todo estos medios digitales, así como la información que proporcionan (Kotler et al., 2016). Es por ello, que la revolución digital ha tenido un auge en los últimos años y ha modificado no solo a las organizaciones de bienes sino también a las de servicios.

En un principio, para las organizaciones de educación superior, el Marketing solía estar enfocado en las necesidades de aprendizaje de sus estudiantes (consumidores), orientado primordialmente en generar programas educativos que satisfagan las expectativas de los padres (clientes) pero, como se ha mencionado anteriormente, las organizaciones de servicios también se han incorporado al Marketing Digital, dando un enfoque diferente desde sus estrategias de mercadotecnia hasta las de comunicación.

(Chaffey & Smith, 2013) mencionan que el uso y aplicación de tecnologías digitales favorecen positivamente al Marketing en las organizaciones puesto que, al contar con un enfoque adecuados en las estrategias se puede incrementar el conocimiento de los usuarios y tener una mejor comunicación y retroalimentación con los mismos, lo cual,

permite que la planificación de servicios se realice conforme a los requerimientos específicos de los usuarios.

Por otra parte, realizar campañas digitales planificadas sistemáticamente, bien estructuradas y enfocadas al consumidor es todo un reto para las organizaciones actuales puesto que, éstas en la mayoría de los casos ignoran que piensa su consumidor y/o cliente por lo que, se suele tener una retroalimentación muy escasa de las estrategias implementadas ya que, no se conoce realmente hacia que público va dirigida la información y al desconocer lo anterior, se pierden diferentes recursos como tiempo y dinero.

Bajo este contexto, la Minería de Opiniones puede funcionar como una herramienta eficaz para analizar las opiniones que tienen los usuarios de las campañas digitales del Grupo Educativo Univo en sus redes sociales, ya que éstas contienen información fundamental para la toma de decisiones conforme a la planificación de estrategias de mercadotecnia y comunicación. Particularmente, se busca identificar la polaridad de los comentarios expresados por los usuarios de la campaña digital (estos pueden ser positivos, negativos o neutrales), con el fin de identificar qué es lo que se piensa de dicha campaña y por ende, de la institución. De tal manera que una vez reunida dicha información se pueden tomar decisiones sobre el Marketing Digital de las próximas campañas *online*.

Conforme a lo anterior, las razones principales por las que se busca aplicar la Minería de Opiniones en el Marketing Digital se puede resumir a continuación:

1. Permitirá desarrollar un plan estructurado con etapas el cual, pueda adquirir y determinar la polaridad de comentarios en una campaña de Marketing Digital.
2. Facilitará el procesamiento de información de las redes sociales de Grupo Educativo Univo.
3. Permitirá detectar la polaridad de las opiniones y a su vez, identificar las áreas de oportunidad conforme al Marketing Digital de la institución.
4. Contribuirá a la generación de proyecto estratégicos de Marketing Digital y comunicación del grupo educativo al tener datos reales conforme a la opinión de sus usuarios, identificando necesidades de los mismos.

Por otra parte, este proyecto tendrá diversos impactos, entre los cuales los más importantes pueden ser:

- Impacto tecnológico: Al implementar en el uso de tecnologías de información una herramienta de lingüística computacional.
- Impacto social: Este se verá reflejado en el uso de un instrumento que permita el desarrollo de las organizaciones educativas particulares ya que, al identificar el uso correcto o incorrecto de estrategias de Marketing Digital puede afectar de manera directa a sus resultados económicos.
- Impactó económico: Al utilizar las redes sociales como canal dinámico para la implementación de campañas *online* permite tener un gran alcance de usuarios en pocas horas sin invertir grandes recursos.
- Impacto ambiental: Al aprovechar los recursos digitales en una campaña, se reduce de manera drástica la impresión de medios tradicionales impresos como por ejemplo, banners y lonas que una vez utilizados, muy pocas veces son reutilizados y casi siempre son tirados a la basura contribuyendo con la contaminación de la zona.

Por último y no menos importante, el uso de una herramienta de Minería de Textos como lo es Minería de Opiniones (herramienta que un principio ha sido utilizada para resolver problemas informáticos y de ciencias computacionales) ha revolucionado la manera de resolver temáticas de mercadotecnia y específicamente de Marketing Digital. A nivel internacional se han hecho investigaciones pero la mayoría de ellas enfocadas a la programación y estructura de los elementos de Minería de Opiniones para la investigación del impacto de ciertas aplicaciones o sitios web y muy pocas de estas han sido enfocadas a empresas de servicios. En México existen pocas investigaciones al respecto, ya que solo grandes corporativos y por supuesto, empresas internacionales instaladas en el país como Oracle utilizan este tipo de conocimientos y sólo de manera interna. Conforme a lo anterior, el Grupo Educativo Univo una organización, que ofrece la educación como su servicio principal, al utilizar esta herramienta se pone en vanguardia y a la altura de universidades internacionales puesto que, este tipo de técnicas son utilizadas en institutos de renombre como por ejemplo el Instituto Tecnológico de Massachusetts, Universidad de Harvard, Universidad de Stanford, entre otras.

1.6. Propuesta de solución A priori

Mediante la Minería de Opiniones se extraerán los datos generados por las campañas *online*, permitiendo mediar el impacto de las estrategias digitales implementadas en las diferentes plataformas sociales de Grupo Educativo Univo.

1.7 Metodología de la investigación

A continuación, se presentas una tabla en la que se comparan las etapas y pasos que desarrollaron diferentes autores en sus investigaciones referentes a la Minería de Datos, Textos y Opiniones:

Autor	Etapa 1	Etapa 2	Etapa 3		Etapa 4	Etapa 5	Etapa 6
	Selección Extracción	Pre- procesamiento	Procesamiento		Clasificación de opiniones / Análisis de Sentimientos	Resumen	Validación
			Agrupamiento	Categorización			
(Turban et al., 2011)							
(Liu, 2012)							
(Montesinos García, 2014)							
(Vásquez Rojas, 2014)							
(Balazr Thenot, 2015)							
(Miranda et al., 2016)							
(Mars & Gouider, 2017)							
(Othman et al., 2017)							
(Tjahyanto & Sisephaputra, 2017)							
(Tavakoli et al., 2018)							
(Deshmukh & Tripathy, 2018)							

Tabla 1.- Comparación del uso de etapas con diferentes autores sobre la Minería de Datos, Textos y Opiniones.

Como se observa en la tabla anterior, las etapas en las que coinciden la mayoría de los autores son:

1.- Pre-procesamiento: En esta etapa una vez que la información ha sido recabada, ésta es depurada a través de una serie de procesos que la van transformando de manera gradual. Cada uno de estos procesos tiene una finalidad específica, algunos procesos que se destacan en esta etapa son: tokens, stopwords, cases, corrección de términos, filtrado de términos.

2.- Procesamiento: La etapa del procesamiento es un punto medular en el modelo de afinidad debido a la aplicación de dos procesos:

- Agrupamiento
- Categorización

3.- Clasificación de opiniones: En este paso se valoran la información y para ello se le aplica la denominada polaridad, a través de la cual se clasifica el mensaje en función de la intención que tenga el autor al realizarlo, pudiendo ser este positivo, neutro o negativo. Esto permite controlar el sentimiento de los usuarios respecto a la información analizada.

4.- Validación: Por último se valida la información dependiendo de los objetivos específicos a alcanzar, en esta etapa se gestionan todos los posibles aspectos a mejorar.

Analizando anteriormente, la presente metodología se basa en el análisis de diferentes artículos referentes a la Minería de Opiniones en donde los autores de los mismos presentan investigaciones relacionadas con plataformas digitales.

La metodología se llevara a cabo para el desarrollo del presente proyecto se observa en la ilustración 2:

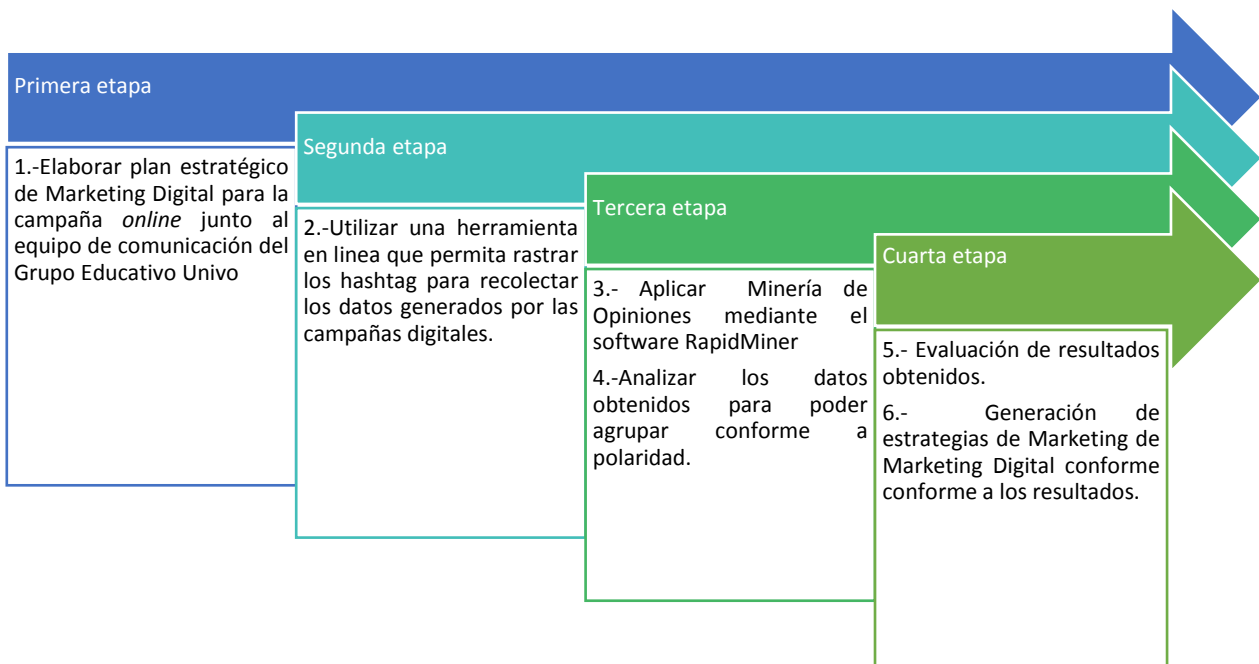


Ilustración 2.- Metodología del proyecto

A continuación, se detallaran cada uno de los puntos que componen las etapas de la figura 2:

Primera etapa

1.- Elaborar plan estratégico de Marketing Digital para las campañas de Marketing Digital junto al equipo de comunicación del Grupo Educativo Univo.

Este paso consiste en colaborar junto con el Grupo Educativo Univo en la elaboración de un plan estratégico de comunicación que se llevará a cabo con el fin de aumentar el tráfico en redes sociales del grupo educativo. El plan permitirá renovar acuerdos con patrocinadores locales de Grupo Educativo Univo para llegar a un común acuerdo en donde se puede cambiar publicidad por servicios con el fin, de promocionar dichos beneficios entre los usuarios y así aumentar el número de comentarios e interacciones.

Segunda etapa

- **2.- Utilizar una herramienta en línea que permita rastrear los hashtags para recolectar los datos generados por las campañas**

Para identificar todos los datos que se generaron a partir de la estrategia de comunicación mencionada anteriormente, se utilizará los siguientes hashtags para identificar la campaña en las diferentes plataformas: #CaminoAlexito y #TriviaUniVO. Estas etiquetas de texto permitirán que un software *online* monitorice las publicaciones y posteriormente analizarlas.

Tercera etapa

- **3.- Aplicar la Minería de Opiniones Mediante el software RapidMiner**

Se realizará la Minería de Opiniones a través del software RapidMiner, el mismo que permitirá analizar a profundidad la polaridad de las mismas.

- **4.-Analizar los datos obtenidos para agrupar conforme a polaridad.**

En esta parte se pretende definir la manera en que se determinarán las polaridades de los comentarios recabados en las redes sociales, es decir, si estos son positivos, negativos o neutrales. En otras palabras, volver la información cualitativa en cuantitativa.

Cuarta etapa

- **5.- Evaluación de resultados obtenidos.**

Se evaluará el desempeño que tuvieron las campañas digitales, con el propósito de medir el impacto del Marketing Digital, así como conocer que piensan los usuarios de la misma y por ende de la institución.

- **6.- Generación de estrategias de Marketing Digital conforme a los resultados**

Por último, una vez obtenida la información sobre el nivel de impacto de la campaña *online*, así como la polaridad de la misma, se desarrollarán estrategias

enfocadas especialmente a las áreas de oportunidad del Marketing Digital del grupo educativo.

1.8. Organización de la Tesis

El presente proyecto de tesis se divide en cinco capítulos que se describen brevemente a continuación:

- Capítulo 1. Generalidades. En este capítulo se presentan las generalidades del proyecto, entre estas se consideran el planteamiento del problema, los objetivos del proyecto, la justificación del proyecto, la propuesta de solución y la metodología que se empleó para el desarrollo del proyecto.
- Capítulo 2. Marketing Digital. En este apartado se presenta la primera parte que compone el marco teórico en el que se llevó a cabo el proyecto, mismo que sirvió como base para el fundamento del área de estudio aplicada. Se presentan temas como principios de Marketing Digital, aplicación del Marketing Digital, beneficios del Marketing Digital y estrategias de Marketing Digital.
- Capítulo 3. Minería de Web Social. Este comprende la segunda parte del marco teórico, esto debido a que se plantea que el presente proyecto de tesis integra estrategias de Marketing Digital basadas en aplicación de Minería de Opiniones, esta última disciplina está inmersa bajo la temática de Minería de Web, en este capítulo se detallan tres subcampos de investigación Minería Web: la Minería de uso de la web, Minería de Contenido Web Social y Minería de Estructura Web.
- Capítulo 4. Aplicación de un modelo de Minería Web social utilizando estrategias de Marketing Digital y uso de hashtags. Se presenta la metodología para el desarrollo y aplicación del modelo propuesto, aquí se describen cada una de las actividades que se llevaron para la estructuración y desarrollo de las estrategias de Marketing Digital. También se presentan cada una de las dinámicas de la campaña digital, así como su implementación en

las plataformas digitales de la organización, así como el redireccionamiento de las estrategias y alcance de las mismas conforme a los KPI's de alcance e influencia como *engagement*, *sentiment*, *share voice* y *network size*.

- Capítulo 5. Conclusiones generales. En este capítulo se complementa el proyecto de investigación con el análisis de polaridad de los resultados obtenidos. Posteriormente, se muestra una evaluación para comprobar la efectividad de los mismos. De igual forma se presentan implicaciones, discusión, desventajas del enfoque y trabajos futuros.

1.9 Conclusión

En el presente capítulo se definieron detalladamente las bases que sustentan el desarrollo del proyecto de tesis, por lo tanto, este capítulo sirve para definir la dirección sobre la cual se plantea trabajar, de igual forma se delimita el alcance del proyecto. Lo anterior permite plantear el contexto teórico bajo el cual se fundamenta el presente proyecto, este se describe en los siguientes dos capítulos como se puede observar en la organización de la tesis.

Capítulo 2: Marketing Digital

2.1 Introducción

El Internet, muy aparte de proporcionar conectividad y transparencia a la vida de las personas, ha sido un gran responsable de cambios (Kotler, 2016). Este gigante tecnológico no tan sólo cambio el contexto digital con el que el Internet comenzó su vida en línea sino incluso este llegó a revolucionar el Marketing tradicional y por ende, el concepto, función e inclusión del mismo. El comportamiento de las masas se ha modificado sin lugar a dudas a raíz de la integración de los medios digitales o en otras palabras de las redes sociales, es muy común que las personas busquen noticias en una red social de confianza como Twitter en lugar de que enciendan su televisor para visualizar un canal de noticias como CNN, mismo que era líder de acceso. No es sorpresa que en la actualidad CNN se encuentre incluido en Twitter y que use el mismo como un canal de confianza para su espectador. Otro ejemplo de este enorme cambio ha sido el medio artístico puesto que, YouTube llegó a cambiar la manera de realizar contenido e entretenimiento. Grandes corporativos del giro de entretenimiento como Sony han unificado fuerzas con las plataformas digitales y se han adaptado a dichos cambios para seguir siendo una empresa exitosa. No es sorpresa que las empresas busquen adaptarse al medio digital puesto que, este poderoso cambio influye directamente en los consumidores mismos que son la principal fuente de influencia, superando las comunicaciones de Marketing externas y el uso las preferencias personales (Kotler et al., 2016)

La conectividad *online* ha permitido que las redes sociales redefinan de manera exponencial la forma en que las personas interactúan entre sí, permitiendo de la manera directa que las personas puedan construir relaciones sin barreras geográficas y demográficas. El impacto de las redes sociales ha sido tan grande que permite facilitar colaboraciones globales en innovación. De hecho, todas las redes sociales que adoptan un enfoque de *crowdsourcing* (colaboración abierta) son buenos ejemplos de exclusividad social. Los medios sociales impulsan la inclusión social y proporcionan a las personas el sentido de pertenencia a sus comunidades.

Es muy común los clientes se expresen de lo que piensan de un producto o servicio por medio de las redes sociales. Según (Kotler, 2016) los clientes ya no son objetivos pasivos,

sino que se están convirtiendo en medios activos de comunicación. En el pasado, las empresas solían controlar y manejar lo que se hablaba de ellas de manera interna y resolver los problemas de manera individual pero en la actualidad, con el contenido generado por los clientes en la comunidad *online*, las empresas pierden totalmente el control de la conversación y es en este punto donde se debe estar preparado como organización para reacciones sociales masivas tanto positivas como negativas y tener estrategias digitales para enfrentarlo de la mejor manera posible puesto que censurar errores solamente empeora la reacción de las masas y la organización pierde credibilidad de marca.

Los encargados de los departamentos de Marketing y Comunicación en las organizaciones deben de tener un pensamiento macro del panorama empresarial donde los principales pilares deben ser: la integración horizontal, la inclusión digital y la importancia de la participación social que tienen sus clientes en las plataformas digitales.

2.2 Principios del Marketing Digital

Los negocios y el Marketing han sido transformados por medio del Internet desde 1991, año en el que salió el primer sitio web de la historia. En la actualidad, hay más de 4.000 billones de personas *online* en el mundo. Es decir, el 53 por ciento de la población mundialmente ya cuenta con acceso a Internet (Statista, 2019), lo cual es algo inmenso y gigante. A raíz de este contexto mundial las organizaciones deben enfocarse en el Marketing Digital para tener éxito en el futuro puesto que, la innovación es un hecho con la introducción continua de nuevas tecnologías, nuevos modelos de negocio y nuevas metodologías de comunicación. El gran desafío al que se enfrentan los equipos de mercadotecnia y comunicación de las organizaciones radican en comprender cuales son las innovaciones más pertinentes para la organización y de esta manera, obtener una ventaja competitiva que les permita alcanzar su visión deseada.

2.2.1 El Marketing e Internet

El uso de Internet, los medios digitales y las plataformas *online* han tenido una gran influencia y el Marketing Digital es implementando en varias plataformas digitales para lograr objetivos y llegar a distintas audiencias.

Las organizaciones se han movilizado hacia mundo *online* en todo el espectro de las actividades de Marketing y estas pueden ir desde crear conocimiento hasta ofrecer el servicio de postventa a sus clientes, y utilizan diferentes herramientas y medios digitales como los elementos importantes para crear estrategias de Marketing y que pueden tener un alcance global gracias al Internet. (Echeverría Gustavo, 2009)

Fue un gran para las organizaciones comprender y adaptarse al cambio del concepto tradicional del Marketing en el cual, la concepción de las 4 P's, que fue desarrollado en los años 60's por el profesor de contabilidad estadounidense Jerome McCarthy para escribir, de manera concreta las áreas más importantes de una empresa en donde se deben de tomar decisiones estrategias de sus productos o servicios.

A continuación se describirá el concepto denominado Marketing Mix o Mezcla de Mercado según (Espinosa Roberto, 2014):

El Marketing Mix se utiliza para englobar a los cuatro componentes básicos del Marketing: producto, precio, plaza y promoción. Estas cuatro variables son conocidas como las 4P's por su aceptación anglosajona (Product, Price, Place y Promotion). Las 4 P's pueden considerarse como variables tradicionales con las que una organización puede conseguir sus objetivos comerciales para ello, es totalmente indispensables que la cuatro variables del Marketing Mix se combinen con coherencia y trabajen conjuntamente para complementarse entre sí.

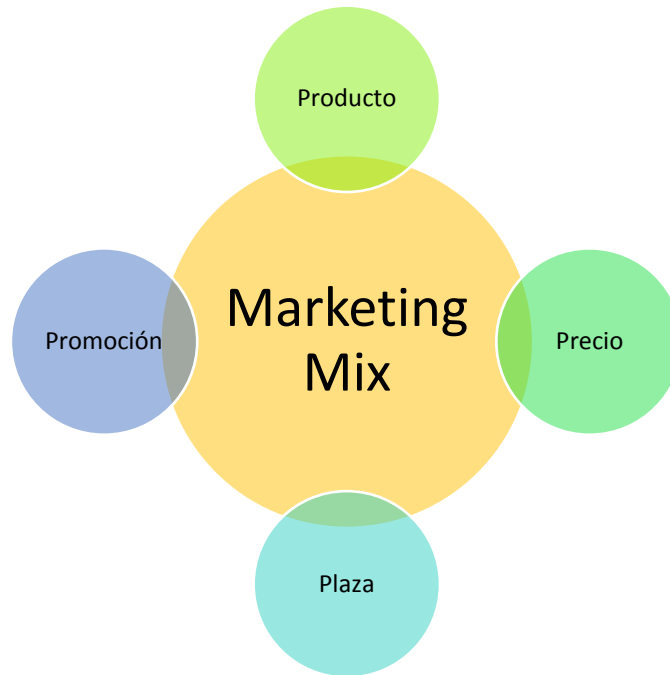


Ilustración 3.- Marketing Mix (Jerome McCarthy ,1960)

- **Producto:** Es la variable por excelencia del Marketing Mix ya que engloba tanto los bienes como los servicios que comercializa una empresa. Es el medio por el cual se satisfacen las necesidades de los consumidores. Por tanto el producto debe centrarse en resolver dichas necesidades y no en sus características tal y como se hacía años atrás.
- **Precio:** El precio es la variable del Marketing Mix por la cual entran los ingresos de una empresa. Antes de fijar los precios de un producto se debe estudiar ciertos aspectos como el consumidor, mercado, costes, competencia, etc. En última instancia es el consumidor quien dictaminará si se ha fijado un precio justo, puesto que comparará el valor recibido del producto adquirido, frente al precio que ha desembolsado por él.

- **Plaza:** En términos generales la plaza consiste en la distribución consiste en un conjunto de tareas o actividades necesarias para trasladar el producto acabado hasta los diferentes puntos de venta. La distribución juega un papel clave en la gestión comercial de cualquier compañía. Es necesario trabajar continuamente para lograr poner el producto en manos del consumidor en el tiempo y lugar adecuado. No hay una única forma de distribuir los productos, sino que dependerá de las características del mercado, del mismo producto, de los consumidores, y de los recursos disponibles.
- **Promoción:** La promoción en términos generales son las estrategias de comunicación de las empresas para dar a conocer los productos que pretenden satisfacer las necesidades de su público objetivo. Se puede encontrar diferentes herramientas de comunicación: venta personal, promoción de ventas, publicidad, Marketing directo y las relaciones públicas. La forma en que se combinen estas herramientas dependerá de nuestro producto, del mercado, del público objetivo, de nuestra competencia y de la estrategia que hayamos definido.

Hacer un producto que cubra una necesidad del mercado, fijar un precio que los consumidores estén dispuestos a pagar y del que puedan obtener un valor agregado, determinar los mejores puntos, posiciones y canales de ventas, así como realizar la promoción adecuada utilizando una comunicación proactiva era una tarea que las organizaciones no se esperaban que revolucionaria gracias al Internet puesto que éste, representa una parte esencial del mercado, puede utilizarse como un medio de comunicación, información y entrenamiento de los consumidores. La función original de la función del Marketing Mix no cambia con la aparición de los medios digitales, pero sí toma otra dimensión, se plantean nuevas estrategias funciones y sirve de modelos innovadores para lograr objetivos (Echeverría Gustavo, 2009)

2.2.2 Marketing Digital

Según (Chaffey & Smith, 2013) el Marketing Digital se define como lograr los objetivos de Marketing mediante la aplicación de tecnologías digitales. Esta definición parece ser corta pero en realidad concientiza que la tecnología tiene un gran poder cuando las organizaciones invierten en el Marketing de Internet.

En la práctica el Marketing Digital comprende la administración de diferentes actividades que permiten la presencia de una organización en línea como por ejemplo en un inicio, los sitio web, blogs y en la actualidad las redes sociales. A pesar, de que el Marketing Digital está inmerso en una comunidad virtual el uso de medios tradicionales complementa y es parte de la comunicación multicanal de las organizaciones en una campaña digital.

2.2.3 Medios propios, ganados y pagados

Para que una organización pueda desarrollar una estrategia digital solida es fundamental comprender el entorno externo el cual es más complejo y competitivo que nunca, en el cual los clientes tienen una importante presencia en línea. Para desarrollar una estrategia directa y que pueda influir en estos, es necesario exponer los tres medios más influyentes del Marketing:

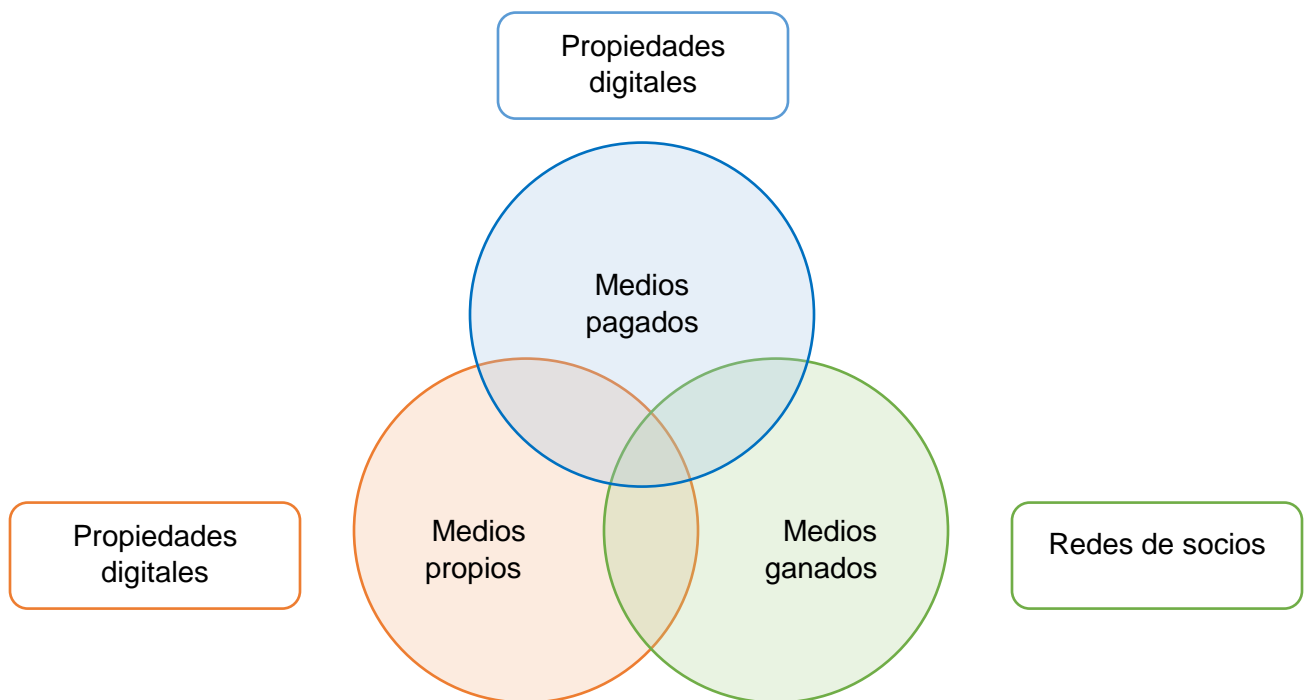


Ilustración 4.-Intersección de los tres principales tipos de los medios en línea (Chaffey & Smith, 2013)

- **Medios pagados:** Usualmente, los medios pagados son comprados o en su defecto se invierte en ellos para obtener visitas, alcance o posicionamiento en buscadores. Estos medios también se pueden encontrar *offline* y generalmente

puede ser la publicidad impresa o televisiva aunque, este tipo de estrategia no es tan recomendable pues el costo de estos es muy elevado.

- **Medios ganados:** Los medios ganados son todos aquellos que por medio de las relaciones públicas se generan estrategias publicitarias en donde se integran a las campañas publicitarias personas influyentes con el objetivo de aumentar la conciencia de marca. A su vez, la publicidad de boca en boca se puede viralizar y estimular a través de plataformas digitales. Los protagonistas de estos medios suelen ser participantes como editores, bloggers, instragramers y otros influyentes.
- **Medios propios:** Es palabras simples los medios propios son todos aquellos que la marca por si misma posee. En línea se puede identificar el sitio web, las redes sociales, blogs, apps. Así mismo, se puede afirmar que la propia presencia de la organización es un medio de comunicación y esta exhorta a la misma para expandir fronteras y se vuelva su propia editora multicanal.

2.2.4 Panorama y tendencia de las plataformas del Marketing Digital

Las organizaciones se han dado cuenta en los últimos años que la manera en que se relaciona el mundo ha cambiado de manera dramática, actividades que comúnmente se hacían cara a cara, actualmente pasan a realizarse en plataformas digitales como por ejemplo, interactuar con la audiencia, realizar alguna búsqueda de un contenido en específico o hacer una transacción económica. El acceso a dichas actividades se realiza por medio de plataformas de hardware para computadoras y móviles por otra parte, (Chaffey & Ellis-Chadwick, 2014) mencionan que aparte de estas plataformas de hardware existen las plataformas de software que los profesionales de Marketing pueden utilizar para llegar a su audiencia de manera directa y positiva a través del Marketing de Contenido y la publicidad puesto que, la digitalización ofrece oportunidades sin precedente desde los grandes corporativos hasta las pequeñas y medianas empresas (Cenamor et al., 2019) puesto que, las plataformas digitales están transformando la manera en que las empresas perciben y construyen una ventaja competitiva (Parker et al., 2016). De hecho, en muchos casos, las plataformas digitales han desempeñado un papel central en muchas propuestas de valor de las empresas al permitirles aprovechar la información misma que puede ser utilizada en su beneficio como por ejemplo, para estrategias digitales (Cenamor et al., 2017)

A continuación, se presentarán algunas plataformas digitales divididas dependiendo los dispositivos electrónicos: computadoras de escritorio, laptops y notebooks (Chaffey & Ellis-Chadwick, 2014):

- Plataforma para computadora de escritorio como navegador: en esta plataforma es la más común y tradicional puesto que, los usuarios tienen acceso a Internet por medio de exploradores como por ejemplo: Google Chrome, Explorer, Mozilla Firefox, entre otros.
- Aplicaciones para computadoras de escritorio: A partir del lanzamiento de Apple Lion muchos usuarios de Apple tienen la posibilidad de descargar aplicaciones gratis a través de sus escritorios en la aplicación de App Store de Apple, en el caso de Microsoft, existe la aplicación Store para descargar de manera gratuita este tipo de aplicaciones e instalarlas en una computadora.
- Plataformas para intercambio de datos en API y basadas en feeds: Diferentes internautas donde se consumen datos a través de feeds RSS, según (Lasso, 2013) Un “Archivo RSS” o “Feed RSS” es un archivo generado por algunos sitios web que contiene una versión específica de la información publicada en esa web. Por otra parte, se pueden tomar en cuenta las actualizaciones que hacen diariamente los usuarios en Twitter y Facebook se pueden considerar una forma de feed o de tráfico continuo en donde se puede aprovechar el uso de anuncios publicitarios.
- Plataformas de Marketing de Video: El flujo de continuo de video suele transmitirse en aplicaciones anteriormente a través, de navegadores y complementos en diferentes aplicaciones.

Además del acceso mediante equipos de escritorio y móviles, hay muchas otras plataformas para comunicarse con los clientes como por ejemplo: plataformas de videojuegos, aplicaciones de quiosco internas o externas, así como la señalización interactiva.

Para las organizaciones los medios digitales ofrecen una gran gama de oportunidades para comercializar productos y servicios a través del ciclo de compra mismo que ha cambiado a través de la implementación del Marketing Digital y por defecto, el comportamiento de compra del consumidor. Este cambio surge a través del impacto que tienen las redes sociales ya que, estas tienen un rol fundamental en el canal de ventas establecido por las empresas. El comportamiento de compra se comprendía hasta algunos años como una especie de embudo mismo que exponía de manera directa que el

acceso a la información era mucho más difícil de encontrar, por otra parte, este mismo mecanismo exhortaba a las organizaciones a invertir cientos de miles y hasta millones en publicidad para lograr que sus marcas fueran reconocidas. Sin embargo, en la actualidad, el comportamiento del consumidor se visualiza de una manera mucho más circular a causa de que, la información es mucho más accesible y las campañas de Marketing Digital son más baratas por medios digitales que por los tradicionales (Giuffra, 2019)

A continuación, se grafica el comportamiento de compra en forma de embudo (antes del acceso a Internet) y de manera circular (cuando los medios digitales forman parte del comportamiento del consumidor):

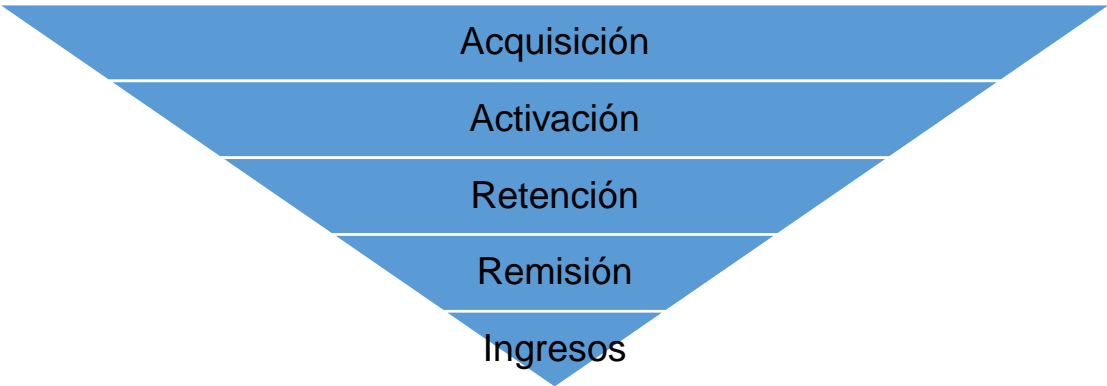


Ilustración 5.- Embudo del conversión, antes de la era digital. Funnel tradicional de Marketing (Giuffra, 2019)

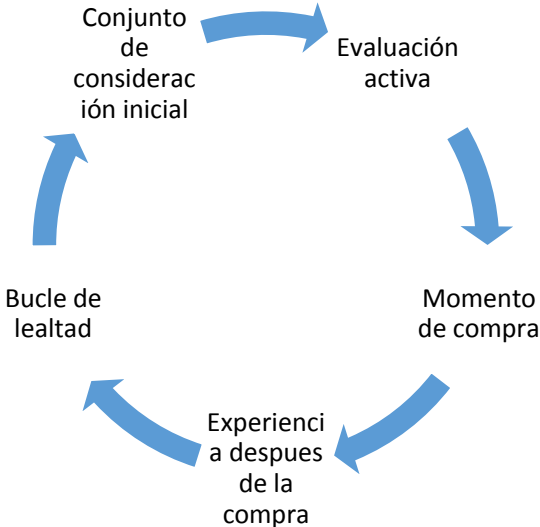


Ilustración 6.- Embudo de conversión en era digital. Funnel del Marketing Digital. (Giuffra, 2019)

El Marketing Digital tiene una mayor efectividad cuando es implementando como una estrategia multicanal es decir, como una estrategia por Internet que permite segmentar a los clientes, así como la selección y el posicionamiento de los medios para que estos puedan tener un resultado positivo.

Para que el Marketing Digital pueda ser efectivo, se necesitan ciertas características en su planificación estratégica. A continuación se exponen las más importantes según (Chaffey & Ellis-Chadwick, 2014):

1. Estar alineado con las estrategias generales del negocio, es decir tener prioridades e iniciativas anuales dependiendo del giro de la organización.
2. Tener objetivos claros para el negocio marca así como para contribución en línea de prospectos estar abiertos a la venta por Internet.
3. Ser coherentes con los tipos de clientes a los que está dirigido el producto y/o servicio.
4. Tener una propuesta de valor atractiva declara con la cual, se pueda comunicar eficazmente a los clientes.
5. Especificar la combinación de herramientas de comunicación en línea y fuera para atraer a los clientes las plataformas digitales concentración de tráfico en las mismas.
6. Contar con un apoyo para el proceso de compra en línea mismo que sea capaz problemas fácilmente.
7. Administrar el ciclo de vida de los clientes en línea a través de etapas de atracción de visitantes al sitio web conversión a clientes, para ña retención y crecimiento de los mismos.

2.3 Aplicación del Marketing Digital

Las organizaciones se enfrentan a un mundo *online* lleno de interacciones donde se ofrecen amplia gama de oportunidades para comercializar productos y servicios del ciclo de compra qué ha cambiado para los consumidores e internautas actuales.

Muchas las funciones de aplicación el Marketing Digital en la era virtual donde la conectividad y el alcance a un solo clic de distancia. Las empresas utilizan la aplicación

del Marketing en Internet como parte de sus estrategias de comunicación, estas actividades se pueden encontrar en su sitio web, sitio web de terceros y en social media:

- 1. Medio publicitario:** En un mundo de la imagen es cada vez más relevante de anuncios publicitarios es necesario considerar como un contenido estratégico para vender, mismo que conecta productos y servicios que busca un usuario pero, ¿cómo conseguir que el anuncio el correcto, delante de la persona adecuada, en un lugar específico, en el momento justo? Muchos autores expertos en publicidad digital, entre los más destacados (Rogers, 2014) menciona que en la era digital, las empresas utilizan las base de datos generada por los usuarios para generar anuncios publicitarios que llamen la atención de los usuarios Y puedan obtener el producto de una manera fácil y rápida.
- 2. Medio de respuesta directa** Las plataformas digitales permiten que las organizaciones implementen tipos de mecanismos estratégicos en los cuales, el consumidor promueve determinada acción en cierto mercado; la publicidad de respuesta trata de impulsar determinada acción por medio de una invitación u orden. Un ejemplo de este tipo respuesta directa es la que utiliza *easyJet*, esta empresa utiliza vínculos patrocinados cuando un usuario está investigando un vuelo en un motor de búsqueda para que, éste sea persuadido y visita directamente el sitio *easyJet* haciendo un clic en este vínculo.
- 3. Plataforma para transacciones de ventas:** El Avance de la tecnología permite el surgimiento de una experiencia en el mercado, con un alto valor para el cliente y a su vez, expandiendo la plaza para los ofertantes. El comercio electrónico es una oportunidad para que las organizaciones tengan una mayor ventaja competitiva y también llegar a mucho más clientes como consecuencia del acceso nuevos mercados. Este tipo de comercialización tiene muchas ventajas mismas que puede ir desde ahorros en los gastos administrativos hasta reducir exponencialmente presupuestos publicitarios.
- 4. Canal de distribución:** Las empresas han encontrado través del comercio electrónico una manera de distribución de sus productos digitales. Como por ejemplo, librerías en línea como *gandhi®* y también plataformas en donde la administración de contenido se realiza por medio de *streaming* como por ejemplo, *Netflix®* y *Spotify®*.
- 5. Mecanismo de servicio al cliente:** El servicio al cliente se ha diversificado puesto que, los usuarios y consumidores están cada vez más presentes en las diferentes

redes sociales en dónde comúnmente realizan reseñas sobre productos y servicios que utilizan es por ello, que las marcas deben estar presentes de plataformas digitales ya que, es muy importante estar donde el cliente tiene mayor presencia para realizar estrategias dirigidas al servicio y la atención del mismo.

- 6. Medio para establecer relaciones:** Puesto que las plataformas digitales tiene un gran peso para los consumidores de hoy en día, la presencia activa de las organizaciones en ellas ayuda a consolidar la marca y que la empresa esté presente en presente en la vida de sus clientes por otra parte, es muy importante que como empresa se adapte a la manera en que se comunican y hablan los usuarios en dichas plataformas así como, segmentar para conocer realmente al consumidor y de esta manera. emplear las mejores acciones en redes sociales y plataformas digitales que puedan aportar beneficios para la organización.

2.4 Beneficios del Marketing Digital

Los cambios en el comportamiento del consumidor requieren que las empresas reconsideren su comercialización conforme a las estrategias en el dominio digital. Actualmente, los beneficios que se tienen de la inmersión de estos esfuerzos en plataformas digitales se conforman a partir de implementar formas innovadoras de comunicación y co-crear contenido con sus clientes y/o consumidores (Tiago & Veríssimo, 2014).

Existen diferentes perspectivas de los beneficios y el uso del Marketing Digital por ejemplo, desde la perspectiva del consumidor, el uso de las Tecnologías de la Información y la Comunicación ofrece una serie de beneficios, entre los que se incluyen eficiencia, comodidad, información más completa y de calidad, mostrar una selección más amplia de productos, precios competitivos, reducción de costos y diversidad de productos (Bayo-Moriones & Lera-López, 2007) En cambio, las redes sociales permiten que los consumidores pueden comunicarse de forma más proactiva. Por ejemplo, a través de las redes sociales en línea, los individuos pueden buscar opiniones de otros sobre productos específicos.

Distintos autores han demostrado las ventajas del enfoque de Marketing Digital puede apoyar de distintas maneras para que las operaciones comerciales para que estás puedan ser rentables según (Ainscough, 1996), la integración de una organización a

las plataformas digitales se lleven a cabo con la finalidad de la publicación de ventas en línea, estudios de mercado y atención al cliente. Por otra parte, autores como (Chaffey & Smith, 2013) hacen notar que el Marketing Digital puede ser utilizado para la ejecución de estrategias digitales a través de la identificación, anticipación y satisfacción. Se explicaran cada una de manera detallada a continuación:

- **Identificación:** Se puede utilizar el Internet para investigaciones de Marketing a fin de conocer necesidades y deseos de los clientes.
- **Anticipación:** El Internet se comporta como un canal adicional por el cual, los clientes tienen acceso a la información, comunicación, compra y venta.
- **Satisfacción:** para que el Marketing Digital sea exitoso se debe de enfocar diferentes recursos en lograr la satisfacción del cliente ya sea a través seguimiento del mismo después de realizar una compra o proporcionarle una manera de comunicación inmediata para dudas y aclaraciones en una plataforma digital misma que pueda estar disponible para el siempre que quiera.

Lo anterior, nos lleve a comprender que tanto la organización como el consumidor tienen grandes beneficios gracias a la adaptación de los negocios a la era digital puesto que, se cuenta con más información para tomarse decisiones, se pueden contactar con proveedores en cualquier lugar y se puede acceder a los precios del mercado en tiempo real. A partir de este gran cambio y adaptación de Marketing autores como Philip Kotler, el padre del Marketing, dijo: *“El manejo inteligente de la información y el uso de interacciones con el cliente, apoyadas por tecnología, se encuentra entre las reglas básicas del Marketing electrónico para la nueva economía”* Es un hecho que dichas reglas tradicionales del Marketing siguen vigentes pero, para que el Marketing Digital sea eficaz este tiene que adaptar algunos conceptos y también entender los modelos de negocios que surgen a partir de la existencia del mismo.

2.5 E-business: Modelos de negocios digitales

Entre mediados y finales de la década de 1990 se produjo un gran aumento en el número de empresas dispuestas a aprovechar las oportunidades que el comercio a través del Internet podría brindarles. Estas empresas fueron denominadas *“dot.com”* o *“punto.com”* en su traducción al español.

Las empresas “punto.com” adoptaron numerosos tipos de modelos de negocio para ganar ingresos y lograr el crecimiento. Estas organizaciones existen porque gracias a la comercialización por Internet algunas pudieron lograr altas valoraciones basadas en el optimismo que rodeaba la economía de Internet durante ese período (Combe, 2006). Casi inevitablemente, la iniciante economía de Internet no podría apoyar la gran cantidad de empresas que fueron atraídos por las perspectivas de hacer dinero rápidamente.

Internet ha transformado el entorno competitivo para las empresas y que los superiores y gerentes en las organizaciones entiendan sus atributos distintivos en para construir modelos efectivos y explotar tantas oportunidades digitales como sean posibles. La economía de Internet funciona por un conjunto diferente de reglas y las organizaciones acostumbradas a los modelos de negocio tradicionales se enfrentaron al desafío de transformar su modelo de negocio o construirlo con base a los atributos de Internet.

El autor (Lee, 2001) propuso en su artículo de investigación : *An analytical framework for evaluating e-commerce business models and strategies*, describe los principales atributos que el Internet representa tanto un reto, como una oportunidad para las empresas. Este debe incluir:

- **Economía del intercambio de información:** Las empresas pueden aprovechar el alcance de la información usando el Internet. El costo de utilizar el Internet como medio informativo es prácticamente cero y el alcance es global.
- **Conectividad e interactividad:** El número de personas que tienen acceso a Internet es cada año mayor cada año a escala global. La conectividad amplía el alcance de la tecnología. La comunicación y la interactividad permite la comunicación bidireccional en tiempo real y cualquier parte del mundo.
- **Red de economías de escala:** Los efectos de red son más fuertes en la economía de Internet. Existen oportunidades para lograr alcanzar objetivos de venta y consumo mediante el acceso a Internet. Las organizaciones pueden tener una base de clientes más amplia electrónicamente a bajo costo. Los costos marginales del envío de información es casi nulo para que las empresas puedan lograr economías. de escala al proporcionar productos o servicios de valor agregado a los clientes de manera eficaz, rápida y eficiente a través de Internet.
- **Velocidad de cambio:** Internet ha acelerado el proceso de transacciones y ha aumentado las expectativas de los clientes. Las empresas necesitan reajustar sus plazos de entrega, los tiempos de respuesta y los tiempos de distribución y

entrega para cumplir las exigentes demandas de los clientes, proveedores y socios en toda la cadena de suministro en la economía de Internet. Las empresas que son capaces de cumplir o superar estas elevadas expectativas y puede crear una lealtad a la marca como ventaja competitiva ante sus competidores.

- **Economía abundante:** La información es un activo valioso en la economía de Internet. Los ingresos pueden ser creados por la difusión de información basada en su valor para los consumidores. Empresas con derechos exclusivos de la propiedad conforme a la información pueden obtener valiosas entradas de dinero puesto que, tanto la reproducción como protección de la información (*copyright*) que se genere a través de Internet se puede traducir como grandes ingresos para las organizaciones. Sin embargo, gran parte de la información que impulsa el la economía de Internet está ampliamente disponible y es omnipresente. Como se ha señalado previamente, la información puede ser reproducida y distribuida a bajo o nulo costo lo cual, permite que gran cantidad de información pierda su valor. Estas se denominan "deseconomías de escala" y puede ocurrir cuando el consumidor se ve tan abrumado por el gran volumen de información.
- **Intercambio de mercancías:** El Internet proporciona un mecanismo para mostrar una gran variedad de productos y servicios sin tener que incurrir en los costos de visualización y presencia que en las tiendas tradicionales. Las facilidades de búsqueda pueden canalizar a los consumidores a los productos y servicios en los que están interesados y el sitio web puede ofrecer servicios adicionales tales como servicios de descuento, entre otras promociones y enlaces a productos complementarios, transacciones, pagos y entrega. En comparación con las formas tradicionales de compras, Internet ofrece una mayor comodidad a un costo menor y potencialmente mejor conforme a la prestación de servicios.
- **Medio de comunicación:** El Internet puede ser usado como un medio de comunicación con clientes y enriqueciendo la relación entre comprador y vendedor. Esto puede llevar a personalizar productos o servicios. El Internet facilita la capacidad de los clientes para determinar el diseño, desarrollo y producción de productos y servicios
- **Contexto industrial:** En la economía tradicional el valor se crea dentro de los sectores industriales, tales como la fabricación o venta al por menor. En la economía basada en Internet el valor es generado en las comunidades en línea misma que trascienden sectores industriales.

Por otra parte, los modelos adoptados por las empresas están diseñados para explotar las ventajas que el uso de Internet puede traer. Existen tres tipos de canales. A través de la cual se desarrolla la actividad de *e-business*. Estos son los canales de comunicaciones, transacciones y distribución (Combe, 2006):

a) Canal de comunicaciones: El intercambio eficiente de información entre compradores y vendedores es la ventaja más importante que ofrece Internet. A través de este medio se demostrado ser un medio eficaz para acceder, organizar y comunicar. Este canal incluye:

- La capacidad de almacenar grandes cantidades de información;
- la facilidad de acceso a la información;
- el alcance de la interactividad;
- la capacidad de proporcionar información bajo demanda;
- la presentación visual mejorada de la materia.

El Internet proporciona información casi instantánea y está disponible durante todo el día en cualquier lugar. Los clientes se benefician de tiempos de respuesta más rápidos y acceso a una gama más amplia de servicios proporcionados por las empresas a través de su página web. Esto, a su vez, fomenta la interactividad entre clientes y vendedores las diferentes formas de comunicación que incluyen correo electrónico, listas de correo, grupos de noticias, salas de chat y redes sociales. Los vendedores también pueden ajustar su estrategia de Marketing en respuesta a los comentarios de los clientes. Esta retroalimentación también ayuda a los vendedores en el desarrollo de nuevos productos, desarrollando la aplicación de productos existentes u ofreciendo productos personalizados. El proceso de construcción de relaciones se ve impulsado y los vendedores pueden beneficiarse de la lealtad del cliente.

b) Canal de transacciones: El Internet también beneficia el proceso de transacciones entre compradores y vendedores, aumenta la eficiencia y puede reducir los costos. Los costos de transacción para los clientes se reducen debido a la facilidad de acceso a información de mercado sobre una amplia gama de productos incluyendo, precio, calidad, disponibilidad y descuentos.

c) Canal de distribución: Las empresas que producen productos digitales pueden usar Internet como canal de distribución para ofrecer suministro instantáneo a una reducción significativa de costo de envío.

Modelos de negocio a negocio y negocio a consumidor

Una vez que se comprendieron de manera detallada la manera en que el Internet llegó a modificar la manera en que las organizaciones han modificado y adaptado sus actividades comerciales es necesario mencionar un aspecto fundamental del modelo de negocios es si la propuesta desarrollada atrae a los consumidores o a las empresas. Por ello, las organizaciones de Marketing por Internet se describen en términos del grado en que una organización realiza las transacciones con consumidores (negocio a consumidor, B2C) o con otras empresas (negocio a negocio, B2B) (Chaffey & Ellis-Chadwick, 2014).

A continuación, se muestra una tabla donde se muestra la función de los modelos de negocio conforme a la adaptación en línea:

	Consumidor	Organización	Gobierno
Consumidor	C2C: Consumidor a consumidor: <ul style="list-style-type: none"> Ebay Igual a igual (Skype) Blogs Redes sociales: Facebook, Instagram, Twitter. Recomendaciones de productos 	B2C: Negocio a Consumidor. <ul style="list-style-type: none"> Transacción: Amazon Creación de las relaciones: Business Partner Creación de marca Propietario de medios: News Corp, Merca 2.0 Intermediarios de comparaciones: TripAdvisor 	G2C: Gobierno a Consumidor: <ul style="list-style-type: none"> Transacciones de gobierno nacional. Información de gobierno nacional. Información del gobierno local Servicios del gobierno local.
Organización	C2B: Consumidor a negocio <ul style="list-style-type: none"> Retroalimentación del consumidor, comunidades y compañías 	B2B: Negocio a negocio <ul style="list-style-type: none"> Transacciones Creación de relaciones Propietario de medios Producciones de negocios Mercados 	G2B: Gobierno a negocios y transacciones gubernamentales: <ul style="list-style-type: none"> Servicios y transacciones gubernamentales: impuestos Normales legales
Gobierno	C2G: Consumidor a Gobierno <ul style="list-style-type: none"> Retroalimentación al gobierno a través de grupos de presión o sitios de individuos 	B2G: Negocio a Gobierno <ul style="list-style-type: none"> Retroalimentación a empresas gubernamentales y organizaciones no gubernamentales 	G2G: Gobierno a Gobierno <ul style="list-style-type: none"> Servicios entre gobiernos Intercambio de información

Tabla 2.- Modelos del negocio enfocados al e-business (Chaffey & Ellis-Chadwick, 2014)

En la tabla anterior se sugieren que las empresas en línea están enfocadas en los mercados negocio a consumidor (B2C). Sin embargo, las comunicaciones negocio a negocio siguen siendo importantes para muchas de estas empresas ya que pueden ocurrir transacciones comerciales, como por ejemplo con eBay Business, o podría ser necesario mantener el servicio negocio a consumidor (B2C) con la publicidad proporcionada por las transacciones: negocio a negocio (B2B); por ejemplo, los ingresos de Google se basan primordialmente en sus AdWords. Por otra parte, el servicio de publicidad también son importantes para sitios como YouTube, Facebook e Instagram y están integrados a modelo B2B (Chaffey & Ellis-Chadwick, 2014).

Diferencia entre comercio electrónico y negocio electrónico

En contraste con lo anterior es importante definir los siguientes conceptos: comercio electrónico (e-commerce) y negocio electrónico (e-business), según el autor (Chaffey, 2011) el comercio electrónico se refiere a las transacciones financieras y de información utilizadas de manera electrónica entre una organización y cualquier tercero con el que tenga tratos de ventas en línea, sino también de las operaciones no financieras, como consultas recibidas en el área de servicio al cliente y envío de correo electrónico, es por ello, que el comercio electrónico está abierto a todas las organizaciones en línea.

El comercio electrónico suele subdividirse en una perspectiva de comercio electrónico de lado de la venta, el cual, se refiere a las transacciones para vender productos a los clientes de una organización, y una perspectiva de comercio electrónico del lado de la compra, la cual se refiere a las transacciones de negocio a negocio para que una organización adquiera de sus proveedores los recursos que necesita. En cambio, el comercio social es una parte cada vez más importante del comercio electrónico para los propietarios de sitios, ya que incorporar reseñas y clasificados en un sitio y vincularse a sitios de redes sociales puede ayudar a comprender las necesidades de los clientes y aumentar la conversión a ventas.

Por otra parte, el negocio electrónico (*e-business*) es parecido al comercio electrónico pero de alcance más amplio y se refiere al uso de la tecnología digital para administrar una gran variedad de procesos de negocios que incorporan el comercio electrónico del lado de la venta y de la compra que se muestran en la siguiente ilustración número 7:

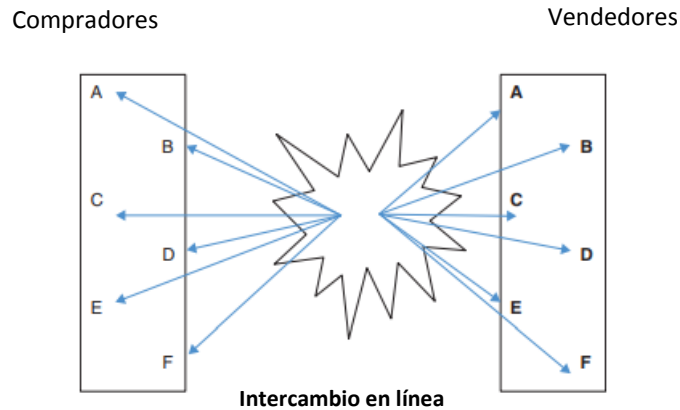


Ilustración 7.- Intercambio en línea según el autor (Combe, 2006)

2.6 Plan de Marketing Digital

El desarrollo y uso de Internet a escala global ha creado oportunidades para consumidores, proveedores y vendedores de productos y servicios para comunicarse entre sí en el mercado en línea. Internet ha sido el catalizador para que las organizaciones extiendan sus actividades más allá de los límites tradicionales que se caracterizaron por limitaciones físicas. Una de estas actividades es el Marketing mismo que funciona dentro de las organizaciones como medio de comunicación, el Internet ofrece una gama de beneficios que pueden ser explotados por las organizaciones para mejorar su estrategia de Marketing (Combe, 2006).

El Internet se caracteriza por la ubicuidad, riqueza y densidad de la información, así como por su alcance, interactividad, personalización. Es importante destacar que Internet une a compradores y vendedores de una manera más rentable que los métodos tradicionales de publicidad y Marketing. A su vez, el propósito del Marketing es identificar, anticipar y satisfacer los requerimientos de clientes rentables. Las variables clave del mercado se deben de considerar al implementar una estrategia de Marketing: producto, precio, plaza y promoción.

Un factor crítico de éxito para las empresas que implementan una estrategia de Marketing es entender las características de los clientes a los que se dirige pero de manera general, el esfuerzo de Marketing está dirigido a uno o más segmentos de clientes. Segmentos de

clientes pueden ser determinados por un rango de criterios incluyendo la edad; género; grupos socio económicos; geografía; empleo; gustos e intereses; factores culturales y estilos de vida.

Internet es un mecanismo adicional disponible para las empresas que facilita la recepción y el envío de información y puede ser utilizado para informar su estrategia de Marketing Digital.

Según el autor (Combe, 2006) el Marketing Digital usa la tecnología de las comunicaciones electrónicas, como Internet para lograr los objetivos de Marketing. Por supuesto, la información en los sitios web. También, empodera a los clientes y les ayuda a lograr a satisfacer sus necesidades.

El Internet ha cambiado la relación entre compradores y vendedores porque la información de mercado está disponible para todas las partes en la transacción. El plan de Marketing Digital es parte de un plan de mayor alcance puesto en marcha por las empresas para lograr los objetivos establecidos. Realizar Marketing Digital forma parte de la estrategia de *e-business* para lograr objetivos, este incluye estudios de mercado y comunicaciones con clientes utilizando Internet. El plan de Marketing Digital será dependiente del nivel de recursos disponible para la organización.

El plan se construirá dentro del presupuesto asignado y oscilará en alcance de acuerdo a los beneficios percibidos que puede acumular para el negocio. El plan de e-marketing está diseñado para identificar y establecer maneras de lograr objetivos de e-marketing. El plan consiste en análisis de la función de la demanda y de los competidores. Estos forman la base para desarrollar e implementar estrategias para mercados objetivo. La mezcla de Marketing es la táctica utilizada para dirigirse a los mercados. Finalmente, hay un elemento de medición del rendimiento y evaluación del mismo. Este modelo de Marketing Digital se expone en la siguiente ilustración fue propuesto por (Combe, 2006):



Ilustración 8.- Modelo del plan de Marketing Digital según (Combe, 2006)

2.6.1 El entorno del Marketing Digital

Un paso muy importante en la implementación de un plan de Marketing Digital es comprender las diferencias clave entre el Marketing Tradicional y el Marketing Digital. Un plan Marketing Digital efectivo se basa en nuevas perspectivas e interacciones de los clientes y a su vez, ofrecer diferentes soluciones para la comercialización de productos y servicios.

Antes de emprender un plan de e-Marketing es necesario abordar algunas cuestiones clave, según (Combe, 2006): Éstas incluyen:

- ¿En qué se diferencia la función de Marketing en el nuevo entorno?
 - a. No hay restricciones geográficas en la línea.
 - b. Los modelos de Marketing electrónico están sujetos a cambios dependiendo del nivel de avance tecnológico. El Marketing tradicional es relativamente estable y más claramente definido;
 - c. Los nichos de mercado se identifican más fácilmente y los mercados son más grandes en línea.

d. El Marketing electrónico está sujeto a una relación de atracción del cliente donde el cliente decide el momento y el lugar de interacción.

- ¿Cuáles son los factores críticos de éxito para el Marketing Digital?

a) Las estrategias iniciales son vitales;

b) Colaboración con socios para aportar valor agregado a los clientes;

c) Los clientes deben percibir valor agregado;

d) Innovación en productos y servicios, por ej. Personalización;

e) Maximizar la interactividad del cliente.

- ¿Cuál es la nueva posición organizacional para la función de Marketing?

a) El desarrollo del producto está impulsado por el Marketing y los sistemas de información;

b) La función de Marketing está integrada con otras funciones tales como TI y desarrollo de negocios;

c) ¿cómo se perciben los clientes dentro del nuevo ambiente comercial?

d) Los clientes tienen mayor conciencia de la organización y sus productos y servicios;

e) El comportamiento del cliente cambia en línea, por ejemplo: Mayor impulso comprando. Esto puede ser monitoreado y usado para fines de Marketing.

f) Los clientes son percibidos más como individuos en lugar de grupos.

2.6.2 Análisis de demanda

Una vez que los negocios electrónicos entienden las diferencias clave entre los tradicionales, el Marketing Tradicional y el Marketing Digital son necesarios para construir un plan de Marketing Digital. Un plan de Marketing eficaz requiere un entendimiento de la demanda dinámica del mercado. Esto puede implicar el análisis la demanda actual y proyectando la demanda futura de todo el mercado o restringiendo el alcance del análisis para adquirir conocimiento de la demanda para segmentos específicos del mercado. El análisis de la demanda se centra en una serie de cuestiones clave, incluyendo:

- El nivel de conectividad a Internet de los clientes;
- El nivel de interactividad de las personas con acceso a Internet;

- El número de clientes que compran productos o servicios a través de Internet;
- El número de clientes que acceden a sitios web pero que no compran a través de Internet;
- Identificar las barreras que puede tener la compra por Internet.

La segmentación de clientes ayuda a comprender mejor los diferentes tipos. Algunos productos van dirigidos a conjuntos de clientes, tales como tarifas aéreas de descuento. Sin embargo, muchos productos y servicios están dirigidos a mercados con características específicas.

Los productos o servicios pueden ser diseñados o adaptados para adaptarse a mercados objetivos particulares. Los segmentos del mercado incluyen:

- Grupos socioeconómicos: estos dividen a la población y consumidores en seis grupos amplios como se describe en productos y servicios basados en la información.
- Ubicación geográfica: ayuda a dirigir productos y servicios según tipos de localización donde los consumidores residen. La información importante incluye si los consumidores residen en áreas que son urbanas o rurales, tienen alta o baja densidad de población, o si tienen acceso a infraestructura y puntos de distribución.
- Género: productos y servicios pueden ser dirigidos a clientes particulares basados en género. Esta información esta enlaza a otros datos clave sobre el comportamiento del cliente, como el tipo producto, uso, fidelidad, patrones de compra y los tipos de beneficios buscados, y variables psicográficas como el estilo de vida, aptitudes y actitudes, rasgos de personalidad, intereses y actividades.
- Edad: el perfil de edad de los clientes o grupos de clientes es importante para el e-Marketing. Esto da lugar a segmentos de mercado tales como niños, adultos jóvenes, de mediana edad, personas jubiladas, etc.
- Gustos y modas: los intereses conforme a gustos y modas es expresado por clientes particulares, o grupos de clientes miso que, proporcionan información valiosa para la planeación de Marketing Digital de los negocios electrónicos.
- Sectores empresariales: en el sector B2B los clientes se dividen por sectores como servicios financieros, comercio minorista, transporte, salud, sector público, ocio, etc. B2B implica la creación de un portal web para comunicarse con

compradores y proveedores, que sustenta alianzas en el proceso de compra y logística. Los negocios electrónicos exploran el mercado para vender productos o servicios a otras empresas. Las empresas pueden especializarse en el suministro de productos para reforzar su cadena de suministro.

- Demografía: los clientes se pueden dividir en una mezcla de criterios geográficos, de edad y género.

2.6.3 Análisis del competidor

Realizar un análisis de los competidores es una parte integral del desarrollo estratégico para las empresas tanto tradicionales como electrónicas. A veces denominado "inteligencia de la competencia", el análisis implica reunir información sobre una amplia gama de criterios, incluyendo quiénes son los competidores, cuáles son sus productos y servicios que ofrecen, su desempeño en el mercado, los canales de comunicación utilizados y sus estrategias de Marketing. El análisis de la competencia ayuda a determinar el impacto que las acciones del competidor tendrán conforme al rendimiento de un negocio.

Gran parte de la información requerida para el análisis de la competencia se realiza a partir del conocimiento de proveedores y clientes, así como la observación tecnológica de la industria por los empleados. Otras fuentes de información pueden incluir material promocional de la competencia, anuncios, informes sobre la industria, informes del gobierno, comunicados de prensa y estudios académicos. Las empresas a menudo compran productos de los rivales para analizarlos conforme a efectividad y para buscar formas de mejorar adaptándolos como medio para crear una ventaja competitiva a través de la diferenciación y la innovación. Los sitios web de las empresas rivales son también una valiosa fuente de información. Estos muy a menudo proporcionan información sobre nuevos productos, precios, disponibilidad, calidad, descuentos, promociones y clientes.

2.6.4 Fijación de objetivos

El plan de e-marketing se basa en objetivos claramente definidos. Estos objetivos pueden incluir aumentar las ventas, establecer nuevos productos en el mercado, ampliando la

cartera de clientes, mejorar el servicio al cliente, diversificando la gama de productos, reducción de costos, o cualquiera de los otros beneficios asociados con la empresa

Los objetivos son elegidos, es importante asegurar que sean alcanzables. Muchas empresas han fracasado en su estrategia de *e-business* porque no entienden las capacidades y limitaciones de la tecnología al establecer objetivos (Combe, 2006). La tecnología no puede lograr objetivos en aislamiento de otras funciones dentro de una organización. Los objetivos también deben ser medibles conforme al plan de Marketing Digital. Algunos objetivos, como mejorar la reputación de la empresa puede llevar algún tiempo lograrlo e incluso entonces puede ser difíciles de medir.

Los objetivos de Marketing Digital variarán entre organizaciones. Las empresas que existen solo debido a que por medio del Internet obtienen todos sus ingresos de sus actividades en línea, mientras las empresas tradicionales utilizan Internet como un servicio adicional para los clientes. Las empresas "nacidas en la red" probablemente tengan objetivos más precisos para su estrategia de Marketing Digital, ya que constituye un papel vital en la determinación de la viabilidad de la organización. Las empresas tradicionales invariablemente ven la inversión en servicios en línea como fuente de ingresos adicionales y un medio para mejorar la relación con el cliente. La viabilidad de la organización no depende totalmente del elemento *e-business* de sus operaciones.

2.6.5 Auditoria del plan de Marketing Digital

La auditoría de Marketing Digital debe estar diseñada para revisar el desempeño del plan con el cuál, la organización está inmersa en el mundo digital. En particular, la auditoría de Marketing Digital evalúa el desempeño de los factores integrados en el establecimiento de objetivos.

La efectividad del sitio web de la compañía es una parte clave de la auditoría de Marketing Digital. El análisis del rendimiento del sitio web se centra en tres áreas (Combe, 2006). Éstos incluyen:

- Contribución a la productividad empresarial: el sitio web tiene que contribuir al desempeño económico de la empresa en forma de ingresos adicionales y rentabilidad.
- Eficacia de Marketing: el sitio web tiene que contribuir a las ventas y estar sujetas al seguimiento de medidas de efectividad en atraer, retener y retroalimentar. Los

comentarios de clientes sobre la calidad del servicio y las tasas de satisfacción también figuran en las medidas de efectividad de Marketing.

- Eficacia del usuario: el sitio web debe ser prominente, atractivo y fácil de usar. La efectividad del usuario puede ser medida en el número de visitas que el sitio web recibe sobre un período determinado de tiempo, las páginas particulares del sitio web que se accede, el tiempo que los usuarios permanecen en el sitio web y resultados conforme a compra de productos y servicios. La investigación de mercado en forma de cuestionarios, encuestas y *focus groups* pueden proporcionar información valiosa sobre las percepciones de los usuarios, de la experiencia del sitio web en términos de accesibilidad, navegación, diseño, calidad de la información y facilidad de uso.

2.6.6 Construcción de marca

Una marca es el nombre, término o símbolo dado a un producto para distinguir la oferta de productos similares suministrados por competidores. El *branding* se refiere al proceso de creación y desarrollo de marcas que agregan valor a los consumidores. Las empresas necesitan desarrollar productos o servicios que se diferencian de los rivales. Elegir un nombre para la empresa es un paso importante en el esfuerzo de marca global.

La marca elegida debe:

- Identificar los aspectos únicos de la empresa y sus productos;
- Comunicar los fines y objetivos de la empresa a una audiencia;
- Ayudar a proteger la propiedad intelectual de la empresa.

El sector de *e-business* está poblado por un gran número de organizaciones. Debido a los pequeños costos que implican realizar un negocio en medios digitales y la facilidad de entrada.

Hay tres tipos principales de marcas:

- Independiente: el nombre está diseñado para ser único y memorable;
- Asociativa: el nombre se elige para reflejar similitud conforme a calidad o es una abreviatura del nombre completo de la empresa;
- Descriptiva: describe el nombre del producto o servicio o el proceso productivo.

La gestión eficaz de la marca a través de Internet requiere estrategias de Marketing y promoción muy diferentes a las tradicionales. Una forma de promocionar marcas es participar en un diálogo con consumidores sobre productos o servicios, lo anterior, se conoce como Marketing de uno a uno. Esto permite a las empresas construyan relaciones con sus clientes y obtienen información sobre sus deseos y necesidades. Las empresas pueden desarrollar o adaptar productos para satisfacer esas necesidades. La personalización de productos y servicios agregan valor a la marca. Esto puede ser la base para la fidelización de los clientes. Hay evidencia que sugiere que los consumidores se familiarizan cada vez más con el uso de Internet y solo comprará productos o servicios que se ajusten exactamente a sus necesidades. Para cumplir con estas expectativas las empresas necesitan adoptar estrategias de gestión de marca que van más allá de las características del producto pero también incluyen criterios como valores, creencias, emociones, recuerdos y actitudes.

2.6.7 Marketing Mix

Otros autores como (Chaffey & Smith, 2013) mencionan que una estrategia de Marketing bien planificada debe estar basada una reformulación de la mezcla de mercadotecnia en función a este nuevo mundo digital. Dicha mezcla está basada en siete P's en lugar de cuatro, las últimas tres son las más relevantes, tanto en el mundo digital como en la mercadotecnia B2B y de servicios. La presente propuesta de Marketing Digital estará expuesta a continuación:

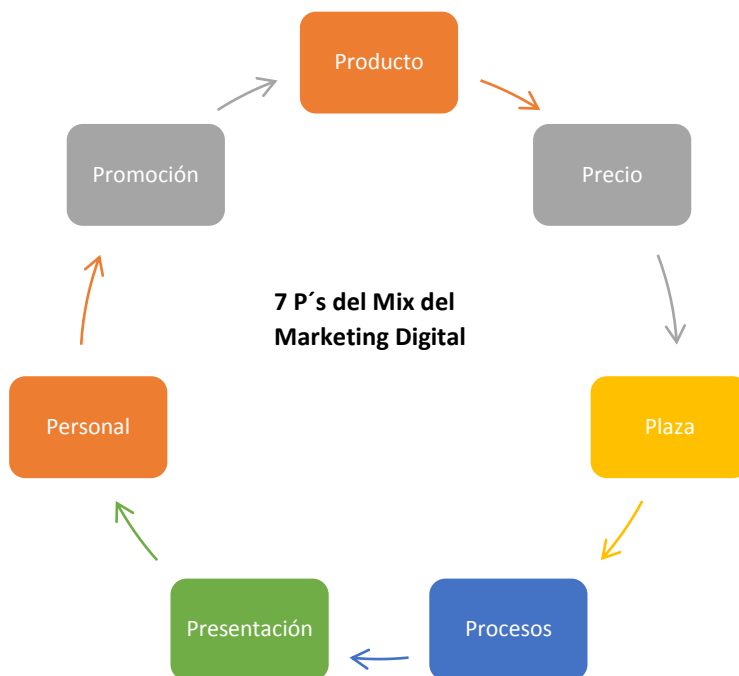


Ilustración 9.- 7 P's del Mix del Marketing Digital según (Chaffey & Smith, 2013)

a) Producto: es todo aquello que la empresa o la organización realiza o fabrica para ofrecer al mercado y satisfacer determinadas necesidades de los consumidores (Psicología y Empresa, 2010). No obstante, al referirnos a un producto, éste no se limita a un bien tangible; sino que abarca también una idea, una organización o un servicio.

La versión digital de un producto pretende extender la capacidad informativa y persuasiva del mismo, a través de nuevos y diferentes atributos de los ya conocidos en el mundo físico (empaques, presentaciones, variedad, sabores, etc.) (Chaffey & Smith, 2013).

b) Plaza: Se trata de todo el esfuerzo que realiza la organización social para poner a disposición de los usuarios los productos sociales, como los lugares físicos que se tienen que establecer o la elaboración de alianzas estratégicas con organizaciones públicas y privadas de la comunidad para que los productos sociales lleguen a través de sus estructuras al usuario final. La plaza es el lugar tangible en donde se ofrece un producto al consumidor; es decir, implica los diferentes canales de distribución. Dentro de la estrategia de Marketing, se debe privilegiar cada canal de distribución, siempre facilitando que el producto en cuestión sea lo más accesible que se pueda, al consumidor.

c) Precio: Se puede entender que el precio es un equilibrio entre la oferta y demanda de un producto y servicio o en su defecto, la transacción casi inmediata sucedida al momento de consumir un producto y servicio ofertado. El concepto del precio en el mundo digital es transparente y democrático en su totalidad puesto que, las herramientas en línea para comparar precios y ofertas genera un entorno mucho más agresivo conforme a competencia.

d) Procesos: Se refiere a los procedimientos que debe realizar el usuario o servicio. De parte de la institución o marca, se trata de que la entrega de productos y servicios se lleve a cabo de una manera fácil y rápida para el usuario final.

e) Presentación: a presentación o presencia física "se refiere a la fachada de los bienes inmuebles en donde se ofrece el producto social, la apariencia de los espacios exteriores e interiores de la organización prestadora del producto social" (Pérez Romero, Marketing Social, Teoría y Práctica., 2004). En este sentido, es preciso prestar atención en cada factor inclusivo en el bien inmueble, pues, conforma toda una experiencia, sensación u opinión del cliente con respecto a la organización (Figueroa Rojas, 2015).

En la mezcla digital, prácticamente cualquier persona puede tener un sitio web y que este sea una tienda virtual. Lo que es fácil de validar en el mundo físico es muy complicado en el mundo digital. Es muy importante tener la evidencia física ante cualquier sitio web, tienda en línea o perfil de redes sociales. La calidad, el diseño, los certificados, avales e interacción con los clientes, representan la tan importante evidencia física (Chaffey & Smith, 2013).

f) Personal: El personal es la imagen de toda empresa; se trata de todas las personas que se encargan de facilitar los servicios al cliente.(Figueroa Rojas, 2015). En pocas palabras, el personal "es el elemento humano de toda organización (Romero Servín, 2007). El funcionamiento exitoso de los nuevos medios digitales depende de las personas, quienes constituyen los llamados medios ganados. Otros recursos como el "*Crowdsourcing*" se detonan con apoyo voluntario de los consumidores. La motivación, la comunicación clara y los incentivos adecuados forman parte integral de esta nueva mezcla de mercadotecnia (Chaffey & Smith, 2013).

g) Promoción: La promoción se define como la función de Marketing relacionada con la comunicación persuasiva, hacia audiencias objetivo, de los componentes del

programa de Marketing para facilitar el intercambio entre el fabricante y el consumidor, y ayudar a satisfacer los objetivos de ambos (Burnett, 2012).

En el marco digital, promoción se hace vasta en alternativas para generar curiosidad, expectativas, activar clientes, vender y generar lealtad a lo largo de los medios sociales y digitales existentes. Las posibilidades se multiplican, con el uso de geo localización, programas de lealtad, cupones e mecánicas. También la publicidad digital, tanto en sitios web como en motores de búsqueda (Chaffey & Smith, 2013).

Por otra parte, es muy importante tomar en cuenta los desafíos del desarrollo y la administración de la estrategia de Marketing Digital, estos están inmersos bajo el marco de la administración de la estrategia de Marketing por Internet que se puede observar en muchas organizaciones (y que se deben manejar de manera inmediata), se exponen a continuación (Chaffey & Ellis-Chadwick, 2014):

- Responsabilidades poco claras para las actividades de Marketing por Internet.
- Se establecen objetivos no específicos para el Marketing por Internet.
- Se asigna un presupuesto insuficiente para el Marketing por Internet porque se subestima la demanda de servicios en línea por parte de los clientes y los competidores ganan participación en el mercado a través de actividades en línea superiores.
- El presupuesto se desperdicia cuando las diferentes partes de una organización experimentan utilizando diferentes herramientas o proveedores sin resultados óptimos como por ejemplo, economías a escala.
- Las nuevas propuestas de valor en línea no se ha desarrollado porque se trata a Internet como “otro canal para comercializar”, sin revisar las oportunidades para ofrecer servicios en línea mejorados y diferenciados.
- Los resultados del Marketing Digital no se miden, ni se revisan adecuadamente, así que no se puede realizar acción alguna para aumentar la efectividad.
- Se adopta un enfoque experimental en el lugar de un planificado para utilizar las comunicaciones electrónicas con poca integración en las comunicaciones de Marketing en línea y fuera de línea.

2.6.8 Modelo para la planificación de Marketing Digital: SOSTAC ®

Existen diferentes metodologías para realizar un plan de Marketing, el autor (Tabuenca, 2018) afirma que hay una muy completa que puede utilizarse para casi todo tipo de negocio y aporta buenos resultados a la organización. La metodología SOSTAC® fue desarrollada por PR Smith en 1990 que puede ser utilizada tanto para multinacionales como en *start-ups*.

El acrónimo de la metodología proviene de las seis etapas de la misma: **S**ituation, **O**bjectives, **S**trategy, **T**actic, **A**ction y **C**ontrol y que, conforme a su traducción en español es: Análisis de Situación, Objetivos, Estrategia, Tácticas, Acciones y Control.



Ilustración 10.- SOSTAC®: etapas del modelo. SOSTAC® es una marca registrada de PR Smith (www.prsmith.org)

1. Análisis de la situación

Toda estrategia debe partir de un análisis de situación actual. Se debemos hacer un estudio profundo del negocio ya que, así se tendrá una visión global de lo que se tiene internamente y a su vez, del entorno competitivo.

La organización de hacerse las siguientes preguntas:

- ¿quiénes somos?,
- ¿qué hacemos?,
- ¿cómo es nuestro proceso comercial?

Son solo algunas de las preguntas que este análisis debe incluir.

Algunas herramientas para hacer este análisis son: DAFO, análisis de la competencia (*benchmark*), análisis de los canales digitales, etc.

2. Objetivos

Una vez que la situación de la organización en la que se encuentra la organización es clarificada se debe de seguir la acción más importante de la metodología: Definir los objetivos estratégicos.

Es importante mencionar que los objetivos no solo deben estar enfocados en ventas y clientes potenciales. Se recomienda ampliamente que los objetivos incluyan cada una de 5's del Marketing Digital:

Vender (Sell): La razón para emprender actividades de Marketing Digital es llamar la atención sobre la marca y, en última instancia, atraer el crecimiento. Ya sea a través de la venta convencional de un producto o servicio, la venta de una membresía, la suscripción o la obtención de una donación. El objetivo es el crecimiento, por lo que sea cual sea el crecimiento de la organización, se debe mantener como punto de referencia para determinar si las actividades de Marketing Digital logran tener éxito o no.

Aporta valor (Serve): El sitio web y las redes sociales son los lugares en los que las personas buscan tener una conexión con la organización y buscan ayuda, por lo que existen para servir y agregar valor a la audiencia.

Para las organizaciones que operan de manera comercial, el sitio web consiste en permitir que los clientes encuentren productos e información relevante para realizar sus compras de forma rápida y sencilla.

En el sentido más básico, las organizaciones de membresía existen para servir a una comunidad al proporcionar un proceso de unión sencillo que proporciona acceso a los valores de la membresía. Estos valores incluyen a menudo contenido y servicios relevantes y de calidad que los harán regresar y renovarse con facilidad año tras año.

Proporcionar información de manera pública a una organización desempeña un papel importante para ayudar a ganar y mantener la confianza

Opinion (Speak): Cuando una organización se acerca a los clientes haciéndoles un seguimiento, haciéndoles preguntas, se preocupa por creación de un diálogo, foros de seguimiento, grupos, blogs y tweets, la mayoría de tiende a responder el diálogo de forma estructurada permitiendo recolectar ideas de clientes que puedan ser utilizadas anteriormente.

Ahorrar (Save): Gracias a los programas de automatización, CRM's, chatbot, gestores de autoadministración, etc. Usar los recursos digitales posibles y fáciles de gestionar ahorra tiempo y recursos financieros.

Amplifica (Sizzle): *Sizzle* se refiere a todo lo que haga la organización en línea debe estar enfocado a aporta a la marca. Como por ejemplo introducir nuevos elementos o vías de comunicación. La idea es aportar valor a cada cosa que se añada al producto de una manera pensada y que sea positivo para la experiencia del usuario.

3. Estrategias

Cuando se definen los objetivos del plan de Marketing Digital el siguiente paso es definir cómo se va a lograr, a partir de esto se conoce como plan estratégico.

Es importante haber hecho bien el análisis de situación del negocio porque este ayudará a hacer estrategias coherentes y alcanzables según el estado de nuestra empresa. Cada objetivo debe de tener al menos 1 estrategia. Desarrollar una tabla en la que concrete la estrategia para alcanzar cada objetivo es una buena manera de organizarlo.

4. Tácticas

Las tácticas son las acciones de cómo alcanzaran lo propuesto en las estrategias. Para el Marketing Digital se han agrupado las tácticas en 10 grupos se podrían catalogarse como los más relevantes pero que según el tipo de negocio pueden ser menos o incluso más (Tabuenca, 2018):

- Social Media
- PPC
- SEO
- Email Marketing
- Inbpund Marketing

- Content Marketing
- WEB
- Display/RTB
- Mobile
- RRPP

5. Acciones

Esta quinta fase se centra en el plan de acción. Se debe tener definido un buen plan de ejecución con tiempos iniciales, acción y termino a su vez, los responsables de llevar cada acción a cabo y los reportes necesarios para evaluar los resultados. Existen herramientas de coordinación de equipos como Trello, Zoho Project, etc. que facilitan este tipo de gestión.

6. Control

Última fase pero no menos importante. En esta etapa se analiza si todos los objetivos planteados están cumpliéndose, ver si las estrategias y tácticas definidas fueron las acertadas y si deben o no cambiarse para el próximo plan a desarrollarse.

Se puede establecer controles trimestrales o depende el negocio hasta mensuales gracias a los medios digitales que permiten ir verificando en directo los resultados de tus acciones y tomar decisiones correctivas rápidamente.

2.7 Conclusión

Como se puede comprender, el Marketing evoluciona conforme a las necesidades del cliente y en los últimos años la convergencia tecnológica entre el Marketing Tradicional y el Marketing Digital ha encauzado una transición historia.

El Marketing Digital también conocido como mercadotecnia en línea y e-marketing nace a partir del uso del Internet y plataformas digitales con la finalidad de promover la comercialización de diferentes productos y servicios. (Zurita Guerrero, 2017) explica que el Marketing Digital reside principalmente en la aplicación de tecnologías digitales para impulsar diferentes actividades de Marketing encaminadas a conseguir mayor rentabilidad

y fidelización de clientes, a través de estrategias enfocadas en tecnologías digitales y del perfeccionamiento de un enfoque planeado, con el fin de optimar el discernimiento del usuario, entregándole comunicación conectada definida, así como servicios en línea que concuerden con sus necesidades individuales (Thomas J. Russell & W. Ronald Lane, 2005)

Según los profesionales del Marketing deben adoptar el cambio hacia un panorama empresarial más horizontal, inclusivo y social. El mercado se está volviendo más inclusivo. Las redes sociales eliminan las barreras geográficas y demográficas, permitiendo que las personas se conecten y se comuniquen, y que las empresas innoven a través de la colaboración. Los clientes se están volviendo más orientados horizontalmente. Se vuelven cada vez más cautelosos con las comunicaciones de Marketing de las marcas y confían en cambio en el factor "F": *friends* (amigos), familias, fanáticos y *followers* (seguidores) Finalmente, el proceso de compra del cliente se está volviendo más social de lo que ha sido anteriormente. Los clientes están prestando más atención a su círculo social al tomar decisiones.

Las empresas deben responder de manera rápida y eficiente a los cambios en su situación interna y ambientes externos. Mucho dependerá de la capacidad de los altos ejecutivos al determinar qué tecnología es apropiada para ayudarlos a alcanzar sus objetivos. Es probable que las organizaciones exitosas sean las que pueden adquirir trabajadores clave que aprovechen las capacidades de las nuevas tecnologías efectivamente. Esto significa usar Internet y otras tecnologías para compartir conocimientos, crear nuevos productos y servicios, agregar valor a los clientes y mejorar la eficiencia interna.

Capítulo 3: Minería de la Web Social

3.1 Introducción

En este capítulo se exponen las temáticas relacionadas a la Minería de Web Social, misma que se encarga de extraer, analizar e identificar patrones de los datos generados en plataformas digitales. Este enfoque de Minería de Datos se basa en teorías y metodologías de optimización de datos a través modelos y algoritmos que sean capaces de representar gráficamente patrones significativos a gran escala. Bajo este contexto es donde se encuentra la importancia de dichos datos puesto que, día a día se crea nueva información y a gran velocidad, esto se debe a que hay más personas creando perfiles a través de las redes sociales y estas mismas consumen algún tipo de entretenimiento y cada vez que se genera algún tipo de actividad en línea esta es guardada en diferentes bases de datos a nivel mundial. La transferencia de datos en el mundo virtual permite promover y estimular la acción comunicativa de las personas bajo un contexto de bajo costo y libertad de expresión (Álvarez Tabares & Rodríguez Guerra, 2012)

Los consumidores actuales, suelen compartir opiniones, sentimientos, anhelos y frustraciones sobre servicios y productos que ofrecen las organizaciones y que tienen presencia digital. Para las organizaciones es elemental saber qué es lo que piensan sus consumidores y clientes potenciales es por ello, que los datos generados por ellos pasa a ser información valiosa misma que es extraída, clasificada y analizada para la toma de

decisiones en las empresas puesto que, el saber que “piensa” o “siente” los consumidores y clientes potenciales es indispensable para la planeación estratégica de Marketing y comunicación.

Al extraer la información de las plataformas digitales es necesario aplicar técnicas de Minería de Datos para extraer el conocimiento de los datos web, incluyendo documentos, hipervínculos y sitios (Velasco, 2013). La información que es extraída de la Web puede ser vista como un área de investigación interesante debido a que, ayuda extraer el conocimiento de diferentes sitios. De acuerdo con el objetivo de análisis, la Minería Web puede dividirse en tres tipos según los autores (Ramesh, Dursun, & Turban, 2015):

- Minería Web de Contenidos
- Minería Web de Estructura
- Minería Web de Uso

La presente investigación se enfocará en el uso de la Minería Web que a través de la Minería de Opiniones misma que pretende analizar lo que los clientes y consumidores del grupo educativo piensan de las publicaciones realizadas en sus diferentes plataformas digitales bajo la ejecución de la campaña digital “El Camino al Éxito”, al entender la opinión de los usuario, este conocimiento puede aplicarse en personalización Marketing y diseño de contenido e incluso como apoyo para tomar decisiones de planeación futuras.

3.2 Minería de la Web

El Internet ha cambiado para siempre el panorama de los negocios tal como se conocía, las empresas mismas que, enfrentan cada vez más oportunidades (como lo llegar a clientes y mercados que nunca pensaron que eran posible) y un desafío mayor (un mercado competitivo globalizado y en constante cambio). Las organizaciones que tienen la visión y las capacidades para lidiar con un entorno tan inestable se benefician enormemente de él, mientras que otros que se resisten tienen dificultades para sobrevivir. Tener una presencia comprometida en Internet ya no es una opción: es un requisito comercial dado que, los clientes esperan que las compañías ofrezcan sus productos y/o servicios a través de Internet.

Las plataformas digitales no solo permiten que los usuarios compren productos y contraten servicios a través de redes, sino que también hablan sobre las organizaciones, campañas publicitarias y comparten sus experiencias transaccionales y de uso con otros

en distintos sitios web. El crecimiento de Internet y sus tecnologías habilitadoras ha facilitado la creación, la recopilación y el intercambio de datos, información y opiniones. Las organizaciones, privadas y públicas, recopilan continuamente datos, información y conocimiento a un ritmo cada vez más acelerado y los almacenan en sistemas computarizados. Es alarmante el hecho de que estas redes se han convertido en un grupo importante de datos no estructurados que pertenecen a una gran cantidad de dominios, incluidos negocios, gobiernos y salud. Mantener y utilizar estos datos e información se vuelve extremadamente complejo, especialmente a medida que surgen problemas de escalabilidad. Además, el número de usuarios que necesitan acceder a la información continúa aumentando como resultado de una mayor confiabilidad y disponibilidad del acceso a la red, especialmente a Internet. (Ramesh, Dursun, & Turban, 2015) Por otra parte, la creciente dependencia de las redes sociales exige técnicas de extracción de datos que puedan facilitar esta actividad de los datos no estructurados y colocarlos dentro de un patrón sistemático para ser analizados posteriormente (Injadat, Salo, & Nassif, 2016).

Debido al gran tamaño de la Web, no es factible configurar un almacén de datos para replicar, almacenar e integrar todos los datos en la misma, lo que hace que la recopilación e integración de datos sea un desafío para las organizaciones dado que, esta no solamente crece rápido, sino que su contenido se actualiza constantemente como el que se encuentra en redes sociales, blogs, plataformas digitales de interacción y sitios de retransmisión en directo (*streaming*). Lo anterior, es un desafío que ha provocado diferentes esfuerzos de investigación para mejorar la efectividad y la eficiencia de descubrir y usar activos de datos en la web. Varios motores de búsqueda web basados en índices buscan constantemente en la web e indexan sitios web bajo ciertas palabras clave. La Minería Web es el proceso de descubrir relaciones intrínsecas (es decir, información interesante y útil) a partir de datos web, que se expresan en forma de información textual, de enlace o de uso (Ramesh, Dursun, & Turban, 2015).

Esta es un área de tecnología y práctica empresarial en constante evolución. La Minería Web es esencialmente lo mismo que la Minería de Datos sólo con la diferencia que esta utiliza datos generados a través de Internet. El objetivo es convertir grandes y complejos depósitos de transacciones comerciales, interacciones con los clientes y datos de uso del sitio web en información procesable (en otras palabras, en información y conocimiento) para promover una mejor toma de decisiones en toda la empresa.

Desde una perspectiva general, el Análisis Web puede considerarse una parte de la Minería Web. A continuación, se presenta una taxonomía simple de Minería Web, donde se divide en tres áreas principales: Minería Web de Contenidos, Minería de Estructura Web y Minería de Uso Web.

A continuación, se muestra la ilustración número 11 con dicha relación:

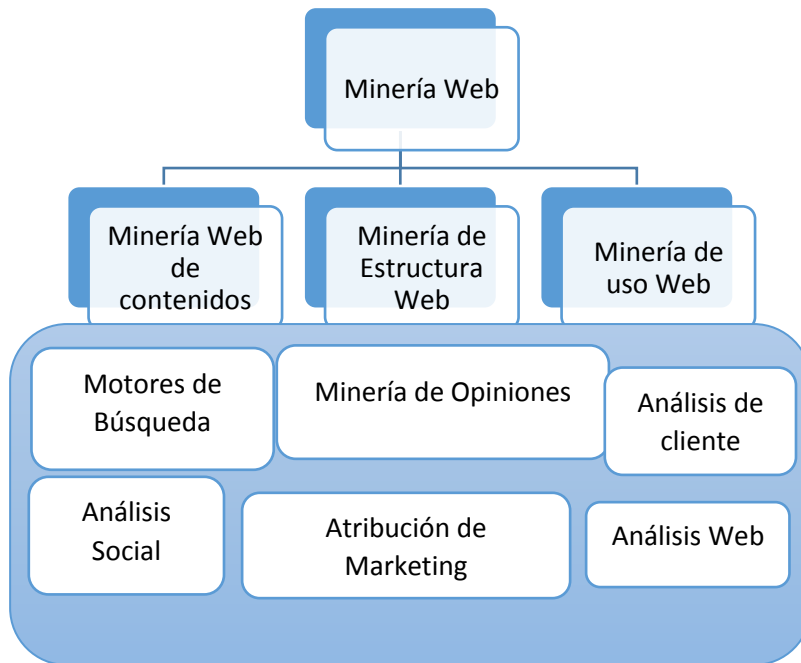


Ilustración 11.-Taxonomía de la Minería Web (Ramesh, Dursun, & Turban, 2015)

El gráfico anterior muestra la relación que tiene la Minería de Web y la Minería de Uso Web misma que tiene relación directamente con Minería de Opiniones ya que, precisamente buscan extraer datos de plataformas digitales que puedan servir como referencia para la toma de decisiones enfocadas a productos y servicios. A continuación se muestra la definición de Contenido Web y Estructura de la Minería Web conceptos que proporcionan un mejor entendimiento de dichas temáticas.

3.2.1 Contenido web y estructura de la Minería Web

La Minería de Contenido Web se refiere a la extracción de información útil de las páginas web. Estos documentos se pueden extraer en algún formato legible para computadoras a partir de las técnicas automatizadas puedan extraer cierta información de las páginas

web. Los rastreadores web (también llamados *spiders*) se utilizan para leer el contenido de un sitio web automáticamente. La información recopilada puede incluir diferentes características de documento similares a las que se usan en la Minería de Texto, pero también puede incluir conceptos adicionales, como la jerarquía de documentos. Tal proceso automatizado (o semiautomatizado) de recopilación y extracción de contenido web se puede utilizar para la inteligencia competitiva (recopilación de información sobre productos, servicios y clientes de la competencia). También se puede usar para recopilar y resumir información, Análisis de Sentimientos, opiniones, recopilación automatizada de datos y estructuración para el modelado predictivo. Es necesario mencionar que una gran cantidad de información en la web proporciona una rica colección de información sobre la relevancia, calidad y estructura de los contenidos de la web, y por lo tanto es una fuente rica para la Minería Web y en su defecto, para las organizaciones.

Por otra parte, la Minería de Contenido Web también se puede utilizar para mejorar los resultados producidos por los motores de búsqueda. De hecho, la búsqueda es quizás la aplicación más predominante de Minería de Contenido Web y Minería de Estructuras Web (Ramesh, Dursun, & Turban, 2015). Una búsqueda en buscador de Internet para obtener información sobre un tema específico por lo general, muestra páginas web relevantes y de calidad y un mayor número de páginas web inutilizables. El uso de un índice de relevancia basado en palabras clave y páginas autorizadas mejorará los resultados de búsqueda y la clasificación de las páginas relevantes.

La Minería de Estructuras Web es el proceso de extraer información útil de los enlaces incrustados en documentos web. Se utiliza para identificar páginas y centros autorizados, que son los pilares de los algoritmos contemporáneos de clasificación de páginas que son centrales para los motores de búsqueda populares como Google. Así como los enlaces que van a una página web pueden indicar la popularidad de un sitio, los enlaces dentro de la página web pueden indicar la profundidad de la cobertura de un tema específico. El análisis de enlaces es muy importante para comprender las interrelaciones entre un gran número de páginas web, lo que lleva a una mejor comprensión de una comunidad internauta de manera general.

3.2.2 Minería de Uso Web (Análisis Web)

La Minería de Uso Web (también llamada Análisis Web) es la extracción de información útil de los datos generados a través de visitas a la página web y transacciones. Algunos

autores como (Masand, Spiliopoulou, Srivastava, & Zaïane, 2002) afirman que al menos tres tipos de datos se generan a través de visitas a páginas web:

1. Datos generados automáticamente almacenados en registros de acceso al servidor, registros de referencia, registros de agente y cookies de los dispositivos del cliente.
2. Perfiles de usuario.
3. Metadatos, como atributos de página, atributos de contenido y datos de uso.

El análisis de la información que es recopilada de los servidores web puede ayudar de manera exponencial a comprender de mejor manera el comportamiento del usuario. El análisis de estos datos a menudo se llama análisis "*clickstream*". Al utilizar las técnicas de Minería de Datos y texto, una empresa podría encontrar patrones interesantes de los flujos de clics y así, extraer información y saber cómo utilizar ese conocimiento generado para mejorar el proceso, mejorar el sitio web y, lo más importante, aumentar el valor para el cliente. La Minería Web tiene una amplia gama de aplicaciones empresariales. Por ejemplo, los autores (Nasraoui, Soliman, Saka, Badia, & Germain, 2008) enumeraron algunas de las aplicaciones más comunes:

1. Determinar el valor de cada cliente.
2. Diseñar estrategias de Marketing cruzado en todos los productos.
3. Evaluar campañas promocionales.
4. Analizar anuncios y correos electrónicos *online* en grupos de usuarios basados en patrones de acceso.
5. Predecir el comportamiento del usuario según las reglas aprendidas previamente y los perfiles de los usuarios
6. Presentar información dinámica a los usuarios en función de sus intereses y perfiles.

Así mismo, existen numerosas herramientas y tecnologías para el Análisis Web en el mercado. Debido a su poder para medir, recopilar y analizar datos de Internet para comprender mejor y optimizar el uso de Internet, la popularidad de las herramientas de Análisis Web está aumentando. La Analítica Web promete revolucionar la forma en que se hacen negocios tanto dentro como fuera de la web. El análisis enfocado hacia este sector digital no es solo puede clasificarse como una herramienta para medir el tráfico web;

también se puede utilizar como un instrumento de investigación comercial y de mercado. Las aplicaciones de Análisis Web también pueden ayudar a las empresas a medir los resultados de las campañas publicitarias impresas o de difusión tradicionales. Puede ayudar a estimar cómo cambia el tráfico a un sitio web después del lanzamiento de una nueva campaña publicitaria (Ramesh, Dursun, & Turban, 2015).

El Análisis Web proporciona información sobre la cantidad de visitantes que hay en un sitio web y la cantidad de visitas a la página en un periodo determinado. A su vez, ayuda a medir el tráfico y las tendencias de popularidad, que se pueden utilizar para la investigación de mercado. Hay dos categorías principales de Análisis Web; fuera del sitio y en el sitio. La analítica web fuera del sitio se refiere a la medición y Análisis Web sobre la organización y sus productos de manera externa del sitio web. Incluye la medición de la audiencia potencial de un sitio web (perspectiva y oportunidad), la publicidad de boca en boca) y el Análisis de Sentimiento y opiniones, perspectivas que se tiene de un producto y servicios en Internet.

De manera histórica, el Análisis Web se ha referido a la medición de visitantes en el sitio. Sin embargo, en los últimos años ha obtenido nuevos horizontes, principalmente que se están produciendo herramientas que abarcan ambas categorías. El Análisis Web en el sitio mide el comportamiento de los visitantes una vez que están en su sitio web. Esto incluye conversiones en un sitio web, por ejemplo, el grado en que diferentes páginas de destino están asociadas con compras en línea.

El Análisis Web en el sitio mide el rendimiento de su sitio web en un contexto comercial. Estos datos recopilados en el sitio web se comparan con los indicadores clave de rendimiento y se utilizan para mejorar la respuesta de la audiencia de un sitio web o campaña de Marketing. A pesar de que *Google Analytics* es el servicio de Análisis Web más utilizado y constantemente surgen nuevas y mejores herramientas que proporcionan capas adicionales de información. Para el Análisis Web en el sitio, hay dos formas técnicas de recopilar los datos. El primer y más tradicional método es el análisis del archivo de registro del servidor, donde el servidor web registra las solicitudes de archivos realizadas por los navegadores. El segundo método es el etiquetado de páginas, que utiliza JavaScript incrustado en el código de la página del sitio para realizar solicitudes de imágenes a un servidor dedicado a análisis de terceros cada vez que un navegador web representa una página (o cuando se produce un clic del mouse). Según los autores (Ramesh, Dursun, & Turban, 2015) ambos recopilan datos que pueden procesarse para

generar informes de tráfico web. Además de estas dos corrientes principales, también se pueden agregar otras fuentes de datos para aumentar los datos de comportamiento del sitio web. Estas otras fuentes pueden incluir correo electrónico, datos de campaña de correo directo, historial de ventas y clientes potenciales, o datos originados en redes sociales.

3.2.3 Métricas de Análisis Web

Utilizando una variedad de fuentes de datos, los programas de Análisis Web proporcionan acceso a una gran cantidad de datos de Marketing valiosos, que pueden aprovecharse para obtener mejores conocimientos para hacer crecer a la organización. Los datos y la información obtenida de analítica web se pueden utilizar para administrar eficazmente los esfuerzos de Marketing de una organización y sus diversos productos o servicios. Los programas de Análisis Web proporcionan datos casi en tiempo real, que pueden servir como base de datos para posteriormente utilizarlos en campañas de Marketing o permiten realizar ajustes oportunos a las estrategias de Marketing actuales. Si bien la analítica web proporciona una amplia gama de métricas, existen cuatro categorías de métricas que generalmente son accionables y pueden afectar directamente los objetivos de la organización (Masand et al., 2002) Estas categorías incluyen:

- Uso del sitio web
- Fuentes de tráfico
- Perfiles de los visitantes
- Estadísticas de conversión

3.2.4 Usabilidad del sitio web

Es importante mencionar la usabilidad del sitio web está completamente integrada al giro de la organización y esta va de la mano con la analítica del mismo, puesto que son muchos los datos que se generan, todos estos datos pueden ser reconocidos a través de indicadores y herramientas que permitan posteriormente almacenarlos para ser analizarlos a beneficio de entender la estadística del sitio web en futuras planeaciones de la organización.

A continuación, se muestran algunos indicadores de usabilidad de un sitio web funcional:

1. Vistas de página: Es sin duda una de las medidas más básicas y generalmente esta métrica es representada por el número de visitas promedio que se tiene en una página web o plataforma digital.

2. Tiempo en el sitio: Es una de las medidas fundamentales en cuestión de interacción con el visitante. De manera general, cuanto más tiempo el internauta pase en el sitio web, mayor impacto positivo tendrá.

3. Descargas: Esto incluye archivos PDF, videos y otros recursos que se ponen a disposición del internauta, es importante considerar qué tan accesibles son estos elementos y de qué manera se están proporcionando.

4. Clics: La mayoría de los programas de análisis enfocados en el sitio web pueden mostrar el porcentaje de clics ha recibido cada elemento en la página web.

3.2.5 Fuentes de tráfico

Los programas que están enfocados en el Análisis Web son una herramienta poderosa y útil para identificar dónde se origina el tráfico web. Los motores de búsqueda, sitios web de referencia y visitas de páginas marcadas son categorías básicas para fomentar el tráfico en línea. Sin embargo, también se puede identificar el tráfico web generado por las diversas campañas publicitarias tanto en línea como fuera de línea de la organización:

- 1. Sitios web de referencia:** Los sitios web que contienen enlaces que envían visitantes directamente a otro sitio web se consideran sitios web de referencia.
- 2. Motores de búsqueda:** Los datos en la categoría del motor de búsqueda se dividen entre búsqueda paga y búsqueda orgánica (o natural). Se debe revisar las principales palabras clave que generaron tráfico web del sitio y ver si están representando de manera correcta los productos y servicios en los que está inmersa la organización.
- 3. Directo:** Las búsquedas directas se atribuyen a dos fuentes. Una persona que guarda una página web en favoritos y haga clic en ese enlace se

registrará como una búsqueda directa. Otra fuente ocurre cuando alguien escribe su URL directamente en su navegador.

4. **Campañas sin conexión:** Si se utilizan opciones de publicidad que no están directamente relacionadas la web, el Análisis Web puede capturar datos de rendimiento si se incluye un mecanismo para enviarlos al sitio web.
5. **Campañas en línea:** Es posible medir la efectividad de una campaña individual simplemente usando una URL similar a la estrategia de campaña fuera de línea.

3.2.6 Perfiles del internauta

Una de las formas en que puede aprovechar sus Análisis Web en una herramienta de Marketing realmente poderosa es a través de la segmentación. Al combinar datos de diferentes informes analíticos, a continuación, se encontrarán algunas formas de analizar el perfil del internauta:

- **Palabras clave:** En su informe analítico, puede ver qué palabras clave utilizaron los internautas en los motores de búsqueda para ubicar el sitio web de la organización.
- **Agrupaciones por contenido:** Dependiendo de cómo agrupe el contenido la organización en su sitio web, se pueden analizar secciones del mismo correspondientemente dirigido a productos, servicios, campañas y otras tácticas de Marketing específicas.
- **Geografía:** Este análisis permite ver dónde se origina geográficamente el tráfico, incluido la ubicación de país, estado y ciudad. Esto puede ser especialmente útil si se utiliza en campañas con orientación geográfica o si desea medir la visibilidad de cierto contenido en una región específica.
- **Hora del día:** El tráfico web generalmente tiene picos al comienzo de un día laboral, durante el almuerzo y en el final del mismo. Es importante analizar estos datos para determinar cuándo el internauta simplemente navega en lugar de comprar y también para tomar decisiones sobre qué horas es conveniente ofrecer servicio al cliente.

Como conclusión de la temática explicada anteriormente y como conector al siguiente subtema, las organizaciones saben de manera certera que conocer a su cliente y comunicarse de manera correcta con él, genera una gran cantidad de datos, que suelen estar disponibles para todos, incluso para los competidores. Debido a esto, las empresas deben de analizar de manera minuciosa estos datos y para realizarlo, existen técnicas de Minería de Texto que permite a la organización obtener información confiable, en tiempo y forma misma que, permite vigilar el comportamiento de sus clientes y consumidores y así tomar decisiones acertadas conforme a la estrategia competitiva de la organización. La relación que existen entre la Minería de Texto y las organizaciones se encuentra mediante el análisis de texto en las principales plataformas digitales y de comunicación en línea como lo es: Facebook, Twitter, Instagram, YouTube, entre otros. Muchos estudios han indicado que existe una fuerte relación entre la rentabilidad financiera y el uso de estas técnicas para el Análisis de Sentimientos dirigido hacia el cliente y/ o consumidor (Germán & Jiménez-Quintero, 2015).

3.3 Minería de Texto

En este capítulo se proporciona una visión general bastante completa de las siguientes temáticas: Minería de Texto y una de sus aplicaciones más populares, el Análisis de Sentimientos, puesto que ambos se relacionan con el análisis empresarial, planeación estratégica y por supuesto, para la toma de decisiones de decisiones. En términos generales, el Análisis de Sentimientos es un derivado de la Minería de Texto, y esta misma es esencialmente un derivado de la Minería de Datos.

3.3.1 Análisis de texto y Minería de Texto conceptos y definiciones

El análisis de texto y la Minería de Texto es un proceso de se adapta a las necesidades de las empresas puesto que, estas generan una gran cantidad de datos, de acuerdo con un informe de *Data never Sleeps*, elaborado por el sistema operativo basado en la nube Domo, una compañía dedicada las tecnologías de información, cada día se crean en Internet más de 2,5 billones de bytes de datos, que van en aumento, especialmente en las plataformas de video como Netflix o Youtube. Conforme a lo anterior, El objetivo general tanto del proceso de Análisis de Texto y Minería de Texto es convertir precisamente datos textuales no estructurados en información que pueda ser procesable mediante la aplicación del procesamiento de lenguaje natural o su nombre en inglés: *Natural Language Processing* (PLN).

Algunos autores mencionan que existe diferencias entre Análisis de Texto y Minería de Texto, puesto que el primero de estos incluye métodos enfocados en la recuperación y extracción de información, Minería de Datos y también Minería Web mientras que la Minería de Texto se centra primordialmente en descubrir nuevas fuentes de conocimiento enfocados en datos textuales. (Anandarajan, Hill, Nolan, 2018)

A continuación, el autor (Ramesh et al., 2015) ilustra en el libro “*Business intelligence and analytics: systems for decision support*” la relación que tienen existe entre el Análisis de Texto y la Minería de Texto junto con otras áreas de aplicación relacionados. A continuación, se muestran:

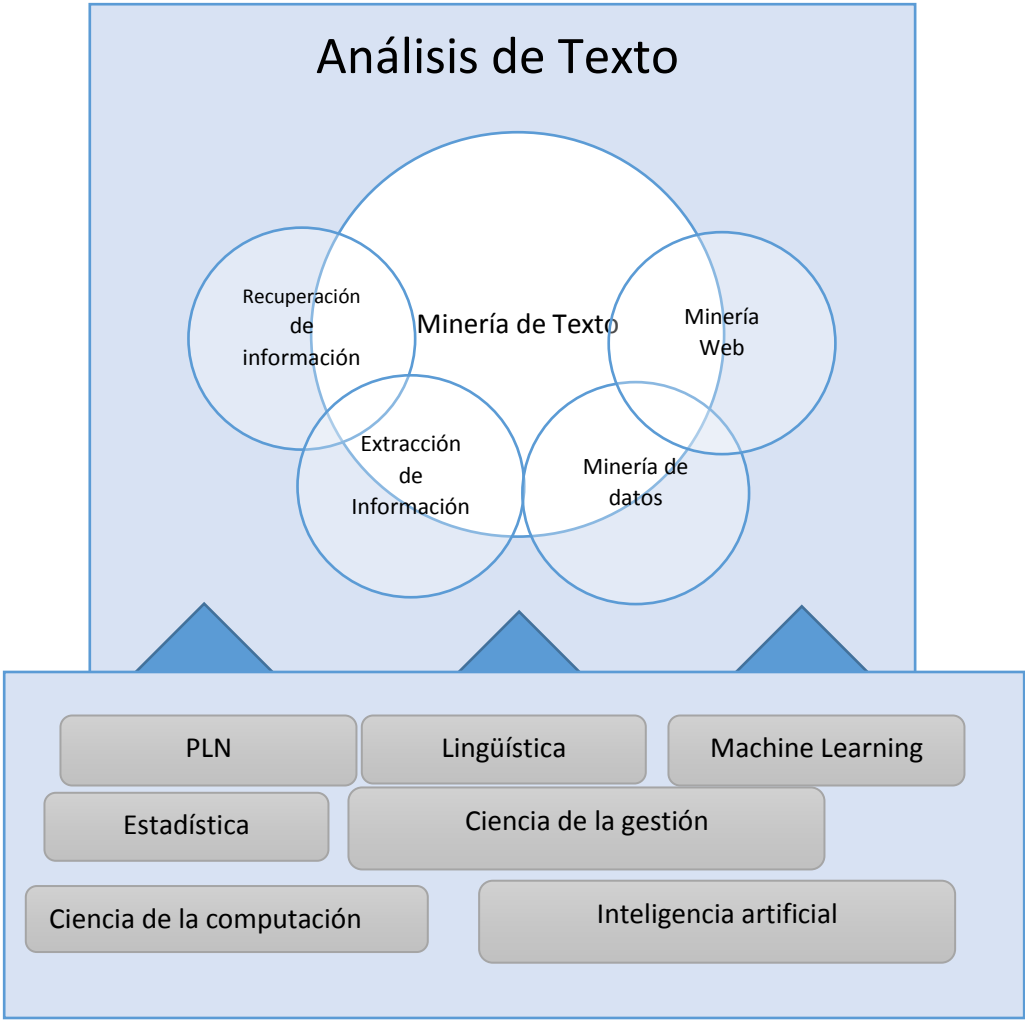


Ilustración 12.- Minería de Texto y áreas de aplicación relacionadas.

La figura número 12 ilustra las relaciones existentes entre el análisis de texto y la Minería de Texto junto con otras áreas de aplicación relacionadas. La parte inferior de la Figura

número 12 numera las principales disciplinas base que tienen una importante participación en el desarrollo de estas áreas de aplicación.

La Minería de Texto es un proceso semiautomatizado de extracción de patrones de grandes cantidades de fuente de datos no estructurados mismos que se caracterizan por no tener un formato específico y se pueden almacenar en múltiples formatos. La Minería de Texto tiene la misma finalidad y propósito que la Minería de Datos puesto que, ambas buscan el proceso para identificar patrones válidos, novedosos y que pueden ser de gran utilidad, en el caso de la Minería de Texto la entrada al proceso proviene de una colección de archivos de datos no estructurados como documentos Word, extractos de texto, archivos xml, cvs, etc. (Ramesh et al., 2015)

De manera esencial, la Minería de Texto tiene diferentes beneficios en las áreas donde se generan grandes cantidades de datos textuales como en la ley, la investigación académica, las finanzas, la medicina, biología, tecnología y Marketing Tradicional y digital por ejemplo, todas las interacciones de tipo texto que proporcionan los clientes cuando se quejan o alagan algún producto y/o servicio, la organización puede tomar ventaja competitiva de lo anterior y aprovechar para identificar objetivamente las características del producto y servicio que consideran menos agradables para mejorarlo cuando se desarrollen nuevas asignaciones del mismo en el futuro. Del mismo modo, la opinión del cliente se encuentra en la divulgación del mercado y los grupos focales en redes sociales que generan grandes cantidades de datos, al no restringir comentarios sobre productos y servicios, los clientes pueden presentar, con sus propias palabras, lo que realmente piensas sobre ellos.

La Minería de Texto tiene diferentes funciones según su área de aplicación a continuación se mencionan las relevantes:

- Extracción de información: Identificación de palabras clave, frases y relaciones entre el texto y secuencias predefinidas del texto mediante la coincidencia de patrones.
- Seguimiento de temas importantes (*trending topics*): Basado en un perfil de usuario y lo que el internauta ve y busca, la Minería de Texto puede predecir temas de interés para el usuario.
- Resumen: Resumir un documento para ahorrar tiempo al lector.
- Categorización: Identificar los temas principales de un documento y luego colocar el mismo en un conjunto predefinido de categorías basadas en esos temas.

- Agrupación: Agrupar documentos similares sin tener un conjunto predefinido de categorías.
- Vinculación de conceptos: Relaciona documentos identificado los conceptos en común, esto permite que los usuarios puedan encontrar información relevante utilizando métodos búsqueda tradicionales.
- Preguntas de respuesta: Encontrar la mejor respuesta a una pregunta dada a través de patrones y algoritmos basados en el conocimiento.

3.3.2 Procesamiento del Lenguaje Natural

La tendencia actual en la Minería de Texto es incluir muchas de las funciones avanzadas que solamente se pueden obtener con técnicas de Procesamiento del Lenguaje Natural (PLN o PLN por sus siglas en inglés). Este campo se enfoca primordialmente en estudio, comprensión, diseño e implementación de los elementos software necesarios para el tratamiento computacional del lenguaje natural, entendiendo como tal el lenguaje humano y abarca parte de la Ciencia de Datos, Inteligencia Artificial (Aprendizaje Automático) y la lingüística.

La evolución del NPL ha evolucionado desde la era de las tarjetas perforadas y el procesamiento por lotes, en el que el análisis de una oración podría más de 7 minutos, hasta la era de Google y similares, en la que millones de páginas web pueden ser procesadas en menos de un segundo. En otras palabras, el PLN permite a las computadoras realizar una gama de tareas difíciles relacionadas directamente con el lenguaje natural en todos los niveles y puede ir desde el análisis y el etiquetado procesarse en menos de un segundo. PLN permite a las computadoras realizar una amplia gama de tareas relacionadas con el lenguaje natural en todos los niveles, desde el análisis y el etiquetado de cada palabra en un texto también conocido como “*partof-speech*” (POS), hasta la traducción automática y los sistemas de diálogo (Young, Hazarika, Poria, Cambriam, 2018)

Como se ha comprendido en el párrafo anterior, el objetivo principal del PLN es ir más allá de la manipulación de texto basada solamente en la sintaxis (conteo de palabras) sino tener una mayor comprensión y procesamiento del mismo a través de restricciones gramaticales y semánticas, así como el contexto. A continuación, se mencionan algunos de los desafíos más comunes asociados con la implementación de PLN:

- **Etiquetar de palabras:** Es complicado marcar términos en un texto como correspondientes a una parte particular del discurso (como sustantivos, verbos, adjetivos, adverbios, etc.) puesto que estos componentes no solo dependen de la definición del término sino también en el contexto dentro del cual se usa.
- **Segmentación de texto:** Algunos idiomas escritos, como chino, japonés y tailandés, no tiene límites de una sola palabra. En estos casos, la tarea de análisis de texto requiere la identificación de límites de palabras, lo que a menudo es una tarea difícil. Así mismo los desafíos en la segmentación del habla surgen cuando se pretenden analizar el lenguaje hablado, porque los sonidos que representan letras y palabras sucesivas se mezclan entre sí, causando confusión.
- **Desambiguación del sentido de las palabras:** Muchas palabras tienen más de un significado, lo que causa problemática en la selección del significado que tiene más sentido conforme al contexto.
- **Ambigüedad sintáctica:** La gramática de los lenguajes naturales es ambigua; es decir, que se necesita muchas veces considerar la estructura de la oración. Elegir la estructura apropiada requiere de fusión semántica y contextual de la información analizada.
- **Entrada imperfecta o irregular:** Los acentos extranjeros o regionales y los impedimentos vocales en el habla y los errores tipográficos o gramaticales en los textos hacen que el procesamiento del idioma sea una tarea aún más difícil.

Por otra parte, el NPL se ha usado para darle sentido a enormes cantidades de texto escrito que millones de internautas publican diariamente en las plataformas digitales. Los textos que se encuentran en estos sitios webs a menudo están escritos de manera que no puede ser “entendido” fácilmente por las computadoras y, por lo tanto, solo puede ser comprendido por otros usuarios humanos y no se puede lograr un análisis automatizado dirigido por el procesador hasta que la información sea transformada en un formato comprensible por la computadora (Balazr Thenot, 2015) y es precisamente por esta razón que el NPL es necesario para tareas relacionadas con la Minería de Opinión o Análisis de Sentimientos para dar sentido a un texto generado por el usuario y, en particular, extraer opiniones de este.

3. 3. 3 Minería de Opinión

Como se ha entendido anteriormente, la Minería de Opinión deriva de la Minería de Textos y es importante recordar que la Minería de Textos no es lo mismo que la Minería de Datos, puesto que, la Minería de Datos consiste en extraer datos ya estructurados en cambio, tanto la Minería de Textos, como la Minería de Opiniones trabajan con información no estructurada. (Conde Sánchez, 2012)

La Minería de Opiniones, de manera específica, se ha aplicado de manera exitosa para el estudio de las redes sociales permitiendo estructurar y organizar los datos de manera general acerca de determinados temas y de forma automatizada, en específico, se utiliza para analizar los sentimientos, opiniones y actitudes expresadas por diferentes internautas en la web. Sin embargo, la Minería de Opiniones, como ya se ha expuesto anteriormente se enfrenta a diferentes dificultades para su estudio y análisis en todo el mundo, al trabajar con datos no estructurados y en diferentes idiomas según los autores Williams, L., Bannister, C., Arribas-Ayllon, M., Preece, A., & Spasić, I. (2015)

Definición

En términos más amplios, la Minería de Opinión es la ciencia del uso del análisis de texto para comprender los impulsores detrás del sentimiento de diferentes clientes y/o consumidores en plataformas digitales. De manera técnica, autores como Pang y Lee, (2008) así como, Bing Liu, (2012) definen este concepto como una tupla, que en lenguaje de programación es una estructura de dato o un conjunto ordenado de valores mismo que, contiene el objetivo al que se dirige la opinión.

Clasificación de opinión

El autor Bing Liu, (2012) menciona en su publicación: "Sentiment Analysis and Opinion Mining", que las opiniones suelen clasificarse en dos principales categorías:

- **Opiniones regulares:** Son todas aquellas opiniones que se refieren a una sola identidad o en su defecto, aspecto de una identidad.

1. Opiniones regulares directas: Son aquellas que se refieren de manera primordial y explícitamente a una entidad o uno de sus aspectos. Son lo más fáciles y simples de manejar y la mayoría de las investigaciones existente se centran en estas.

Ejemplo:

1	La universidad tiene un pésimo servicio al cliente
---	--

En la oración anterior, el aspecto “servicio al cliente” de la entidad “universidad” se juzga directamente con una polaridad negativa.

2. Opiniones regulares indirectas: Estas opiniones son todas aquellas que se expresan indirectamente en una entidad o en uno de sus aspectos, generalmente, elementos y efectos observables.

Ejemplo:

2	Termine el examen en dos horas
---	--------------------------------

En la oración anterior, indirectamente, indica que el aspecto “dos horas” es decir, “el tiempo del examen” de la entidad “examen” fue “corto”, lo que podría considerarse positivo o negativo según el contexto. Sin embargo, “el tiempo del examen” no se mencionó directamente y tuvo que inferirse el hecho de que el emisor de dicha opinión término en dos horas como tal y se indicó.

- **Opiniones comparativas:** Las opiniones comparativas: Es una opinión que apunta hacia un grado de similitud o diferencia entre dos o más unidades o en su defecto, entre dos o más aspectos. Estas se subdividen, al igual que las opiniones regulares, la subdivisión es la siguientes:

1. Opiniones comparativas directas: Esta establece de manera directa e explícitamente el grado de diferencia o similitud entre dos o más elementos.

Ejemplo:

3	El servicio al cliente la universidad “A” es mejor que de la universidad “B”.
---	---

El ejemplo muestra claramente que aspecto “servicio al cliente” de la universidad “A” es mejor que la universidad “B”. Por otra parte, si se informa de manera adicional,

sería posible establecer la polaridad de ambas universidades en una escala absoluta.

2. Opiniones comparativas indirectas: Establece indirectas que señalan el grado de diferencia o similitud entre dos o más elementos sin mencionar, explícitamente el aspecto que se está comparando.

3.4 Proceso de Minería de Opinión

- **Etapas 1. Extracción de datos**

La primera etapa para cualquier proceso de Minería de Opinión es la adquisición de datos y consiste de manera general en extraer el corpus que se va a extraer para obtener las opiniones. Existen dos enfoques para extraer opiniones. El primero es a través de la programación de aplicaciones en un sitio web y el segundo es a través de rastreadores web para capturar los datos de un sitio deseado (Balazr Thenot, 2015) como, por ejemplo, Twitter o Facebook.

Enfoques para lograr esta tarea. El primero es a través de la interfaz de programación de aplicaciones (API) y el segundo es con rastreadores web. En ambos enfoques presentan algunas ventajas y desventajas, algunos autores comparan han comparado ambos métodos (Kaijie Guo, Liang Shi*, Weilong Ye, Xiang Li, 2014) y se puede comprender, que existe una compensación entre el uso de cualquiera.

A continuación, se muestra la comparación que hacen los autores anteriores hacia los métodos de extracción de datos:

Con el enfoque basado en API, la implementación es fácil, los datos recopilados se ordenan y es poco probable que cambie su estructura, sin embargo, presenta algunas limitaciones dependiendo del proveedor de datos. Por ejemplo, las consultas de búsqueda a la API de Twitter están limitadas a 180 por 15 minutos.

El método está directamente relacionado a la disponibilidad de una API ya que, no todos los sitios web pueden proporcionar este acceso o en su caso no tienen las funcionalidades necesarias. Por otra parte, los enfoques que se basan en rastreadores son mucho más difíciles de implementar y los datos que se obtienen suelen estar menos estructurados, pero tienen una ventaja y es que no tienen restricciones. Además, los

rastreadores web pueden priorizar la extracción de datos subjetivos y relevantes para el objetivo con el que se esté trabajando en actualidad.

- **Etapa 2.- Preprocesamiento de texto**

El concepto de procesamiento de texto tanto para el Procesamiento de Lenguaje Natural (PLN) y la Minería de Opinión (MO) es diferente. El término de “pre-procesamiento” suele adaptarse al pre-procesamiento que la aplicación o la investigación se refiere.

El pre-procesamiento en ambas definiciones tiene la tokenización como otras etapas del proceso en común, pero de manera general, los objetivos de estas no suelen ser los mismos puesto que, el proceso de Minería de Opinión ocupa las tareas necesarias para el proceso central y estas tareas no son exclusivas para el método de procesamiento de PLN, en el primer caso se incluye tokenización, segmentación de oraciones, *streaming*, *tagging*, entre otros y, para el preprocesamiento de PLN se ocupan tareas generadas posteriormente (Balazr Thenot, 2015).

Se podría decir que el preprocesamiento de texto aplicado de PLN es más como un subconjunto de la metodología del procesamiento de texto de Minería de Opinión. Se debe tener en cuenta que en la metodología del procesamiento de texto enfocado en la Minería de Opinión incluye pasos del procesamiento de texto aplicado a PLN y algunos subprocesos de análisis de léxico y sintaxis.

A continuación, se muestra una figura que muestra la relación anterior:

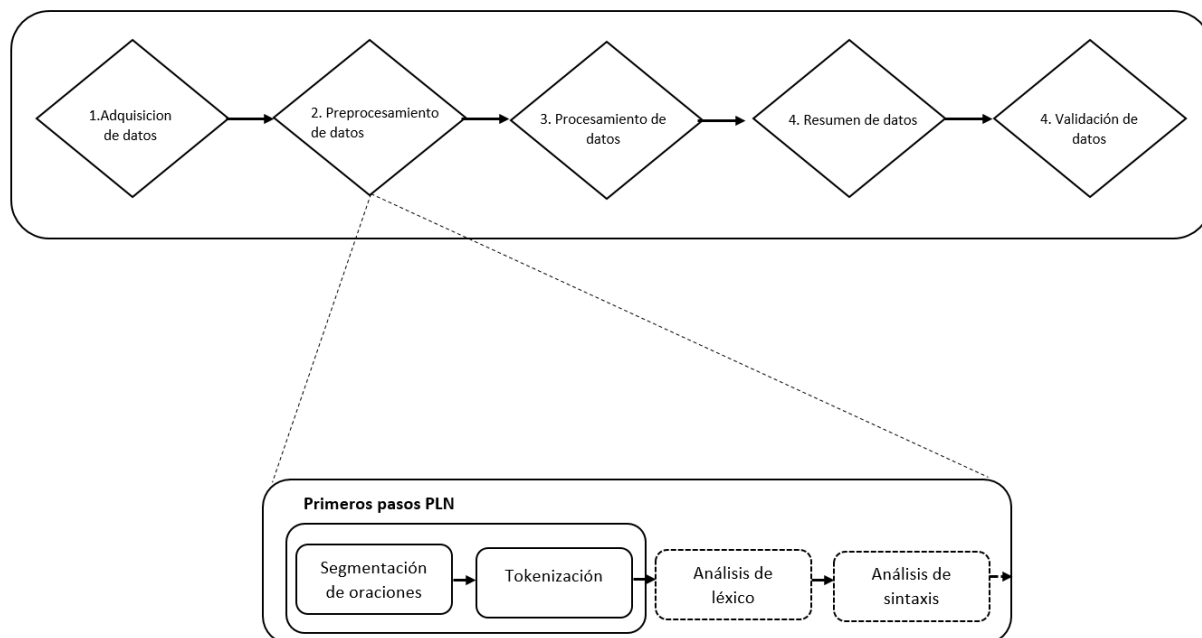


Ilustración 13.-Procesamiento del lenguaje natural en el procesamiento de texto como paso en el proceso de la Minería de Opinión

Para que lo anterior pueda clarificarse de manera efectiva, es necesario mencionar que el procesamiento de texto aplicado a PLN se compute de dos subprocesos:

- 1) Segmentación de oración
- 2) Segmentación de palabras (Tokenización)

El orden de utilizar ambos procesos dependerá de la aplicación en la que serán utilizados, en el caso específico de esta tesis se realizarán los siguientes pasos:

- **Conversión de mayúsculas y minúsculas:** este subproceso también es conocido como *case conversation* o *lowercasing*. Es una de las formas más simples y efectivas de procesamiento de texto puesto que, en la mayoría de los casos, las palabras en mayúsculas no coincidirán con las mismas palabras con letras minúsculas y viceversa y es aplicable a la mayoría de problema de Minería de Texto y PLN así mismo, puede ayudar en casos donde el conjunto de datos no es muy grande y puede mejorar significativamente con la consistencia del texto, convirtiendo la entrada de palabras en un formato único para facilitar su análisis. A continuación, se muestra un ejemplo de cómo las minúsculas resuelven la problemática anteriormente expuesta, donde las mismas palabras con diferentes

combinaciones de mayúsculas pueden causar confusión en el análisis, es por ello que se asignan minúsculas para unificar la palabra:

Raw Wod (palabra sin conversión)	Case conversation
México	méxico
México	
MÉXICO	

Ilustración 14.-Ejemplo de Case Conversation

- Tokenización:** La tokenización es el proceso de segmentar el texto en ejecución en palabras y/ u oraciones. Es decir, es responsable de fragmentar el texto en caracteres, palabras y frases. Naturalmente, antes de realizar cualquier procesamiento de texto este debe de segmentarse en unidades lingüística como palabras, puntuación, números alfanuméricos, etc. Esto ayuda significativamente a reducir la dimensionalidad del espacio de características y a promover la eficiencia del sistema en cuestión de clasificación de texto.

Un token, lingüísticamente hablando significativo y metodológicamente útil, detectar tokens significativos depende de la capacidad para comprender los patrones para una selección significativa.

La tokenización estándar (identificar los espacios en blanco) puede parecer una tarea fácil si se compara con otros procesos de lenguaje natural. Sin embargo, los errores que suelen cometerse en esta fase se verán reflejados en fases posteriores y pueden causar problemas en el análisis del texto.

- Palabras de detención:** Este proceso de análisis de texto también conocido como identificación de *stopwords*, en él se reduce la complejidad y la dimensionalidad a partir de eliminar las palabras más triviales, como pronombres, preposiciones y artículos, comparando cada *token* con una lista de palabras vacías. Las palabras vacías son palabras que se consideran que no transmiten un significado significativo y este proceso simplemente los elimina. Ejemplos de estas palabras son:

Ejemplo de stopwords
De, la, el, en, y, muy, pero, con, sin, ni, antes, a, acá, ahí, al, algo, así

Ilustración 15.- Ejemplo de Stopwords

La forma más común de llevar a cabo esta tarea es comparar cada token con una lista de palabras vacías y eliminar las que se identifiquen. Para compilar esta lista, el procedimiento habitual usar listas de *stopwords* en el idioma que se desea realizar el análisis.

El contexto de recuperación en este subproceso se refiere a ordenar cada término que aparece en los corpus por su frecuencia y enlistar los que son más frecuentes como palabras clave.

- **Corrección de términos** (*terms correction*): Este proceso busca corregir la ortografía y es posible que se puedan subdividir las tareas:
 - 1) **Palabras alargadas**: Identificar las palabras que estén “alargadas” es decir, que tengan un máximo de dos caracteres repetidos. Es esencial primero corregir este problema y posteriormente lidiar con errores ortográficos.
 - 2) **Corrección conforme a deletreo**: Este el proceso de corregir la ortografía de una palabra. En análisis de texto se pueden usar diferentes bibliotecas para realizar esta tarea.
- **Filtrado de términos**: Puesto que para la computadora el entender una oración puede ser una tarea difícil y también que la dimensionalidad de la base de datos puede ser muy amplia, un filtrado puede reducir las dimensiones del documento de manera efectiva. En este proceso se compara el texto con un vocabulario de dominio específico. Existen diccionarios de sinónimos en diferentes idiomas para el filtrado y se debe utilizar el que sea más apropiado para el análisis (Conde Sánchez, 2017).

Etapas 3.- Procesamiento

Análisis de Sentimiento: Para identificar la polaridad de las opiniones, es necesario realizar un Análisis de Sentimientos, es una técnica que utiliza el procesamiento de lenguaje natural, el análisis de texto y herramientas computacionales para clasificar los comentarios de los diferentes consumidores (Montesinos García, 2014) y en el caso específico de esta tesis, los internautas. Generalmente permite explicitar el conocimiento y en los últimos años se ha enfocado en el análisis de las redes sociales debido a que, permite estructurar la idea general de la opinión acerca de determinados temas. Una ventaja evidente de este proceso es que se pueden extraer datos de distintas plataformas digitales a nivel mundial y también cantidades considerables de los mismos. No obstante,

este proceso de Minería de Texto puede encontrar muchos obstáculos debido a que un texto puede contener diferentes matices gramaticales, variaciones culturales como jerga lingüística, comprensión contextual y también errores gramaticales. (Bannister, 2015)

Metodología para la Minería de Opiniones

Dado a que los datos extraídos no siempre están bien definidos para su análisis, los siguientes autores (Aggarwal & Zhai, 2012), (Bing Liu, Lei Zhang, 2012) proponen la siguiente metodología para estructurar y procesar a los mismos de una manera efectiva:

- Extracción y agrupación de entidades: Se extrae la entidad de cada comentario, es decir: tema, producto, servicio, persona, evento u organización (Conde Sánchez, 2017) Posteriormente, se deben agrupar las expresiones de entidades en clústeres.
- Extracción y agrupamiento de aspectos: Se extraen todos los aspectos de la opinión, particularmente a las características que describen a la identidad posteriormente, al igual que la primera tarea, se agrupan de acuerdo a sus similitudes.
- Extracción del emisor de la opinión y el tiempo: Consiste en extraer el emisor, es decir identificar quien ha expresado la opinion y cuando realizo esta actividad.
- Clasificación del sentimiento del aspecto: Se determina la positividad y la negatividad de los aspectos que componen la opinión.
- Generación de los quintuples de la opinión: En esta tarea se generan 5 variables en función al tratamiento de la información en las tareas anteriores.

Detención de polaridad

La polaridad de una oración como se ve representado en las tareas anteriores, deriva de la clasificación de sentimientos y consiste primordialmente en determinar la postura expresada de dicha oración. Es decir, si es positiva, negativa o neutral hacia el tema general. Autores como (Eguchi, Lavrenko, 2006) señalan que las etiquetas de polaridad pueden ser utilizadas simplemente para resumir el contenido de unidades de texto con opiniones sobre un tema, ya sean positivas o negativas, o solo para recuperar elementos de una orientación de sentimiento.

Existen 3 enfoques sobre los cuales se puede determinar la polaridad de un texto, estos se llevan a cabo a través del aprendizaje computacional, los enfoques semánticos y el enfoque híbrido.

- Aprendizaje computacional: Este consiste en la utilización de algoritmos mediante la colección de textos, en donde cada texto se puede presentar a través de un vector de palabras o *bag of words*, combinados con características semánticas que pretenden modelar la estructura sintáctica de una frase, su intensidad, negación, ironía o subjetividad. (Villena-Román, 2015) El autor (Conde Sánchez, 2017) menciona que la detención de polaridad se enfoca primordialmente en la identificación de dos categorías que son positiva y negativa.
- Enfoque semántico: El enfoque utiliza diccionarios de términos o lexicons, estos generalmente tienen una orientación semántica es de polaridad u opinión. Fundamentalmente estos sistemas procesan el texto y lo clasifican en palabras, eliminando aquellas que no tengan aportación significativa, posteriormente se compara el texto con los términos de lexicón para determinar la polaridad del texto, mediante la sumatoria de valores de polaridad de cada términos (Villena-Román, 2015), (Conde Sánchez, 2017).
- Enfoque híbrido: Corresponde a la combinación de los dos enfoques anteriores para generar un sistema eficiente e innovador (Conde Sanchez, 2017).

Es muy importante identificar que palabras o frases se clasifican como positivas o negativas es por ello, que es necesario definir el sentimiento si este es positivo o negativo. A continuación se muestran de oraciones donde se pueda evaluar la postura de la emoción. En los ejemplos se pueden detectar sentimientos negativos y positivos.

	Sentimientos positivos	Sentimiento negativos
Emoción	Yo estoy feliz	Yo estoy triste
Evaluación	¡Buena idea!	¡Mala idea!
Postura	Ella paga sus cuentas	Ella no paga sus cuentas

Tabla 3.- Ejemplo de sentimientos positivos y negativos

Existen diferentes herramientas para determinar la polaridad de un texto, estas ayudan a determinar en análisis a partir de connotaciones tanto positivas como negativas. Una de ellas es SentiWordNet (Conde Sánchez, 2017). Esta herramienta se crea a partir de

clasificadores cuyos métodos son semi-supervisados tienen la finalidad de determinar una polaridad previa. La polaridad previa consiste en detectar una característica previa directamente en relación con las palabras a analizar y la subjetividad tiene que ser alta. No obstante, esta característica no puede definir la polaridad total del texto de una manera general sino que solamente determinar la polaridad de las palabras (Villar & García, 2013), (Sánchez, 2017).

A continuación se explicará de manera detallada que es y cómo funcionan este tipo de herramientas:

- SentiWordNet es un recurso léxico en el que cada conjunto de palabras de WordNet está asociado a tres puntuaciones numéricas Obj (s), Pos (s) y Neg (s), que describe si son positivos o negativos los términos contenidos en el conjunto datos. Uno de los usos más sobresalientes SentiWordNet es enriquecer la representación de texto.

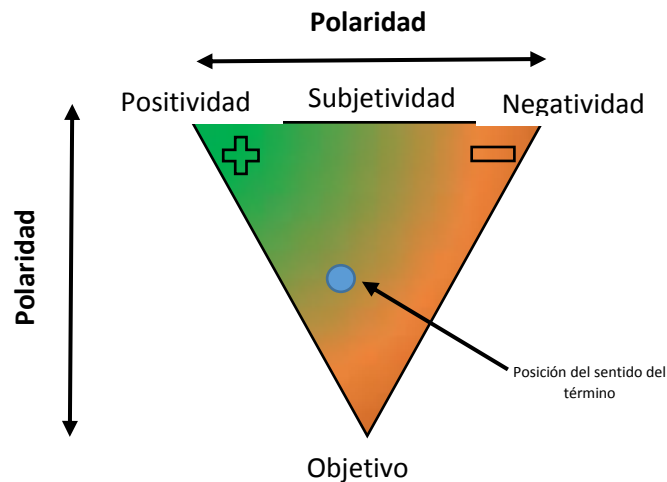


Ilustración 16.-Aplicación de SentiWordNet en Minería de Opinión

- Procesamiento del lenguaje natural a través de SentiWords ofrece un recurso que contiene aproximadamente 155,000 palabras en inglés. Las palabras se asocian con una puntuación de sentimiento incluida entre -1 y 1. Las palabras tienen enfoques de lematización en Python y también están alineadas con listas de WordNet que incluyen adjetivos, sustantivos, verbos y adverbios.
- WordStant es un diccionario de sentimiento que incluye más de 9,164 patrones de palabras negativas y 4,847 de palabras positivas, pero el sentimiento no se puede medir en estas dos listas. El sentimiento negativo se mide usando las siguientes

dos reglas en su lugar. La primera regla es palabras negativas no precedidas por una negación dentro de tres palabras en la misma oración. La segunda reglas de palabras positivas procedidas por una negación dentro de tres palabras en la misma oración. Las reglas para el sentimiento positivo son las mismas.

- SenticNet proporciona polaridad asociada con 50,000 conceptos de lenguaje natural. Una polaridad es un número flotante entre -1 y +1. Menos uno es la negatividad extrema y más uno es una positividad extrema.
- Consulta Lingüística y Word Count (LIWC) es un programa información para el análisis del lenguaje. Es compatible con inglés, alemán, español, holandés e italiano. Esta lista de palabras comerciales le permite extraer alrededor de 60 categorías de palabras diferentes, incluyendo emociones positivas, negativas, agresión, procesos afectivos, ansiedad, blasfemias, etc.

Etapa 4.- Validación

Una vez que se ha analizado los datos con cierto algoritmo implementando es muy importante la evaluación precisa de dichos algoritmos. Algunos autores mencionan que en el mundo de la investigación no siempre se reúnen esfuerzos para realizar una evaluación precisa y sistemática de las técnicas de medición estándar que han sido implementadas en dichas investigaciones. Muchos desarrolladores de algoritmos suelen proporcionar información de cómo evaluar sus algoritmos y aunque están por desarrollarse marcos y enfoques globales y genéricos para evaluar los diferentes algoritmos (Powers, 2011)

Al evaluar un algoritmo de clasificación de texto, hay dos errores comunes:

- 1) Falsos positivos: Todos aquellos que fueron etiquetados con cierta clase y no lo son.
- 2) Falsos negativos: Todos aquellos que fueron etiquetados como no perteneciente a la clase y en realidad, si lo son.

A continuación, se muestra una tabla donde se puede entender mejor lo anterior

	Correcto	No correcto
Seleccionado	Verdadero positivo	Falso positivo
No seleccionado	Verdadero negativo	Falso negativo

Tabla 4.- Selección de polaridad de términos.

Algunos autores como el caso de Elgueta Morales (2017) menciona que, para entender el rendimiento de los clasificadores se usan métricas para la recuperación de información y en diferentes investigaciones se describen de la siguiente manera:

- True Positives (*tp*): elementos a los que el clasificador asignó la clase relevante y esta era correcta.
- False Positives (*fp*): elementos a los que el clasificador asignó la clase relevante y esta no era correcta.
- False Negatives (*fn*): elementos a los que el clasificador asignó la clase no-relevante y esta no era correcta.
- True Negatives (*tn*): elementos a los que el clasificador asignó la clase no-relevante y esta era correcta.

Es importante que, para evaluar el rendimiento de los clasificadores utilizados en el proceso de análisis, se usen métricas utilizadas en recuperación de información, las cuales, muchas veces son adaptadas en función de los casos correctamente e incorrectamente clasificados.

La evaluación de clasificadores se puede interpretar como valores que dados según el documento estudiado y las coincidencias que muestre el clasificador sobre la efectividad del análisis, es posible analizar mediante el cálculo de rendimiento de acuerdo a las siguientes métricas:

- **Precision:** En ella se representa la porción de documentos que son clasificados correctamente para la clase. En otras palabras, son los textos correctamente clasificados sobre el total de textos clasificados, para una clase data.
- **Recall:** Esta métrica representa la porción de documentos de clase "A" que son clasificados correctamente. Es decir, los textos clasificados correctamente sobre el total de textos de esa categoría.
- **F-measure:** También llamada Medida- F, combina las medidas de precisión y Recall a partir de la media armónica pondera de estos dos valores.

La situación ideal es aquella en la que existe una precisión y exhaustividad alta (es decir muy cercana a 1). A esta situación se la denomina utilidad teórica. Con el objeto

de ponderar y ver cuán lejos están ambas medidas de la utilidad teórica, suelen emplearse los valores de ambas métricas combinadas en una media armónica denominada valor-F.

$$F_{\beta} = (1 + \beta^2) \cdot \frac{\text{Precisión} \cdot \text{Exhaustividad}}{(\beta^2 \cdot \text{Precisión}) + \text{Exhaustividad}}$$

Si β es igual a uno, se está dando la misma ponderación (o importancia) a la precisión que a la exhaustividad, si β es mayor que uno se le debe dar importancia a la exhaustividad mientras que si es menor que uno se le da importancia a la precisión.

$$\text{Precision} = \frac{TP}{TP + FP}$$

TP = True positive

TN = True negative

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

FP = False positive

FN = False negative

$$F1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

3.4 Minería de Opiniones con RapidMiner

El diseño de modelos predictivos se puede decir que es uno de los temas más relevantes en cuestión del análisis cuantitativo de datos, este está presente en diferentes plataformas basadas en redes neuronales hasta en modelos de generación de código de manera automatizada.

Actualmente, uno de los enfoques dirigido hacia sistemas computacionales es implementar estrategias viables que puedan analizar una gran cantidad de datos en poco tiempo y de manera efectiva, muchas de estas acciones están enfocadas en explorar el rastreo e interpretación de polaridad de opiniones en diferentes mercados.

Los softwares de Minería de Datos brindan diferentes ventajas como por ejemplo, son más simples y específicos al analizar grandes cantidades de datos. Además, utilizan pocos recursos y herramientas analíticas y estas, casi siempre están disponibles en las

mismas plataformas puesto que están diseñadas para realizar todas las funciones y tareas para para posteriormente analizar resultados.

Existen numerosos softwares y plataformas para el desarrollo de sofisticados modelos estadísticos y el análisis predictivo de grandes paquetes de datos: SAS, *Oracle Advanced Analytics*, KNIME, WEKA, IBM, *Neural Designer*, etc. En el caso específico de esta tesis se ha decidido trabajar con RapidMiner puesto que, este software permite preprocesar la base de datos, aplicar parámetros y operadores, así como, analizar los datos con diferentes algoritmos y por último pero no menos importante, está disponible para personal académico.

RapidMiner es un software el cual, tiene muchas herramientas de Minería de Datos ampliamente usadas y probadas a nivel internacional en aplicaciones empresariales, de gobierno y academia. Implementa más de 500 técnicas de pre-procesamiento de datos, modelación predictiva y descriptiva, métodos de prueba de modelos, visualización de datos, entre muchos otros. Los operadores disponibles cubren una inmensa variedad de tareas; desde conectividad con bases de datos y otras aplicaciones estadísticas hasta algoritmos avanzados de clasificación, segmentación, predicción, Text Mining y análisis de procesos. Otro de sus grandes atractivos es el ser una herramienta de fuente abierta o por su término en inglés “*open source*” muy empleada en el mundo académico y gratuita para usos no comerciales.

A continuación se muestra una imagen del software, visto de primera estancia y listo para utilizarse:

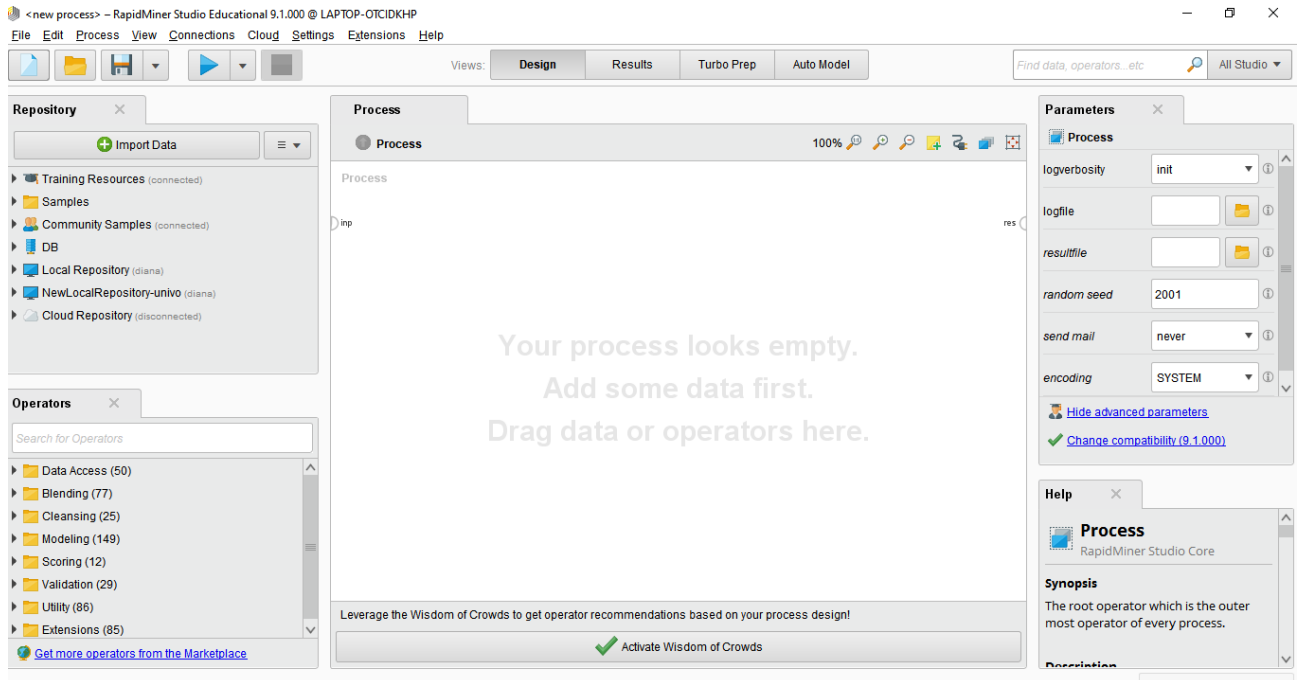


Ilustración 17.- Programa RapidMiner Studio, versión 9.1

3.4.1 Características

La versión de RapidMiner con la que se trabajara actualmente en esta tesis es la versión Studio 9.1. Se decidió trabajar con este software puesto que se puede trabajar con temáticas referentes a ciencias computacionales como Minería de Datos y Minería Web social sin necesidad de tener conocimientos previos de programación puesto que, el software es fácil, intuitivo y permite aprender a usarlo a través de diferentes tutoriales. Cabe mencionar que este software tiene una versión académica misma que permite un fácil acceso y uso del mismo.

Además, diferentes características integran a la herramienta de software RapidMiner, autores como (Beltran; Poveda, 2010) mencionan a continuación las más sobresalientes:

- Desarrollado en Java
- Multiplataforma
- Representación interna de los procesos de análisis de datos en ficheros XML
- Permite el desarrollo de programas a través de un lenguaje de script
- Puede usarse de diversas maneras:
 - a. A través de un GUI
 - b. En línea de comandos

- c. En *batch*
 - d. Desde otros programas a través de llamadas a sus bibliotecas
- Extensible
 - Incluye gráficos y herramientas de visualización de datos

3.4.1 Construcción del modelo

3.4.1.1 Operadores

RapidMiner para la ejecución de todas sus funcionalidades pone a disposición del usuario un set de operadores, mismo que están disponibles para el tratamiento de datos.

- Módulo *Import*: Este módulo contiene operadores de importación y están categorizados en las siguientes secciones:
 1. *Data Access* (50)
 2. *Blending* (77)
 3. *Cleansing* (25)
 4. *Modeling* (149)
 5. *Scoring* (12)
 6. *Validation* (29)
 7. *Utility* (86)
 8. *Extensions* (85)

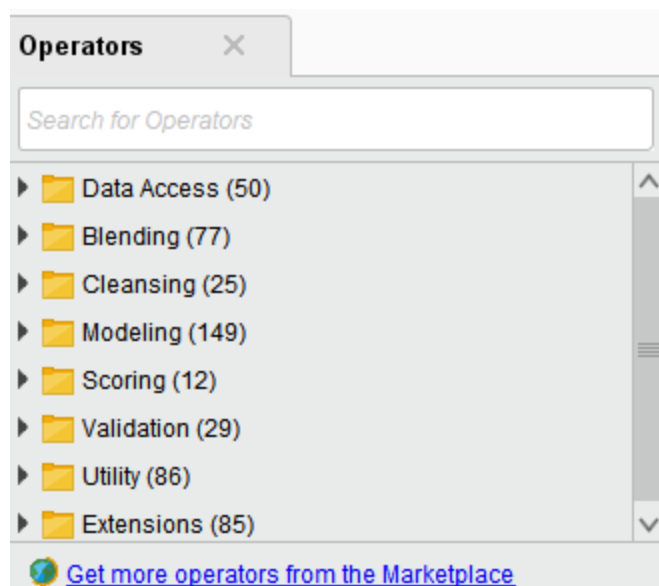


Ilustración 18.- Operadores en RapidMiner, versión 9.

Estos operadores pueden ser ubicados dependiendo del área de trabajo y análisis que desea realizar y después, a partir de esto, se debe llamar el objeto de trabajo es decir, la base de datos, un algoritmo de modelo, especificación de funciones y parámetros etc.

Al analizar el sentimiento de datos recopilados en un archivo CSV, se deben de comprender los aspectos principales mencionados y el tipo de sentimiento (positivo, negativo o neutral) para ello, se debe de realizar un análisis de correlación sobre las palabras y los aspectos para posteriormente, visualizar resultados.

En el caso de la construcción del modelo para el Análisis de Sentimientos se deben de tomar en cuenta el siguiente procedimiento:

- **Obtención de datos:** Estos datos se pueden obtener de diferentes fuentes y estas pueden ser automatizadas y no automatizadas, como se explicó al principio de este capítulo en el subtema **3.2.1 Contenido web y estructura de la Minería Web** y **3.2.2 Minería de Uso Web (Análisis Web)**. En el caso del análisis realizado en esta tesis, se cuenta con una base de datos en Excel, misma que se puede trabajar en RapidMiner con la función de Crear Documento.
- **Preparación de datos:** Las técnicas de preprocesamiento de datos que provee RapidMiner en la mayoría de análisis es la transformación *Numerical to Binominal*. El operador numérico a binominal cambia el tipo de atributos numéricos a un tipo binominal (también llamado binario). Este operador no solo cambia el tipo de atributos seleccionados sino que también asigna todos los valores de estos atributos a los valores binominales correspondientes. Los atributos binominales pueden tener solo dos valores posibles, es decir, 'verdadero' o 'falso'. Si el valor de un atributo está entre el valor mínimo y máximo especificado, se convierte en 'falso', de lo contrario, 'verdadero'. Los valores mínimos y máximos se pueden especificar mediante los parámetros min y max respectivamente. Los límites predeterminados están establecidos en 0.0, por lo tanto, solo 0.0 se asigna a 'falso' y todos los demás valores se asignan a 'verdadero' de forma predeterminada. En el panel Operadores, se debe explorar y seleccionar este operador.
- **Entrenamiento:** Este paso depende muchas veces de otros operadores como por ejemplo: *Set Role*, *Nominal to Text* y *Process Documents*.

- Validación: Este operador realiza una validación simple, es decir, divide aleatoriamente los datos en un conjunto de entrenamiento y un conjunto de prueba y evalúa el modelo. Este operador realiza una validación dividida para estimar el rendimiento de un operador de aprendizaje (generalmente en conjuntos de datos no vistos). Se utiliza principalmente para estimar con qué precisión un modelo (aprendido por un operador de aprendizaje particular) se desempeñará en la práctica.
- Evaluación: La evaluación en RapidMiner se puede realizar con el operador de *Cross Validation*. Se utiliza principalmente para estimar con qué precisión un modelo (aprendido por un operador de aprendizaje en particular) se desempeñará en la práctica. El operador de validación cruzada es un operador anidado. Tiene dos subprocesos: un subproceso de capacitación y un subproceso de prueba. El subproceso *Training* se usa para entrenar un modelo. El modelo entrenado se aplica luego en el subproceso de prueba. El rendimiento del modelo se mide durante la fase de prueba.

La entrada de datos se divide en k subconjuntos de igual tamaño. De los k subconjuntos, un solo subconjunto se retiene como el conjunto de datos de prueba (es decir, la entrada del subproceso de prueba). Los restantes subconjuntos $k - 1$ se utilizan como conjunto de datos de entrenamiento (es decir, entrada del subproceso de entrenamiento). El proceso de validación cruzada se repite k veces, con cada uno de los k subconjuntos utilizados exactamente una vez como datos de prueba. Los k resultados de las k iteraciones se promedian (o combinan) para producir una sola estimación. El valor k se puede ajustar utilizando un parámetro numérico. La evaluación del rendimiento de un modelo en conjuntos de pruebas independientes arroja una buena estimación del rendimiento en conjuntos de datos no vistos.

- Aplicación del modelo: El operador debe de entrenar con un modelo primero en un *ExampleSet*, que a menudo es un algoritmo de aprendizaje. Posteriormente, este modelo se puede aplicar en otro *ExampleSet*. Por lo general, el objetivo es obtener una predicción sobre datos no vistos o transformar datos aplicando un modelo de preprocesamiento. El conjunto de ejemplos sobre el que se aplica el modelo debe ser compatible con los atributos del modelo. Esto significa que el *ExampleSet* tiene el mismo número, orden, tipo y función de Atributos que el *ExampleSet* utilizado para generar el modelo.

A continuación se muestra un modelo generalizado de análisis de datos en RapidMiner.

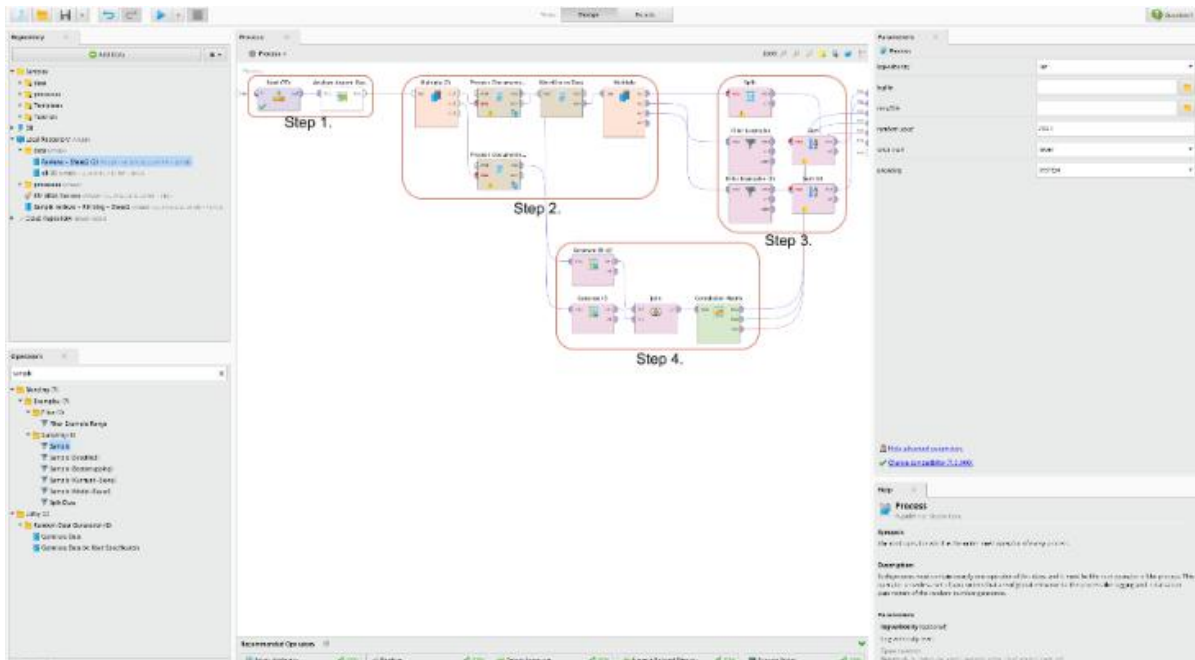


Ilustración 19.-Procedimiento de análisis de datos en RapidMiner

3.5 Conclusión

Este capítulo desarrolló el marco teórico en el que se posiciona la tesis: Minería de Opiniones. En él, se resaltan temas importantes para el entendimiento de esta temáticos en la cual, se explica que la Minería de Texto es parte de la Minería Web y que en los últimos años estas trabajan en conjunto puesto que, el crecimiento del Internet deriva diversas fuentes de tráfico.

Muchos de los métodos y técnicas para extraer la información y procesarla pero, se debe de tener en cuenta que la decisión de cuál ocupar, siempre dependerá de los intereses de la organización y también si estos métodos son eficientes para el tipo de análisis que se desea realizar. Bajo este contexto, la Minería de Textos surge como una herramienta que permite conocer la opinion de un producto y/o servicio en las diferentes plataformas

existentes en Internet puesto que, esta permite el tratamiento de texto, carente de una estructura definida.

Esta herramienta puede ser de gran utilidad para la evaluación de diversos indicadores dentro de toda una variedad de datos para entender mejor la opinión de las personas como por ejemplo: se puede detectar información crítica que pueda afectar a una organización, mejorar la experiencia de un usuario, lanzar nuevos productos y servicios y por supuesto, analizar al mercado. Los efectos de aplicación referentes a esta tesis, los datos analizados con la Minería de Opiniones representa la información necesaria para el desarrollo de estrategias digitales para una empresa de servicios.

Capítulo 4: Aplicación de un modelo de Minería de Web Social utilizando Marketing Digital y *Hashtags*

4.1 Introducción

En este capítulo se explica de manera detallada el desarrollo de la metodología propuesta anteriormente en donde se integran dos áreas de oportunidad y de gran impacto para las organizaciones de hoy en día: Marketing Digital y Minería Web Social, específicamente, bajo el contexto de Minería de Opiniones. La integración de estas dos temáticas, implicó el desarrollo de distintas metodologías, técnicas y prácticas enfocadas en la creación y evaluación de una campaña *online* y que a partir de detectar la polaridad de las opiniones, se generen estrategias factibles que estén basadas en parámetros e indicadores confiables los cuales, tienen métricas de medición y cuantificación para retroalimentar el Marketing Digital de la organización.

Este capítulo se divide en tres etapas; en la primera se muestra la realización del plan de Marketing Digital mismo que al ser implementado permitió aumentar el tráfico y la actividad en las redes sociales pertenecientes a la organización a través de una campaña digital y así, generar opiniones de los internautas que siguen dichas plataformas digitales. Posteriormente, la extracción de estos datos, se pudo realizar a través de un rastreador de etiquetas en línea (*hashtracker online*). Una vez que se obtuvieron los datos, se desarrolló el modelo de Minería de Opiniones en el software RapidMiner para detectar la polaridad de los comentarios generados en dicha campaña y conforme a esto, realizar las estrategias y recomendaciones de Marketing Digital y comunicación para el grupo educativo.

Es muy importante mencionar que el uso de este modelo de Minería de Opiniones busca evaluar de manera específica el impacto conforme a positividad de dicho plan de Marketing Digital. Esta herramienta es útil para medir el alcance de la campaña de Marketing Digital implementada puesto que, al detectar la polaridad de cada opinión permite un mayor conocimiento conforme a la aceptación y reacción de los internautas hacia las publicaciones. Esto representa una ventaja competitiva en pertinencia a la toma de decisiones de la organización orientadas a la planificación estratégica del plan de Marketing Digital y comunicación.

4.2 Plan de Marketing Digital

4.2.1 Plan Integral de Estrategias para el cumplimiento de metas institucionales del Grupo Educativo Univo

Grupo Educativo Univo cuenta con una oferta educativa medianamente diversificada en tres áreas del conocimiento, lo que permite establecer tres facultades: Económico-Administrativa, Humanidad e Ingeniería. Las tendencias actuales han impulsado que el grupo educativo se mantenga en constante actualización conforme a tendencias que permitan estructurar procesos internos productivos y flexibles.

Desde sus inicios Grupo Educativo Univo se ha caracterizado por implementar estrategias innovadoras las cuales, se pueden difundir a través de medios internos y externos.

A su vez, el interés del grupo educativo radica en la generación y aplicación del conocimiento para que este dé respuesta a las necesidades para el desarrollo humano de la sociedad. Lo anterior, se puede lograr con la mejora continua de su modelo educativo, mismo que cubra las necesidades del entorno.

Es importante mencionar que el modelo educativo de Grupo Educativo Univo se fundamenta de la visión, misión y decálogo de valores del grupo educativo mismo que se mencionan a continuación:

Misión

Grupo Educativo Univo, está ubicado en la Zona Centro del Estado de Veracruz, es una institución privada que ofrece educación superior, con el objetivo de formar profesionales con conocimientos actualizados, habilidades y actitudes; que les permitan analizar, innovar, emprender y dirigir organizacionales nacionales y extranjeras, comprometidos con la mejora continua y con el alto sentido de responsabilidad ambiental y social.

Considerando como ejes rectorales la educación continua de sus cuerpos académicos, la investigación privada, el uso de nuevas tecnologías, la vinculación con el sector productivo y de servicios, así como el fomento de las relaciones internacionales.

Visión 2020

Consolidarse como un grupo educativo con reconocimiento a nivel nacional por su prestigio académico y sus relaciones internacionales, fundamentando en la operación de

un modelo educativo consistente y con participación social, la acreditación de sus programas académicos y la certificación de sus procesos

Decálogo de valores

- Ser Comprensivo
- Ser Honesto
- Ser Incluyente
- Ser Justo
- Ser Leal
- Ser Libre
- Ser Íntegro
- Ser Responsable
- Ser Solidario
- Ser Tolerante

Así mismo, para lograr los objetivos estratégicos se deben implementar acciones enfocadas en Cuatro Ejes Rectorales:

- Generación y aplicación del conocimiento
- Visión Social
- Internacionalización y Calidad
- Gestión y gobierno

El Grupo Educativo Univo cuenta con indicadores y metas internas divididas en las siguientes áreas: personal académico, plan de estudios, investigación, estudiantes, formación integral, vinculación y extensión, calidad académica, movilidad internacional, infraestructura y gestión administrativa y financiamiento.

Conforme a lo anterior, se estructura del Plan Integral de Estrategias para el cumplimiento institucional de Grupo Educativo Univo y de este plan se deriva el plan de comunicación y el plan de Marketing y en su defecto, de este último el plan de Marketing Digital. A continuación, se explica de manera gráfica:

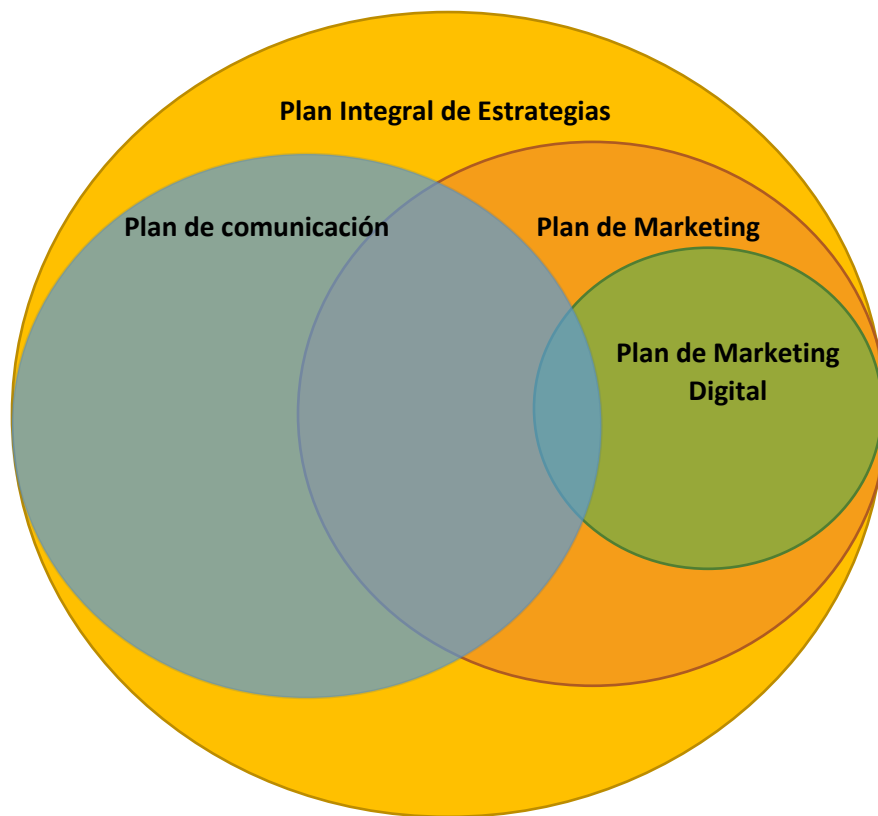


Ilustración 20.- Integración e interacción del plan Integral de estrategias, plan de comunicación, plan de Marketing y plan de Marketing digital de Grupo Educativo Univo.

4.3 Plan de Marketing Digital

El Plan de Marketing de Grupo Educativo Univo se llevó a cabo conforme a las cinco etapas del modelo SOSTAC® (PR Smith, 1990): Situación, Objetivos, Estrategias, Tácticas y Acciones.

4.3.1 Análisis de situación

El análisis de la situación es la primera parte del plan de Marketing Digital y se analizó tanto internamente como externamente. El análisis consistió en realizar un análisis de los competidores directos e indirectos así como, un análisis FODA (Fortalezas, Oportunidades, Debilidades y Amenazas) del Marketing Digital de la organización.

4.3.1.1 Análisis de competencia

El posicionamiento de Grupo Educativo Univo está en constante cambio puesto que, al ser una industria enfocada a los servicios educativos su campo es muy competitivo y la amenaza de competidores directos e indirectos es muy alta.. A continuación, se muestra un análisis de competencia del grupo educativo, en este se determinaron cuáles son sus productos y servicios que ofrecen, su desempeño en el mercado, los canales de comunicación utilizados y sus estrategias de Marketing de la competencia. Esta está dividida por pública y privada:

Escuela	Servicios que ofrece	Desempeño en el mercado (% de población en redes sociales)	Canales de comunicación	Estrategias de comunicación
Grupo Educativo Univo	Primaria, secundaria, preparatoria, licenciaturas y posgrado	Grupo Educativo Univo cuenta con una comunidad de 8,831 internautas mismos que, representan aproximadamente el 0.15% de la población que se encuentra en el rango de edad (10– 60 años), el cual es representada por 5, 730,283 adolescentes, hombres y mujeres de Veracruz según INEGI, 2017.	Redes sociales (Facebook, Twitter, Instagram, Youtube), página web, correo institucional, teléfono fijo y oficinas.	Marketing Tradicional (banners, lonas) y Marketing Digital (publicaciones y dinámicas en redes sociales)

Tabla 5.- Análisis Grupo Educativo Univo.

Competencia Directa				
Escuela	Servicios que ofrece	Desempeño en el mercado (% de población en redes sociales: Facebook)	Canales de comunicación	Estrategias de Marketing
UGM	Kínder, Primaria, Secundaria, Bachillerato, licenciaturas, ingenierías, maestrías	UGM cuenta con una comunidad de 15,267 internautas mismos que, representan aproximadamente el 0.26% de la población que se encuentra en el rango de edad (10– 60 años), el cual es representada por 5, 730,283 adolescentes, hombres y mujeres de Veracruz según INEGI, 2017.	Redes sociales (Facebook, Twitter, Instagram, YouTube), página web, correo institucional, teléfono fijo y oficinas.	Marketing tradicional (banners, lonas, BT)L y Marketing Digital (publicaciones y dinámicas en redes sociales)
Idea de Oriente	Licenciaturas y maestría	Idea de Oriente cuenta con una comunidad de 2,201 internautas mismos que, representan aproximadamente el 0.003% de la población que se encuentra en el rango de edad (10– 60 años), el cual es representada por 5, 730,283 adolescentes, hombres y mujeres de Veracruz según INEGI, 2017.	Redes sociales (Facebook, Twitter, Instagram), página web, correo institucional, teléfono fijo y oficinas.	Marketing Tradicional (banners, lonas, BT)L y Marketing Digital (publicaciones redes sociales: no tienen continuidad, ni <i>engagement</i>)
Universidad Sotavento	Licenciaturas, posgrado y doctorado	Sotavento con una comunidad de 8,457 internautas mismos que, representan aproximadamente el 0.14% de la población que se encuentra en el rango de edad (10– 60 años), el cual es representada por 5, 730,283 adolescentes, hombres y mujeres de Veracruz según INEGI, 2017.	Redes sociales (Facebook, Twitter, Instagram), página web, correo institucional, teléfono fijo y oficinas.	Marketing Tradicional (banners, lonas, BT)L y Marketing Digital (publicaciones y dinámicas en redes sociales)

Tabla 6.- Análisis de competencia directa de Grupo Educativo Univo.

Competencia Indirecta				
Escuela	Servicios que ofrece	Desempeño en el mercado (% de población en redes sociales: Facebook)	Estrategias de Marketing	Estrategias de comunicación
U.T.C.V	Licenciaturas	U.T.C.V con una comunidad de 28,531 internautas mismos que, representan aproximadamente el 0.49% de la población que se encuentra en el rango de edad (10– 60 años), el cual es representada por 5, 730,283 adolescentes, hombres y mujeres de Veracruz según INEGI, 2017.	Redes sociales (Facebook, Twitter, Instagram), página web, correo institucional, teléfono fijo y oficinas.	Marketing tradicional (banners, lonas, BT)L y Marketing Digital (publicaciones y dinámicas en redes sociales)
I.T.S.Z	Licenciaturas	I.T.S.Z con una comunidad de 9,996 internautas mismos que, representan aproximadamente el 0.17% de la población que se encuentra en el rango de edad (10– 60 años), el cual es representada por 5, 730,283 adolescentes, hombres y mujeres de Veracruz según INEGI, 2017.	Redes sociales (Facebook, Twitter, Instagram), página web, correo institucional, teléfono fijo y oficinas.	Marketing Tradicional (banners, lonas, BT)L y Marketing Digital (publicaciones y dinámicas en redes sociales)
U.V	Licenciaturas , maestrías y doctorados	I.T.S.Z con una comunidad de 9,996 internautas mismos que, representan aproximadamente el 0.006% de la población que se encuentra en el rango de edad (10– 60 años), el cual es representada por 5, 730,283 adolescentes, hombres y mujeres de Veracruz según INEGI, 2017.	Facebook, página web, correo institucional, teléfono fijo y oficinas.	Publicaciones institucionales en redes sociales.
I.T.O	Licenciaturas, maestrías y doctorados	I.T.S.Z con una comunidad de 36,224 internautas mismos que, representan aproximadamente el 0.63% de la población que se encuentra en el rango de edad (10– 60 años), el cual es representada por 5, 730,283 adolescentes, hombres y mujeres de Veracruz según INEGI, 2017.	Facebook, página web, correo institucional, teléfono fijo y oficinas.	Publicaciones institucionales en redes sociales.

Tabla 7.-Análisis de competencia indirecta de Grupo Educativo Univo.

Al analizar a la competencia se pueden tener diferentes ventajas competitivas en diferentes áreas de oportunidad, una de esta es la mejora de productos y servicios ya que, al identificar que ofrece la competencia en el caso del grupo educativo se pueden comparar las carreras ofrecidas y los planes de estudios de las instituciones. Por otra parte, también se puede conocer mejor al cliente ya que, a través de estos se pueden conocer e identificar las debilidades de la competencia e identificar las fortalezas en la oferta de la propia institución. Así mismo, permite motorizar las actividades y acciones que se están realizando para captar un mayor número de clientes y de esta manera mejorar las estrategias y su planificación.

4.3.1.2 Análisis FODA

Para llevar a cabo el conocimiento del Marketing Digital de la organización, se presenta el siguiente análisis FODA, con el fin de ubicar fortalezas, oportunidades, debilidades y amenazas que rodean a las redes sociales de la organización. Por otra parte, la información recabada en el presente análisis, brinda un punto de partida para realizar futuras propuestas para redireccionar las estrategias de Marketing Digital:

Fortalezas	Oportunidades
<ul style="list-style-type: none"> 1.- Marca existente y conocida. 2.- Continuo <i>branding</i>. 3.- Mayor audiencia y poca inversión 4.- Comunicación a escala. 5.- Equipo de Marketing y comunicación pequeño (ahorro de dinero, tiempo y mano de obra). 6.- Presencia en diferentes redes sociales: Facebook, Instagram, Twitter, Blog, y página web. 7.- Publicaciones constantes. 	<ul style="list-style-type: none"> 1.- Alcance de marca. 2.- Tener bases de datos. 3.- Nuevos servicios: Cursos y Licenciaturas en línea. 4.- Promoción innovadora en medios digitales. 5.- Uso y mantenimiento de redes.
<ul style="list-style-type: none"> 1.- Tendencias y tecnología en continuo cambio. 2.- Comprensión de comportamiento y necesidades de consumidor, cliente usuario. 3.- Análisis de datos adecuadamente. 4.- Almacenamiento de datos de manera segura 5.- Poco <i>engagement</i> y tráfico en plataformas digitales. 6.- Las estrategias de Marketing Digital están enfocadas a Facebook y no cubre todas las redes. 	<ul style="list-style-type: none"> 1.- Crecimiento de la <i>networking</i> enfocada cada vez en sentimientos y opiniones. 2.- Competencia educativa usando las mismas plataformas digitales. 3.- Estrategias de posting de competidores muy parecidas a las del grupo educativo.
Debilidades	Amenazas

Ilustración 21.-Análisis F.O.D.A del Marketing Digital de Grupo Educativo Univo.

4.3.1.3 Objetivos estratégicos

Los objetivos estratégicos del plan de Marketing Digital se derivaron de la aplicación de un análisis CAME (Corregir, Afrontar, Mantener y Explorar) al FODA (Fortalezas, Oportunidades, Debilidades y Amenazas):

Fortalezas: Mantener	Debilidad: Corregir
<ol style="list-style-type: none"> 1 Crear un plan de <i>branding</i> de marca madre cada seis meses. 2 Crear un plan de Marketing Digital para las diferentes redes sociales de la organización cada seis meses. 3 Planificar las publicaciones que se realizaran cada de semanas divididas en eventos y fechas importantes para la organización. 	<ol style="list-style-type: none"> 1 Realizar un estudio de Minería de Opiniones y sentimientos en redes sociales para saber que piensa el internauta, cliente y consumidor sobre la organización en plataformas digitales en un plazo de un año. 2 Implementar publicaciones y dinámicas innovadoras siguiendo tendencias y temas de interés cada semana en las redes sociales de la organización.
Oportunidad: Explorar	: Afrontar
<ol style="list-style-type: none"> 1 Utilizar base de datos para fundamentar las estrategias y dinámicas de Marketing Digital que se implementarán en las redes sociales. 2 Gestionar el lanzamiento de un apartado en la página web para capacitación <i>online</i>. 3 Realizar publicaciones diarias. 	<ol style="list-style-type: none"> 1 Realizar publicaciones enfocadas en tendencia digital y <i>engagement</i> del usuario para incrementar el tráfico en la red social en un plazo de seis meses. 2 Cuantificar los datos obtenidos de la recolección de información de las plataformas digitales de organización.

Ilustración 22.- Análisis C.A.M.E conforme a objetivos estratégicos del plan de Marketing Digital de Grupo Educativo Univo.

Además, los objetivos se clasificaron conforme a los siguientes beneficios de Marketing Digital:

- *Sell*- Vender
- *Serve*- Valor agregado
- *Speak*- Opinion

- Save- Ahorro
- Sizzle- Diversificación

A continuación se presenta dicho agrupamiento:



Ilustración 23.- Clasificación de objetivos conforme a beneficios de Marketing Digital de Grupo Educativo Univo.

4.3.1.4 Plan de Marketing Digital: estrategias y tácticas

La estrategia de Marketing Digital se debe centrar en las acciones que se realizarán en línea (Chaffey & Smith, 2013). Idealmente, la estrategia debe estar basada en los objetivos y en los recursos disponibles, así como, en el contexto de las estrategias de la competencia, tanto existentes y potenciales).

La estrategia de Marketing Digital debe estar compuesta por los siguientes elementos:

- Segmentos
- Mercado objetivo
- Objetivos
- Posicionamiento
- Secuencia
- Adquisición y retención de clientes
- Herramientas tácticas
- Integración
- Redes sociales (incluye estrategia de contenidos)

Así mismo, se deben de tomar en cuenta los indicadores clave de rendimiento (KPI'S) y en el caso específico de Grupo Educativo Univo se seleccionaron los que conforman alcance e influencia: *Share Voice* (Publicidad boca en boca), *Sentiment* (Opiniones y Sentimientos), *Network Size* (Crecimiento de la red social) y *Engagement* (compromiso-“enganchamiento”).

Para determinar cuál indicador de rendimiento es más importante para el grupo educativo y a partir de esto, estructurar el plan se realizó un Proceso de Análisis Jerárquico junto con al equipo de Marketing y comunicación de esta institución educativa. El proceso de Análisis Jerárquico o (AHP) por sus siglas en inglés fue diseñado por Thomas L. Saaty y este se enfoca en resolver problemas de criterios múltiples y el resultado del AHP es una jerarquización con prioridades que muestran la preferencia general para cada una de las alternativas de decisión.

A continuación se muestra dicho análisis:

- Asignación de importancia: Las comparaciones pareadas son bases fundamentales para el AHP, puesto que este utiliza una escala subyacente con valores del 1 al 9 para calificar preferencias relativas de los elementos. Se presentan las calificaciones numéricas que se recomiendan para las preferencias verbales expresadas por el equipo de Marketing.
- Conforme a los indicadores de Marketing digital se establecieron las siguientes calificaciones numéricas para cada uno:

Asignación de importancia	
	Importancia
Share Voice (SV)	3
Sentiment (S)	2
Network size and growth (NG)	4
Engagement (EN)	1

Tabla 8.- Asignación de importancia

no muy importante	1
Ligeramente importante	3
Más importante	5
Muy importante	7
Fundamental	9

Tabla 9.- Escala de clasificación del 1 al 9 conforme a atributos.

- Matriz de comparación: Una Matriz cuadrada contiene comparaciones pareadas de alternativas o criterios. Una vez que se calcula lo que denomina como prioridad de cada uno de los elementos que se comparan. A esta parte se le conoce como sintetización. El proceso matemático preciso se requiere para realizar tal sintetización implica el cálculo de valores y vectores característicos.

	SV	S	NG	EN
SV	1	1/3	3	1/9
S	3	1	7	1/3
NG	1/3	1/7	1	1/5
EN	9	3	5	1
SUMA	13 1/3	4 1/2	16	1 2/3

Tabla 10.- Consenso de clasificación conforme a variables.

- Matriz de prioridades: Se consideran las prioridades para cada criterio en término. Se denomina matriz de prioridades a la que resume las prioridades para cada alternativa en cada criterio, en otras palabras los valores deben de estar normalizados.

	SV	S	NG	EN	Suma	Wi	%
SV	0.08	0.07	0.19	0.07	0.40	0.10	10%
S	0.23	0.22	0.44	0.20	1.09	0.27	27%
N	0.03	0.03	0.06	0.12	0.24	0.06	06%
EN	0.68	0.67	0.31	0.61	2.27	0.57	57%
SUMA	1.00	1.00	1.00	1.00	4.00	1.00	100%

Tabla 11.- Normalización de valores

Comprobación	
Landa MAX	4.46
RCI	0.02
IC	0.02

Tabla 12.- Comprobación

El AHP ofrece un método para mediar un grado de consistencia entre los indicadores seleccionados por el equipo de Marketing. Si el grado de consistencia es aceptable, puede continuarse el proceso de decisión. Se calcula la razón de consistencia (RCI). Esta razón está diseñada de manera que los valores que exceden de 0.10 son señales de juicios inconsistentes, es probable que en estos casos se deban reconsiderar los valores asignados en el principio y modificarlos. En el caso de la comprobación del AHP de los indicadores de Marketing Digital es aceptable ya que es de 0.02.

Conforme a lo anterior, se determinó que el indicador de alcance e influencia en el cuál, la institución educativa enfocara esfuerzos es: *Engagement* (Compromiso) puesto que este obtuvo un porcentaje de importancia del 57% y posteriormente, Sentiment (Opinion y Sentimientos de sus internautas y clientes) con un porcentaje de 27%.

Una vez haber analizado el ambiente externo e interno de Grupo Educativo Univo conforme a redes sociales, se deben elaborar los objetivos específicos conforme a un análisis CAME y clasificar a los mismos junto con los beneficios de Marketing Digital conforme a estrategia, además de obtener a relevancia de indicadores clave de rendimiento de alcance e influencia, el plan de Marketing Digital de esta institución educativa. La orientación estará enfocada primordialmente a conocer el segmento y mercado objetivo es decir, conocer profundamente que piensan los clientes, consumidores e internautas del grupo educativo a través de las redes sociales de misma, además de implementar herramientas y acciones dinámicas para integrar a los usuarios y posicionar a Grupo Educativo Univo en cada una de sus redes sociales, generando contenido de interés para sus internautas:

Acciones y tácticas	2019												Control de estrategia
	EN	FE	MA	AB	MA	JUN	JUL	AG	SE	OC	NOV	DI	
Sitio web y redes sociales	x	x	x	x	x	x	x	x	x	x	x	x	Publicaciones diarias y actualización de información
Minería de Opinión		x	x	x	x								Se realizó una dinámica en línea para obtener datos sobre las opiniones de los internautas en Facebook
							x	x		x			Se realizaran encuestas en el mercado objetivo sobre las redes sociales de universidad
Publicidad	x	x	x	x	x	x	x	x	x	x	x	x	Banner publicitario en redes sociales y anuncio pagado
Relaciones publicas	x	x	x	x	x	x	x	x	x	x	x	x	Noticias sobre licenciaturas, vida interna universitaria y nuevas publicaciones
Correo directo	x	x	x	x	x	x	x	x	x	x	x	x	Boletines electrónicos, información sobre licenciaturas, promociones, descuentos (fortalezas, becas, periodos de inscripción)
SISE	x	x	x	x	x	x	x	x	x	x	x	x	Actualización de información para alumnos, cuerpo docente y administrativo
Promociones de venta	x					x	x	x				x	Acción en línea en correos electrónicos, sitios web y redes sociales
Exhibiciones y eventos	x	x	x	x	x	x	x	x	x	x	x	x	Exhibiciones virtuales y transmisión de eventos en línea

Tabla 13.- Plan de Marketing Digital de Grupo Educativo Univo

4.3.1.5 Aplicación de la Minería de Web Social y Minería de Opinión en plataformas digitales

La presente investigación se enfocara en las acciones de Minería de Web Social y Minería de Opinion mismas que, son parte fundamental para obtener datos de plataformas digitales del grupo educativo.

Como primera acción, se planificar estrategias y acciones enfocadas en el objetivo: **“Realizar un estudio de Minería de Opiniones y sentimientos en nuestras redes sociales para saber que se piensa el internauta, cliente y consumidor sobre la organización en plataformas digitales en un plazo de un año”**

Las acciones de este objetivo están marcadas por tres etapas:

- Implementación de estrategias digitales
- Recolección de datos
- Análisis de datos

Las estrategias digitales se planificaron en enero 2019 junto con el equipo de Marketing y comunicación y se implementaron de febrero a noviembre del mismo año. Es importante mencionar que las estrategias están compuestas por tres elementos: dinámicas, cuantificables y medibles.

A continuación se muestra el calendario de las acciones y actividades que se realizaron para obtener y extraer las opiniones:

Acción o actividad	Temática	Mes	Descripción
Trivia 1	Trivia: Piloto	Febrero	Se realizaron Trivias de participación en el Facebook del Grupo Educativo Univo con la intención de crear contenido dinámico y aumentar el tráfico así como, el <i>engagement</i> de los clientes e internautas. Las trivias tienen que ir acompañadas de etiquetas de texto para que las opiniones de los usuarios puedan ser rastreadas
Trivia 2	Día de la mujer	Marzo	
Trivia 3	Semana deportiva	Abril	
Trivia 4	Día del niño (primaria)	Abril	
Trivia 5	Día del niño (secundaria)	Abril	

Videos	Campaña “El camino al éxito”: Oferta educativa y experiencia como egresado	Mayo	Los videos que se realizaron de la campaña “UniVO: El camino al éxito” tenían la intención de generar opiniones y <i>engagement</i> entre los internautas.
Encuesta en Google	Valoración de redes sociales	Mayo- Noviembre	Las encuestas se realizaron en el formulario de google con la intención de entender la perspectiva que tienen los consumidores de las redes sociales del grupo educativo.
Celebración del 31 Aniversario UniVO	Aniversario UniVO	Septiembre	Estas actividades fueron realizadas dentro de la institución pero en sus redes sociales se compartió el evento lo cual, genero diferentes opiniones.
Participación en desfile del 15 de septiembre	Actividad enfocada en relaciones públicas de la institución	Septiembre	La participación en actividades enfocadas a las relaciones públicas siempre han generado opiniones en las redes sociales de la institución.
Concurso calaveritas literarias		Octubre	
Festival de Día de muertos		Noviembre	
Exposición fotográfica “Nocturno”		Noviembre	
UniVoz		Noviembre	
Fashion Brand		Noviembre	

Tabla 14.- Calendario de las acciones y actividades que se realizaron.

La extracción de datos se llevó a cabo a través del rastreador de etiquetas (*hashtags*) en línea: *Social Searcher (hashtag-tracker)*. Esta herramienta sirve como un motor de búsqueda mismo que, permite de manera directa monitorear todas las menciones sociales públicas en redes sociales y web. Mide y rastrea rápidamente lo que la gente dice sobre

su empresa, marca, producto o servicio en un panel de control fácil de usar, que agilizaría sus esfuerzos de seguimiento de presencia en línea y le ahorra mucho tiempo.

De dichas actividades se obtuvo un total de 411 comentarios extraídos de las diferentes plataformas digitales de Grupo Educativo Univo, mismos que posteriormente fueron analizados con el modelo de Minería de Opinión en RapidMiner.

4.4 Aplicación de la Minería de Opiniones

4.4.1 Modelo para la Minería de Opiniones

Dentro del modelo para la Minería de Opiniones se desarrollan cinco principales pasos mismos que estarán esquematizados en la siguiente ilustración:

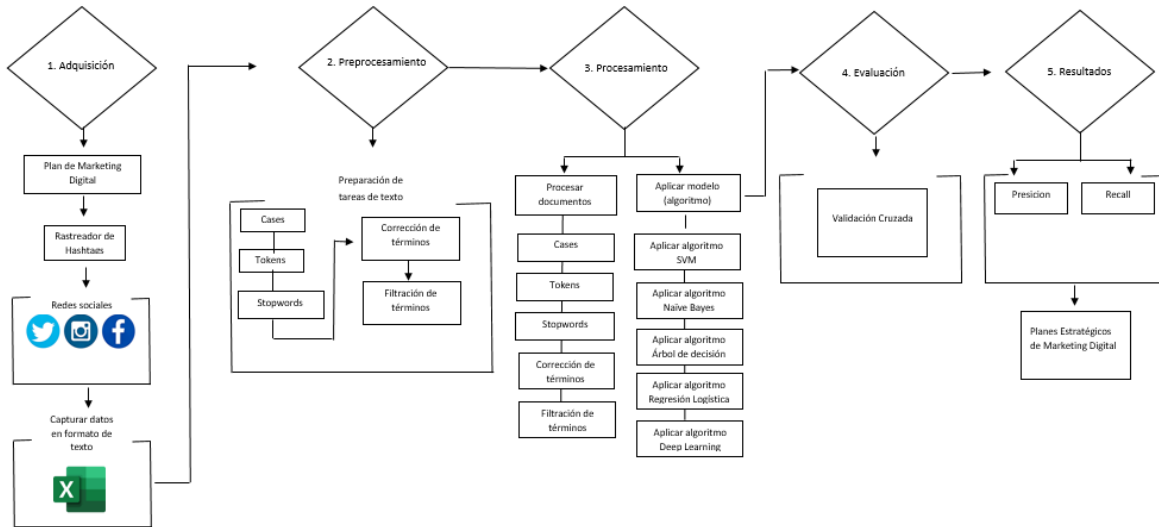


Ilustración 24.- Diagrama de Minería de Opinión para analizar los comentarios.

1.- Adquisición

El paso de adquisición de información se refiere primordialmente a la recopilación de los datos generados por los internautas en la campaña digital. Estos fueron generados a través de la participación de los mismos en redes sociales mediante la expresión de su opinión en ciertas dinámicas y actividades implementadas y fueron extraídos a través del rastreador de hashtags.

Como se ha mencionado anteriormente, este paso es fundamental porque gracias a él se pueden recolectar los datos necesarios para empezar a construir una base de datos misma que es esencial para el Preprocesamiento y Procesamiento puesto que, sin una base de datos adecuada, el análisis de Minería de Opiniones se vería afectado directamente.

Los datos se extrajeron a través de una herramienta en línea que funciona como un monitor de búsqueda de etiquetas de texto (*hashtags*) ya que estos, sirvieron para identificar los comentarios de la campaña digital en las diferentes plataformas digitales. La herramienta que se utilizó fue Social Seacher y ofrece uno de los mejores motores de búsqueda de medios sociales, es una herramienta gratuita pero con funciones Premium. Permite buscar contenido en redes sociales en tiempo real y proporciona un análisis analítico de profundidad de los datos. La herramienta puede monitorear en tiempo real diferentes redes sociales: Facebook, Twitter, Instagram, YouTube, etc. Así mismo, permite guardar las búsquedas realizadas con anterioridad y enviar de alertas al correo electrónico, también ofrece una polaridad generalizada de los tópicos buscados.

A continuación se muestra el panel de la página web de *Social Searcher*:

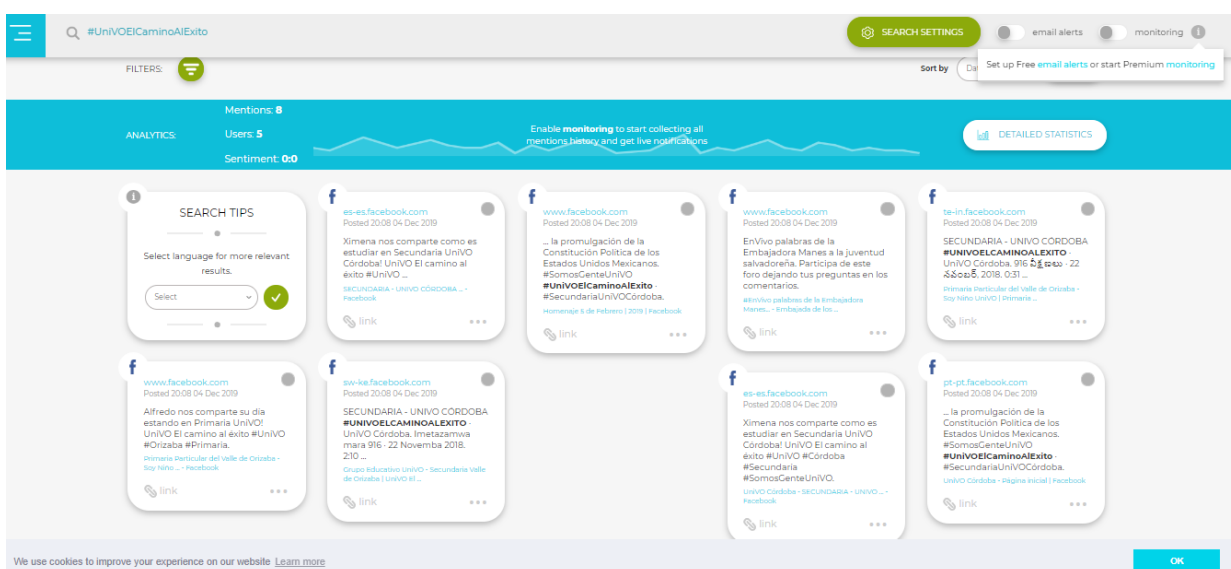


Ilustración 25.- Rastreador de Etiquetas de Texto: Social Searcher

2.- Preprocesamiento

Una vez que se ha recopilado esta información, se debe de refinar, es decir, se transforma de manera gradual a través de diferentes subprocesos. Las tareas del procesamiento son las siguientes:

- **Conversión de mayúsculas y minúsculas:** En muchos casos en la Minería de Opiniones, las palabras en mayúsculas no coincidirán con las mismas palabras en minúsculas y viceversa. El problema es mitigado a través de la conversión de las

palabras de entrada en un formato de caso único, es decir, mayúsculas o minúsculas para facilitar el procesamiento posterior.

- **Tokenización:** Este paso se refiere a remplazar datos confidenciales con símbolos de identificación únicos que permiten retener toda la información de los mismos y posteriormente, se transforman los casos y se filtran palabras vacías o remplazarlas según el caso específico de este Modelo de Minería de Opinión, cada frase es una opinión generada por el internauta en la campaña digital. Este preproceso lo dividirá en tokens.
- **Palabras de detención:** este proceso elimina las palabras más triviales, como los pronombres, las preposiciones y los artículos que se comparan con cada lista de palabras detenidas.
- **Corrección de términos:** este proceso compara cada token restante en un diccionario de idioma local formal. Si la ficha está mal escrita, se puede corregir de manera inmediata. En caso que la palabra no exista, se elimina el token.
- **Filtrado de términos:** Dado que la dimensión del espacio de texto en muchos casos puede ser alta se puede dificultar el procesamiento, es por ello que un filtrado puede reducir las dimensiones de los documentos y, por lo tanto, el costo computacional. Este proceso compara el texto con un vocabulario de dominio para filtrar y/o reducir el volumen de texto.

Después de aplicar los procesos mencionados con anterioridad, el texto disminuye de manera considerable en volumen, dejando solo el texto más relevante para el siguiente paso de procesamiento. Por lo tanto, las palabras restantes se concentran y cuentan en una matriz frecuente de conjuntos de elementos.

3.- Procesamiento

Durante esta etapa se aplican sucesivamente dos subprocesos principales: agrupamiento y categorización. Estos se explican en las siguientes secciones:

- **Agrupación:** Es una tarea de Minería de Datos para organizar datos en grupos según algunas métricas de similitud en las que estas estén involucradas. En el agrupamiento, la métrica de similitud es una función que mide la distancia entre los centros de los grupos. En el dominio de Minería de Opiniones, el enfoque de agrupación está dirigido hacia determinar polaridades, determinando si es positiva o negativa con respecto a un tema específico.

A continuación, se explicará el funcionamiento de los algoritmos utilizados para la clasificación de opiniones en el Modelo de Minería de Opiniones en RapidMiner para posteriormente, comparar su funcionamiento:

Support Vector Machine

Support vector machine (SVM) es altamente usado para la clasificación y detección de sentimientos. SVM se basa en métodos kernel, los cuales toman los datos y los ponen dentro de un espacio de características apropiado. De esta manera usan algoritmos lineales para determinar patrones no lineales. El método se basa principalmente en vectores donde, usando aprendizaje computacional, logra tomar decisiones de límite entre dos categorías separándolas lo más posible una de otra.

SVM fija el criterio de separación entre las dos clases a clasificar y éste debe estar lo más alejado posible del criterio a clasificar. La distancia, del punto de decisión, al punto más cercano es el margen del clasificador. Es así, como el método queda definido por una función de decisión que involucra un subconjunto de características o datos (support vectors) que definirán la posición del separador. Es de esta manera, la decisión del límite o margen es bastante importante ya que los datos que queden en torno a este tendrán una menor probabilidad de ser catalogados correctamente.

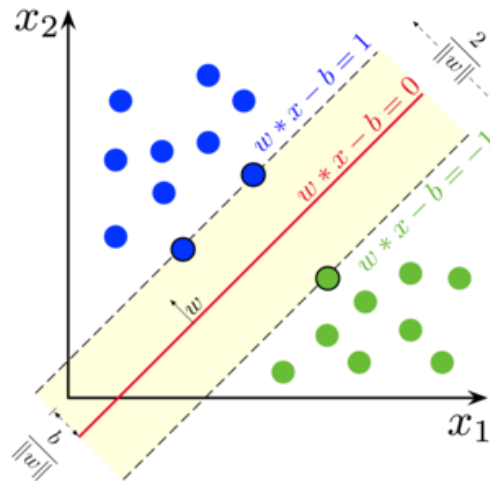


Ilustración 26.-Los vectores son separados por el hiper-plano maximizando la separación entre las clases.

Naïve Bayes

El algoritmo clasificador Naïve-Bayes (NBC), es un clasificador probabilístico simple con fuerte suposición de independencia. Aunque la suposición de la independencia de los atributos es generalmente una suposición pobre y se viola a menudo para los conjuntos de datos verdaderos. A menudo proporciona una mejor precisión de clasificación en conjuntos de datos en tiempo real que cualquier otro clasificador. También requiere una pequeña cantidad de datos de entrenamiento. El clasificador Naïve-Bayes aprende de los datos de entrenamiento y luego predice la clase de la instancia de prueba con la mayor probabilidad posterior. También es útil para datos dimensionales altos ya que la probabilidad de cada atributo se estima independientemente (Chandra et al., 2007).

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

Posterior = $\frac{\text{prior} \times \text{likelihood}}{\text{evidence}}$

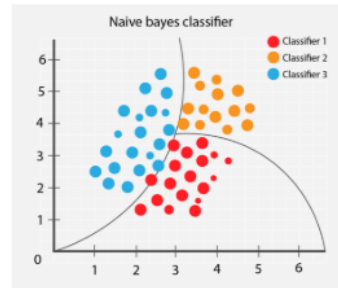


Ilustración 27.- Formula y ejemplo del comportamiento del algoritmo Naïves-Bayes.

Árbol de decisión

Es una herramienta de modelado predictivo general que tiene aplicaciones que abarcan varias áreas diferentes. En general, los árboles de decisión se construyen mediante un enfoque algorítmico que identifica formas de dividir un conjunto de datos en función de diferentes condiciones. Es uno de los métodos más utilizados y prácticos para el aprendizaje supervisado. Los árboles de decisión son un método de aprendizaje supervisado no paramétrico utilizado tanto para tareas de clasificación como de regresión. El objetivo es crear un modelo que prediga el valor de una variable objetivo mediante el aprendizaje de reglas de decisión simples inferidas de las características de los datos.

Las reglas de decisión son generalmente en forma de declaraciones y en cuanto más profundo es el árbol, más complejas son las reglas y mejor se ajusta el modelo.

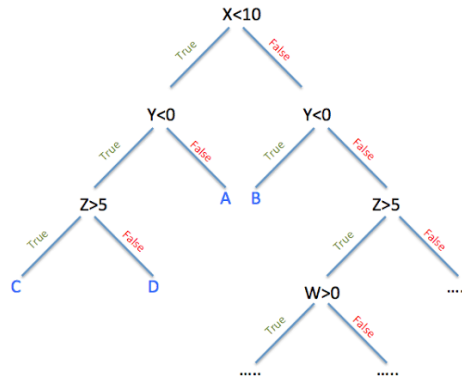


Ilustración 28.- Ejemplo del comportamiento de algoritmo Árbol de Decisión.

Regresión Logística

La Regresión Logística es uno de los algoritmos más comunes, simples y utilizados en *Machine Learning* y es útil para resolver problemas de clasificación binaria puesto que, es un método estadístico para predecir clases binarias. El resultado o variable objetivo es de naturaleza dicotómica. Dicotómica significa que solo hay dos clases posibles. En el caso de la aplicación de esta tesis, sirve para detectar la polaridad de una opinión (positiva/ negativa).

La Regresión Logística describe y estima la relación entre una variable binaria dependiente y las variables independientes y su función es utilizada en el núcleo del método, es decir, la función logística también llamada función *Sigmoide*, esta es una función curva en forma de S que puede tomar cualquier número de valor real y asignar a un valor entre 0 y 1. Si la curva va a infinito positivo la predicción se convertirá en 1, y si la curva pasa el infinito negativo, la predicción se convertirá en 0. Si la salida de la función Sigmoide es mayor que 0.5, se puede clasificar el resultado como 1 o “sí”, y si es menor que 0.5 se clasifica como 0 o “no”.

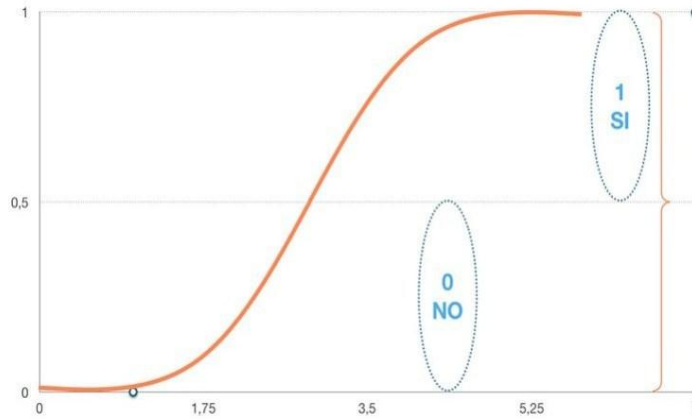


Ilustración 29.- Función Sigmoide

Deep Learning (Red Neuronal)

Una red neuronal profunda (DNN) es una red neuronal artificial (ANN) con varias capas ocultas entre las capas de entrada y salida. Al igual que en las ANN poco profundas, los DNN pueden modelar relaciones no lineales complejas. El propósito principal de una red neuronal es recibir un conjunto de entradas, realizar cálculos progresivamente complejos en ellas y dar salida para resolver problemas del mundo real como la clasificación. Esta red cuenta con una entrada, una salida y un flujo de datos secuenciales. Las redes neuronales se utilizan ampliamente en el aprendizaje supervisado y en los problemas de aprendizaje por refuerzo. Estas redes se basan en un conjunto de capas conectadas entre sí. Principalmente, se utiliza el método de descenso de gradiente para optimizar la red y minimizar la función de pérdida.

Es importante mencionar que en esta parte se definió la manera en que se determinarán las polaridades de los comentarios recabados en las redes sociales, es decir, si estos son positivos, negativos o neutrales. En otras palabras, la información cualitativa paso a ser cuantitativa.

- **Categorización:** También conocida como clasificación de texto, ayuda a identificar el tema al que pertenece un documento teniendo en cuenta algunos temas predefinidos. Actualmente, la tarea de clasificación de texto está más orientada al uso

de métodos de aprendizaje supervisados para clasificar un conjunto de documentos en una o más categorías. Por lo tanto, el objetivo es aplicar algún clasificador a documentos no etiquetados y realizar la clasificación automáticamente en documentos no etiquetados desconocidos.

Para obtener la polaridad de un comentario existen 2 métodos más usados. El primero de ellos es usar métodos de aprendizaje computacional (*machine-learning approaches*) y el segundo es utilizar diccionarios léxicos. Durante esta etapa, el aprendizaje computacional busca analizar la información de manera automática de una forma supervisada, basándose primordialmente en sets de entrenamiento, los cuales posteriormente pretenden ser utilizados para catalogar el resto de las opiniones encontradas en la plataforma digital realizando pruebas y luego validándolas. Las principales técnicas de este método son: Support Vector Machine (SVM), Naive Bayes y clasificadores de máxima entropía. De esta manera, se utiliza la categoría gramatical de las palabras, la presencia y frecuencia de algunos términos y su composición semántica (Georgios Paltoglou, 2012).

La mayoría de estos métodos, sin embargo, van acompañados de algún diccionario que entrega información a priori de los términos para obtener las polaridades respectivas. En algunos casos, estos diccionarios son realizados por personas y en otros se ocupa un sistema semiautomático.

Aplicación del modelo de Minería de Opiniones en RapidMiner

A continuación se muestran el modelo que se construyó en RapidMiner Studio para el análisis de opiniones del grupo educativo, así mismo, es fundamental para la construcción del mismo se deben de cumplir ciertas condiciones:

- Contar con una computadora con un sistema operativo actualizado, en el caso de esta investigación se utilizó una HP pavilion 15-cw00071a con Windows 10 como sistema operativo.
- Se debe tener la versión más actualizada de RapidMiner, para esta investigación se trabajó con la versión 9.1.
- RapidMiner debe de contar con las siguientes extensiones:

1.- Web Mining: La extensión web proporciona acceso a varias fuentes de Internet, como páginas web, canales RSS y servicios web. Además de los operadores para acceder a

esas fuentes de datos, la extensión también proporciona operadores específicos para manejar y transformar el contenido de las páginas web para prepararlo para su posterior procesamiento.

2.- Text Processing: Esta extensión de texto admite varios formatos de texto, incluidos el texto sin formato, HTML o PDF, así como otras fuentes de datos. Proporciona filtros estándar para tokenización, derivación, filtrado de palabras vacías o generación de n-gramas para proporcionar todo lo necesario para preparar y analizar textos.

A continuación se mencionarán las etapas y se mostrarán de manera gráfica con las cuales están conformado:

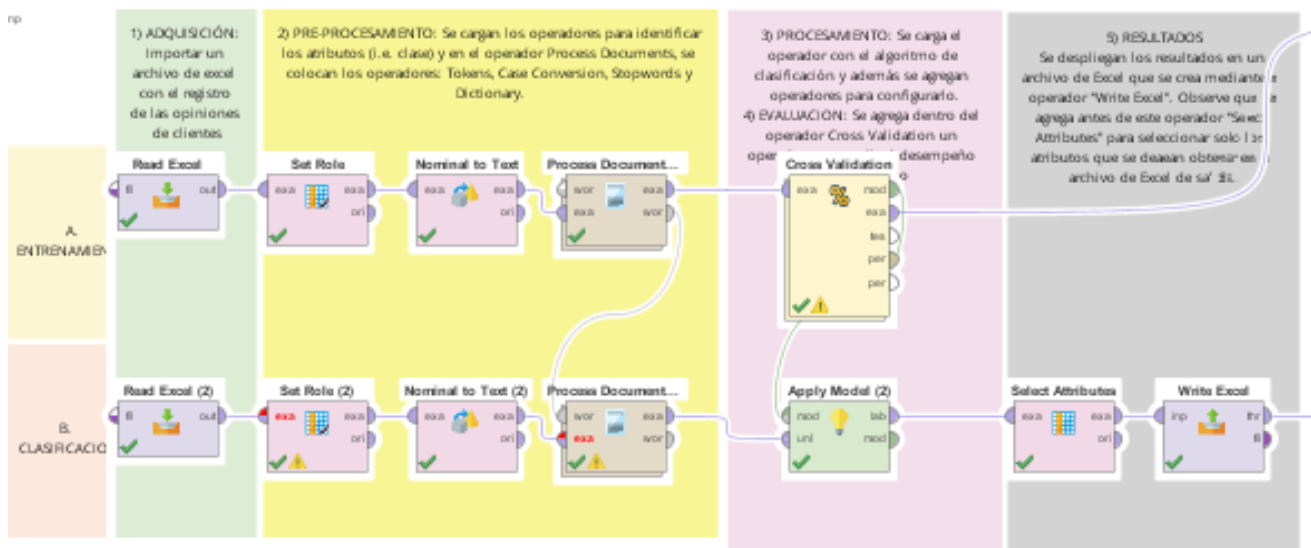


Ilustración 30.- Modelo de Minería de Opinión en RapidMiner Studio 9.1

El modelo se divide en secciones principales:

- A. entrenamiento y;
- B. clasificación

Y estas, a su vez, están conformadas por 5 etapas:

- 1) Adquisición:** En la parte de adquisición se utiliza el módulo de "Read Excel".
 - Operador "Read Excel" ("Leer Excel"): Este operador se puede usar para cargar datos de hojas de cálculo de Microsoft Excel. Este operador puede leer datos de

Excel 95, 97, 2000, XP y 2003. El usuario debe definir cuál de las hojas de cálculo en el libro de trabajo debe usarse como tabla de datos. La tabla debe tener un formato tal que cada fila sea un ejemplo y cada columna represente un atributo. Se debe tener en cuenta que la primera fila de la hoja de Excel puede usarse para nombres de atributos que pueden indicarse mediante un parámetro. La tabla de datos se puede colocar en cualquier lugar de la hoja y puede contener instrucciones de formato arbitrarias, filas vacías y columnas vacías. Los valores de datos faltantes en Excel deben indicarse con celdas vacías o con celdas que solo contengan

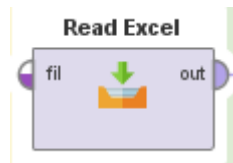


Ilustración 31.- Operador Read Excel

2) Procesamiento en el modelo de Minería de Opinión en RapidMiner: En la etapa de procesamiento se utilizan los operadores y atributos de “*Set Role*”, “*Nominal to Text*” y “*Process Documents*”. A continuación se explicará la función de cada uno de ellos:

- “*Set Role*” (“Establecer Rol”): Este operador se utiliza para cambiar la función de uno o más atributos. Su función primordialmente se describe cómo otros operadores manejan este atributo. El rol predeterminado es regular, otros roles se clasifican como especiales. Un conjunto de ejemplos puede tener muchos atributos especiales, pero cada rol especial solo puede aparecer una vez. Si se asigna un rol especial a más de un Atributo, todos los roles se cambiarán a normales excepto el último Atributo.

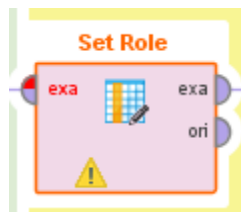


Ilustración 32.- Operador Set Role

- “*Nominal to Text*” (“Nominal a Texto”): El operador Nominal a texto convierte todos los atributos nominales en atributos de cadena. Cada valor nominal se usa

simplemente como un valor de cadena del nuevo atributo. Si falta el valor en el atributo nominal, también falta el nuevo valor.

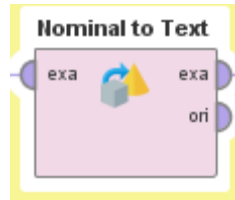


Ilustración 33- Operador Nominal to Text

- “*Process Documents*” (“Procesar Documentos”): Genera vectores de palabras a partir de atributos de cadena. Este operador puede contener otros operadores. En el caso específico de este modelo para la Minería de Opiniones cuenta dentro de él los siguientes operadores: “*Tokenize*”, “*Transform Cases*”, “*Replace Tokens*” y “*Filter Stopwords*” todos ellos se utilizan para preprocesar los datos.

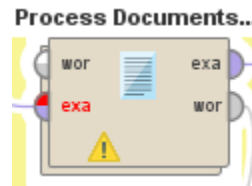


Ilustración 34.- Operador Process Documents

- “*Tokenize*” (“Tokenizar”): Este operador divide el texto de un documento en una secuencia de tokens. Hay varias opciones para especificar los puntos de división. Se pueden usar todos los caracteres que no sean letras. Esto dará como resultado tokens que consisten en una sola palabra.
- “*Transform Cases*” (“Transformación de casos”): Este operador transforma todos los caracteres en un documento en minúsculas o mayúsculas, respectivamente.
- “*Replace Tokens*” (“Reemplazar Tokens”): Este operador permite reemplazar subcadenas dentro de cada token. Por lo tanto, el usuario puede especificar pares arbitrarios qué/ para en el parámetro: *replace_dictionary*. La columna izquierda de la tabla especifica qué se debe reemplazar y el lado derecho el reemplazo. Dado que el reemplazo no se realiza sobre los límites de los tokens, este operador se coloca mejor directamente después

del operador “*TextInput*” o al menos antes del tokenizador. Para especificar qué se debe reemplazar, se pueden usar expresiones regulares.

- “*Filter Stopwords*” (“Filtrado de Stopwords”): Un operador que elimina todos los tokens iguales a una palabra de parada del archivo dado. El archivo debe contener una palabra de parada por línea.

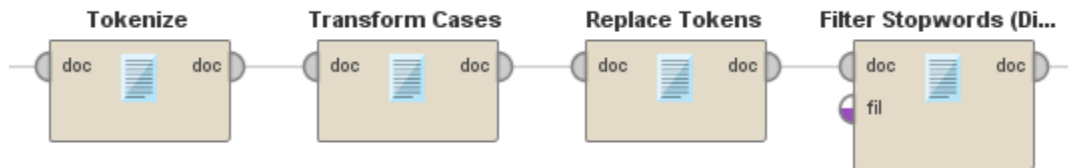


Ilustración 35.- Operadores: “*Tokenize*, *Transform Case*, *Replace Tokens*, *Filter Stopwords*” en el modelo de Minería de Opinión.

3) Procesamiento y evaluación en el modelo de MO en RapidMiner: En estas etapas del procesamiento se lleva a cabo con el operador de “*Crossvalidation*” (“Validación Cruzada”) y “*Apply model*” (“Aplicar el modelo”): Se utiliza principalmente para estimar con qué precisión un modelo (aprendido por un operador de aprendizaje en particular) se desempeñará en la práctica. El operador de validación cruzada es un operador anidado. Tiene dos subprocesos: un subproceso de capacitación y un subproceso de prueba. El subproceso capacitación se usa para entrenar un modelo. El modelo entrenado se aplica luego en el subproceso de prueba. El rendimiento del modelo se mide durante la fase de prueba.

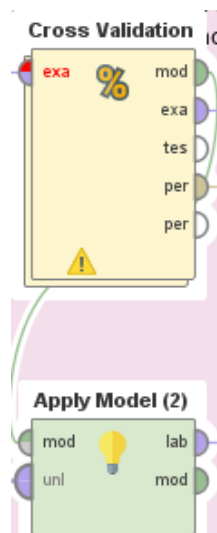


Ilustración 36.- Operador *Cross Validation* y *Apply model*.

Como se mencionó anteriormente, este operador integra al proceso de capacitación: la aplicación del algoritmo (SVM, Naïve Bayes, Regresión Logística, etc.) para posteriormente, aplicar el subproceso de prueba.

El operador de Apply Model precisamente sirve para que el software pueda realizar el entrenamiento con el operador de algoritmo elegido.

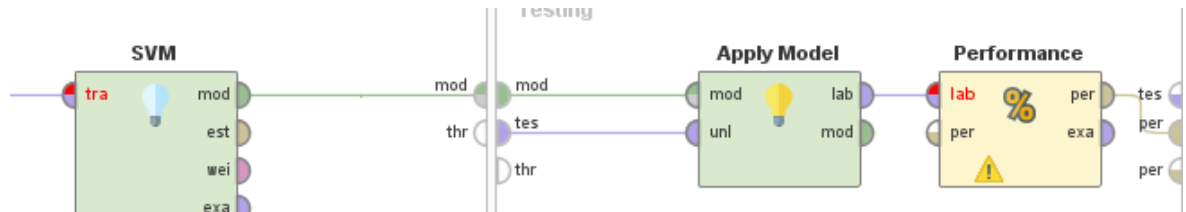


Ilustración 37.- Operadores SVM, Apply Model y Performance en la etapa de capacitación.

- 4) **Resultados:** En esta etapa se despliegan los resultados en un archivo de Excel que se crea mediante el operador "Write Excel". Observe que se agrega antes de este operador "Select Attributes" para seleccionar solo los atributos que se desean visualizar en ese archivo de salida.

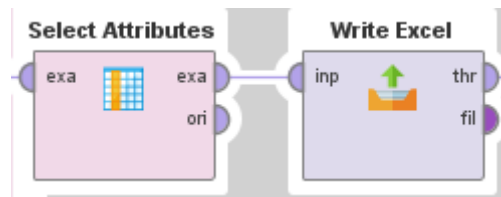


Ilustración 38.- Operadores Selec Atributes y Write Excel.

Conforme a la polaridad de las opiniones analizadas, es fundamental mencionar que se analizó por medio de los algoritmos predictivos mencionados anteriormente (SVM, Naïve Bayes, Árbol de Decisión, Regresión logística y Deep Learning). A continuación se muestran polaridades obtenidas:

Algoritmo	Polaridad Positiva	Polaridad Negativa
SVM	382	36
Naïve Bayes	349	69
Árbol de decisión	407	11
Regresión Logística	387	31
Deep Learning	406	12

Tabla 15.-Polaridades obtenidas.

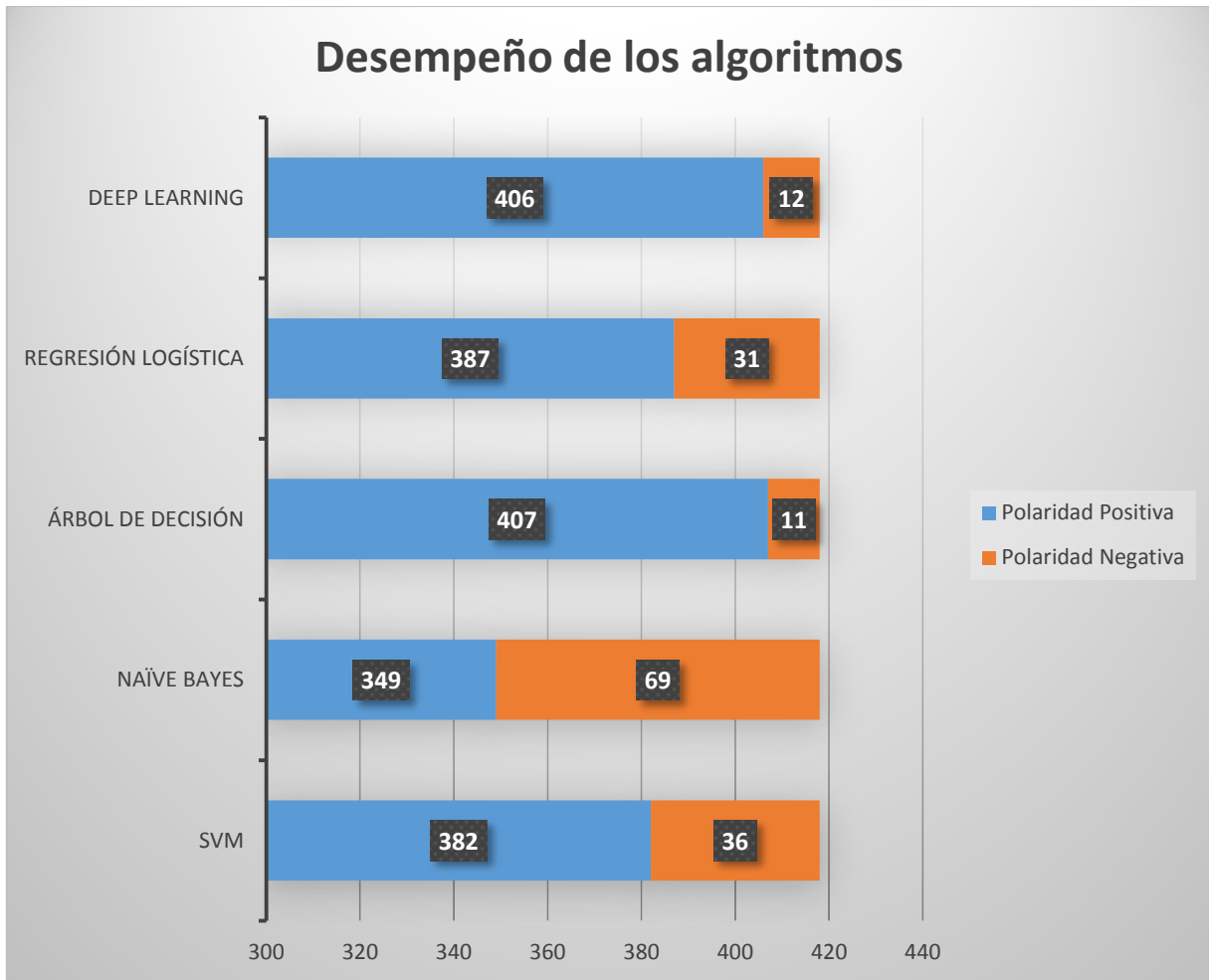


Ilustración 39.- Desempeño de los algoritmos aplicados en el modelo de Minería de Opinión.

5) Evaluación de resultados obtenidos: La evaluación del modelo conforme a su rendimiento puede ser realizada a partir de dos perspectivas:

- Objetivamente: calculando la precisión de la categoría en el modelo y;
- Subjetivamente: con una evaluación cualitativa de un dominio experto.

La primera evaluación puede estimarse calculando la precisión sobre la frecuencia con la que un documento se clasifica correctamente. Esta evaluación cuantitativa generalmente se realiza midiendo el puntaje de error y puede determinarse por las funciones de precisión y recall cuya definición se proporciona en las siguientes ecuaciones (Weiss et al., 2010):

- $\text{Precisión} = \frac{\text{clasificaciones positivas correctas}}{\text{predicciones positivas}}$

- Recall:= predicciones positivas correctas /documentos de clase positiva

Básicamente, la precisión calcula el porcentaje de documentos correctamente clasificados y por otra parte Recall mide el porcentaje de los documentos examinados por este modelo de clasificación. Sin embargo, estas dos medidas pueden ser sensibles a variaciones entre sí. Por esta razón y en el caso específico de la aplicación de este modelo se evaluó el desempeño que tuvieron los comentarios bajo un marco de cinco meses de las campañas digital para tener un mayor control de la información.

Se puede observar en la ilustración número 39 el desempeño de los algoritmos aplicados en el modelo de Minería de Opinión en RapidMiner y se puede observar que la polaridad se inclina hacia tendencia positiva pero, para comprobar cuál algoritmo tiene un mejor funcionamiento conforme al tipo de datos que genera el grupo educativo es necesario realizar una validación por medio de la aplicación de una matriz de confusión para cada algoritmo. Es importante mencionar que para comprobar lo anterior se realizó el método de “*Hold Out*” mismo que consiste en dividir un conjunto de dos partes: “datos de entrenamiento” y “datos prueba”. El conjunto de entrenamiento es en lo que se entrena el modelo, y el conjunto de prueba se usa para ver qué tan bien ese modelo funciona en datos no vistos. Una división común cuando se usa el método de retención es usar el 80% de los datos para el entrenamiento y el 20% restante de los datos para las pruebas. En el caso de la aplicación de esta tesis se aplicará el 90% de los datos para el entrenamiento y el 10% restante de los datos para prueba y realizar a partir de esto una matriz de confusión para cada algoritmo.

A continuación los resultados de la misma para cada algoritmo:

Matriz de confusión	
Algoritmo SVM	
VP	FP
28	10
VN	FN
2	0

Tabla 16.- Matriz de confusión SVM

Precisión: 70%

Error: 30%

Matriz de confusión Algoritmo Naïve Bayes	
VP	FP
13	1
VN	FN
12	14

Tabla 17- Matriz de confusión Naïve Bayes.

Precisión: 67.5%

Error: 32.5%

Matriz de confusión Algoritmo Árbol de Decisión	
VP	FP
27	13
VN	FN
0	0

Tabla 18.- Matriz de confusión Árbol de Decisión.

Precisión: 68%

Error: 33%

Matriz de confusión Algoritmo Regresión Logística	
VP	FP
28	10
VN	FN
2	0

Tabla 19.- Matriz de confusión Regresión Logística

Precisión: 70%

Error: 30%

Matriz de confusión Algoritmo Deep Learning	
VP	FP
27	11
VN	FN
2	0

Tabla 20.- Matriz de confusión Deep Learning

Precisión: 68%

Error: 33%

A continuación se muestra una tabla comparativa en donde se puede comparar la polaridad de todos los algoritmos utilizados, así como la precisión y error de cada uno.

Algoritmo	Polaridad Positiva	Polaridad Negativa	Precisión	Error
SVM	91%	9%	70%	30%
Naïve Bayes	83%	17%	67.5%	32.5%
Árbol de decisión	97%	3%	68%	33%
Regresión Logística	93%	7%	70%	30%
Deep Learning	97%	3%	68%	33%

Tabla 21.- Tabla comparativa de rendimiento de algoritmos.

Como se puede observar en la tabla anterior, los algoritmos que mejor desempeño tienen para el tipo de base de datos conforme a comentarios del grupo educativo son SVM y Regresión Logística con 70% de precisión, posteriormente sigue Árbol de Decisión y *Deep Learning* con un 68% y por último el algoritmo de Naïve Bayes con un 67.5% de precisión. Es necesario mencionar que la diferencia entre los mismos no es grande pero para tener

una clasificación efectiva de debe de utilizar el algoritmo que mejor precisión y menos error tenga conforme al análisis.

6.- Redireccionamiento de estrategias de Marketing Digital conforme a los resultados.

Por último, una vez obtenida la información sobre el nivel de impacto de la campaña *online*, así como la polaridad de la misma, se desarrollarán estrategias enfocadas especialmente las áreas de oportunidad del Marketing Digital del centro educativo.

Es importante mencionar que el análisis anterior arrojó una serie de palabras que más se repetían en las opiniones recolectadas, mismas que proporcionan información valiosa para la planificación de estrategias de Marketing Digital.

La lista de palabras es la siguiente:

Palabra	Sentimiento
Redes sociales	Positivo
Amables	Positivo
Atención	Positivo
Actualizan	Positivo
Confianza	Positivo
Educación	Positivo
Encantar	Positivo
Excelente	Positivo
Estudios	Positivo
Perfecto	Positivo
Nivel	Positivo
Servicio	Positivo
Profesional	Positivo

Tabla 22.- Tabla de aspectos positivos.

Así mismo y bajo el contexto de los aspectos positivos destacados que se obtuvieron en la investigación se realizó también en RapidMiner un procedimiento para formar una nube de palabras la cual, permita ilustrar de manera visual dichos aspectos. A continuación se muestran de manera general los operadores y la relación que tiene cada uno para realizar dicho método:

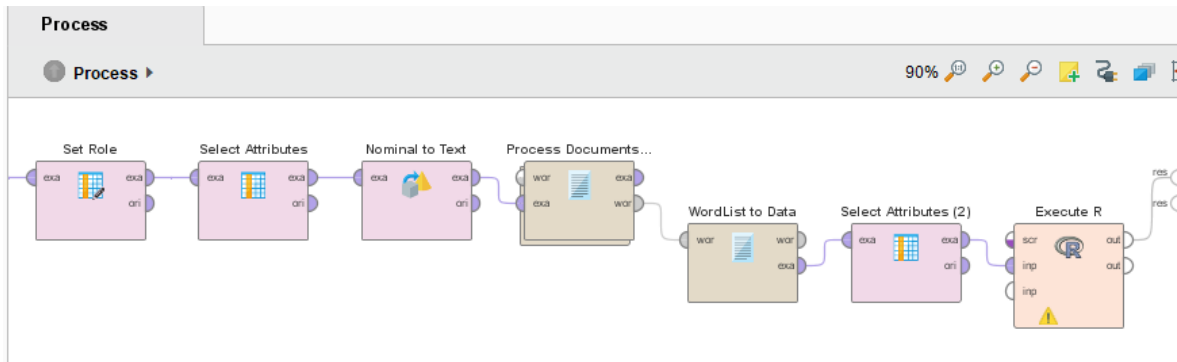


Ilustración 40.- Procedimiento en RapidMiner para crear una Nube de Palabras.

Los operadores que se utilizan para el procedimiento para crear una Nube de Palabras son los siguientes:

1. Set Role
2. Select Attributes
3. Nominal to Text
4. Process Documents
5. WordList to Data
6. Execute R

A continuación se explicará la función de cada uno de ellos:

- “Set Role” (“Establecer Rol”): Este operador se utiliza para cambiar la función de uno o más atributos. Su función primordialmente se describe cómo otros operadores manejan este atributo. El rol predeterminado es regular, otros roles se clasifican como especiales. Un conjunto de ejemplos puede tener muchos atributos especiales, pero cada rol especial solo puede aparecer una vez. Si se asigna un rol especial a más de un Atributo, todos los roles se cambiarán a normales excepto el último Atributo.



Ilustración 41.- Operador Set Role

- “Select Atributes” (Seleccionar Atributos): El operador proporciona diferentes tipos de filtros para facilitar la selección de atributos. Las posibilidades son, por ejemplo: selección directa de atributos. selección mediante una expresión regular o seleccionando solo atributos sin valores faltantes. consulte el tipo de filtro de atributo de parámetro para obtener una descripción detallada de los diferentes tipos de filtro.



Ilustración 42- Operador Select Attributes

- “Nominal to Text” (“Nominal a Texto”): El operador Nominal a texto convierte todos los atributos nominales en atributos de cadena. Cada valor nominal se usa simplemente como un valor de cadena del nuevo atributo. Si falta el valor en el atributo nominal, también falta el nuevo valor.

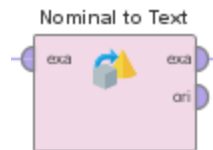


Ilustración 43- Operador Nominal to Text

- “Process Documents” (“Procesar Documentos”): Genera vectores de palabras a partir de atributos de cadena. Este operador puede contener otros operadores. En el caso específico de este modelo para la Minería de Opiniones cuenta dentro de él los siguientes operadores: “Tokenize”, “Transform Cases”, “Replace Tokens” y “Filter Stopwords” todos ellos se utilizan para preprocesar los datos.



Ilustración 44.- Operador Process Documents

- “*Tokenize*” (“Tokenizar”): Este operador divide el texto de un documento en una secuencia de tokens. Hay varias opciones para especificar los puntos de división. Se pueden usar todos los caracteres que no sean letras. Esto dará como resultado tokens que consisten en una sola palabra.
- “*Transform Cases*” (“Transformación de casos”): Este operador transforma todos los caracteres en un documento en minúsculas o mayúsculas, respectivamente.
- “*Replace Tokens*” (“Reemplazar Tokens”): Este operador permite reemplazar subcadenas dentro de cada token. Por lo tanto, el usuario puede especificar pares arbitrarios qué/ para en el parámetro: *replace_dictionary*. La columna izquierda de la tabla especifica qué se debe reemplazar y el lado derecho el reemplazo. Dado que el reemplazo no se realiza sobre los límites de los tokens, este operador se coloca mejor directamente después del operador “*TextInput*” o al menos antes del tokenizador. Para especificar qué se debe reemplazar, se pueden usar expresiones regulares.
- “*Filter Stopwords*” (“Filtrado de Stopwords”): Un operador que elimina todos los tokens iguales a una palabra de parada del archivo dado. El archivo debe contener una palabra de parada por línea.

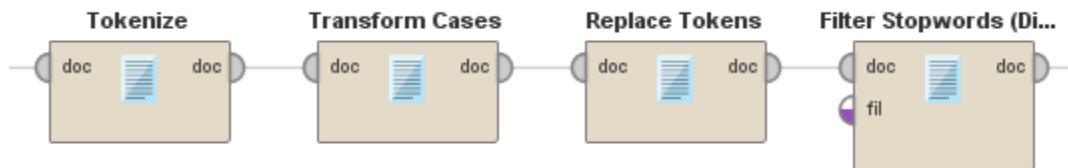


Ilustración 45.- Operadores: “*Tokenize, Transform Case, Replace Tokens, Filter Stopwords*” en el modelo de Nube de Palabras.

- “*WordList to Data*” (Lista de Palabras a Datos): Este operador crea un conjunto de datos a partir de una lista de palabras. El conjunto de datos contiene una fila para cada palabra y atributos para la palabra en sí, el número de documentos en los que ocurrió, el número de documentos etiquetados en los que ocurrió y para cada número de clase que ocurrió en un documento de esta clase. El operador puede ser útil para filtrar una lista de palabras antes de informarla.



Ilustración 46. - Operador WordList to Data

- “Execute R” (Ejecutar R): antes de utilizar este operador, debe especificar la ruta de su instalación de R en Herramientas -> Preferencias -> Scripting de R. Su instalación de R debe incluir el paquete 'data.table'. Este operador ejecuta el script proporcionado a través del puerto o parámetro del archivo de script o el script especificado en el parámetro de script. Los argumentos del script corresponden a los puertos de entrada, donde los conjuntos de ejemplos se convierten en marcos de datos. Análogamente, los valores devueltos por el script se entregan en los puertos de salida del operador, donde los marcos de datos se convierten en conjuntos de ejemplos. La salida de la consola de R se muestra en la Vista de registro (Ver -> Mostrar vista -> Registro).

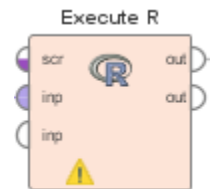


Ilustración 47.- Operador Execute R

A continuación se muestra la nube de palabras obtenida con los aspectos positivos que más se dieron a notar en las opiniones de los internautas, las opiniones aparte de resaltar los aspectos notablemente afectivos se pueden relacionar directamente los servicios que ofrece el grupo educativo como lo es la oferta educativa y también es reconocido el servicio al cliente.

el servicio al cliente mismos aspectos que deben de ser considerados en la planificación de las estrategias de Marketing Digital 2020 de la organización.

Capítulo 5: Conclusiones Generales

5.1 Introducción

En este último capítulo se presentan las conclusiones generales mismas que lograron consolidarse como consecuencia de la recolección de datos de las plataformas digitales del grupo educativo, elaboración del modelo de Minería de Opiniones con el software RapidMiner y la validación de cada uno de los algoritmos. Se detallan además los aportes de la investigación y el planteamiento de trabajos futuros. El objetivo fundamental de esta tesis fue el de abordar la problemática que tenía la organización en sus redes sociales a partir de la implementación de un modelo de Minería de Opiniones para evaluar si la campaña digital era efectiva o no puesto que, no se llevaba un plan de Marketing Digital estructurado y no se habían definido los indicadores a evaluar.

La contribución principal de esta investigación fue realizar esta evaluación en tiempo y forma y con un modelo ajustado a las necesidades de la organización mismo que, podrá utilizar en sus próximas campañas digitales, así como dar una dirección a sus estrategias digitales. La investigación profundizó en los temas de Business Performance Management, Minería de Datos, Minería de Textos, Minería de Opiniones y Marketing Digital.

5.2 Conclusiones generales

- La información extraída mediante el análisis puede ser útil para redireccionar las campañas de Marketing Digital de la organización, sin embargo, es necesario identificar los indicadores y el propósito de la campaña digital, así como las acciones a realizar en la planificación estratégica de Marketing Digital para tener mejores resultados y no tener complicaciones en el análisis de resultados y retroalimentación de la misma.
- El análisis de opiniones que se realizó conforme a polaridad tuvo como resultado una tendencia positiva ya que, cuenta con más del 91% de comentarios positivos y un 9%. La base de datos con la que se realizó este análisis cuenta con un total de 411 datos mismos que fueron obtenidos de las redes sociales del centro

educativo. Por otra parte, es importante mencionar en este punto, que el sesgo en un principio era basto ya que, las dinámicas que se llevó a cabo para obtener las opiniones no permita a los usuarios pudieran expresar libremente su opinión por el miedo tener represalias a partir de esto puesto que, estos pertenecen a este centro educativo.

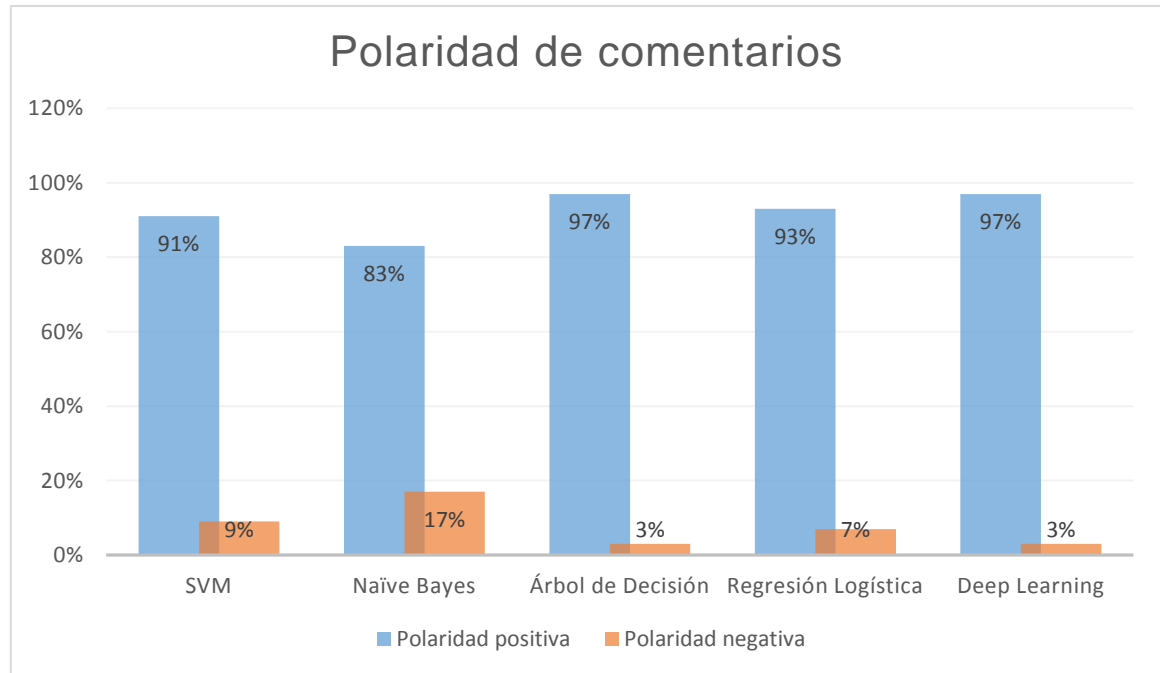


Ilustración 49.- Polaridad de comentarios

Como se puede observar en la ilustración número 49 en donde se muestra la polaridad obtenida en las opiniones conforme a la implementación de los diferentes algoritmos, los que mejor rendimiento tuvieron fueron SVM y Regresión Logística con una precisión arriba del 66% y con un margen de error del 30%. Posteriormente se encuentran los algoritmos de Deep Learning y Árbol de decisión con una precisión del 68% y un margen de error de 33%. Los algoritmos que es menos adecuado para el tipo de información procesada en esta Minería de Opinión es el Naïve Bayes el cual clasifico el 83% como comentarios positivos y el 17% como comentarios negativos y tuvo una precisión del 67.5% y un margen de error de 32.5%.

Respecto al análisis de la Minería de Opinión se muestra una tendencia hacia una tendencia positiva y se obtuvo al analizar los comentarios obtenidos en las plataformas digitales con cinco diferentes algoritmos de los cuales, los más destacados por su rendimiento fueron SVM y Regresión Logística con un 70% de

precisión y un 30% de error. El algoritmo SVM se destaca puesto que es un algoritmo altamente usado en la clasificación y detección de sentimientos.

- La participación de los usuarios en el periodo de evaluación de resultados Febrero- Noviembre fue dinámica puesto que los usuarios han aceptado las actividades *online* y permiten una rápida adaptación a dichas actividades.
- La atención de los usuarios hacia las dinámicas *online* fue variable puesto que, una vez que se acostumbran a ellas pierden el interés fácilmente es por ello, que se han renovado la manera en que se muestran y se realizan las actividades para mantener su atención por mucho más tiempo como por ejemplo, implementar un programa de fidelización *online* el cual, tenga como objetivo la inmersión de los usuarios a las dinámicas digitales y la activación de los mismos conforme a promociones con convenios establecidos de interés al consumidor.
- El equipo de Marketing Digital ha notado que los comentarios; tanto positivos como negativos ayudan de manera exponencial a re-direccionar la campaña *online* de la Universidad puesto que, al saber la polaridad de las opiniones conforme a las trivias implementadas se puede saber si el tiempo de implementación es correcto o en su defecto, si las instrucciones a seguir son claras.

A continuación se muestra el crecimiento en redes sociales en comparación 2018 y 2019 del grupo educativo:

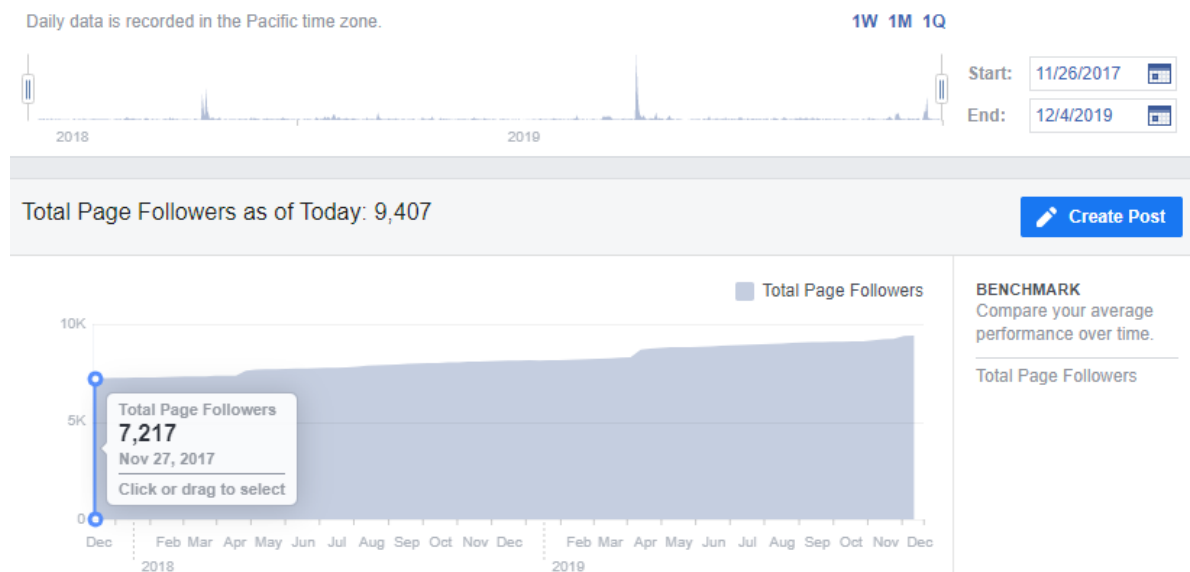


Ilustración 50.- Número de seguidores enero 2018

En enero del 2018 la universidad tenía alrededor de 7,000 *likes* posteriormente a finales del 2019 tiene un total de 9,407 *likes* en Facebook. Este indicador muestra que hubo un aumento de 2000 *likes*. Y se puede observar que donde hubo más aumento de los mismos fue en el 2019.

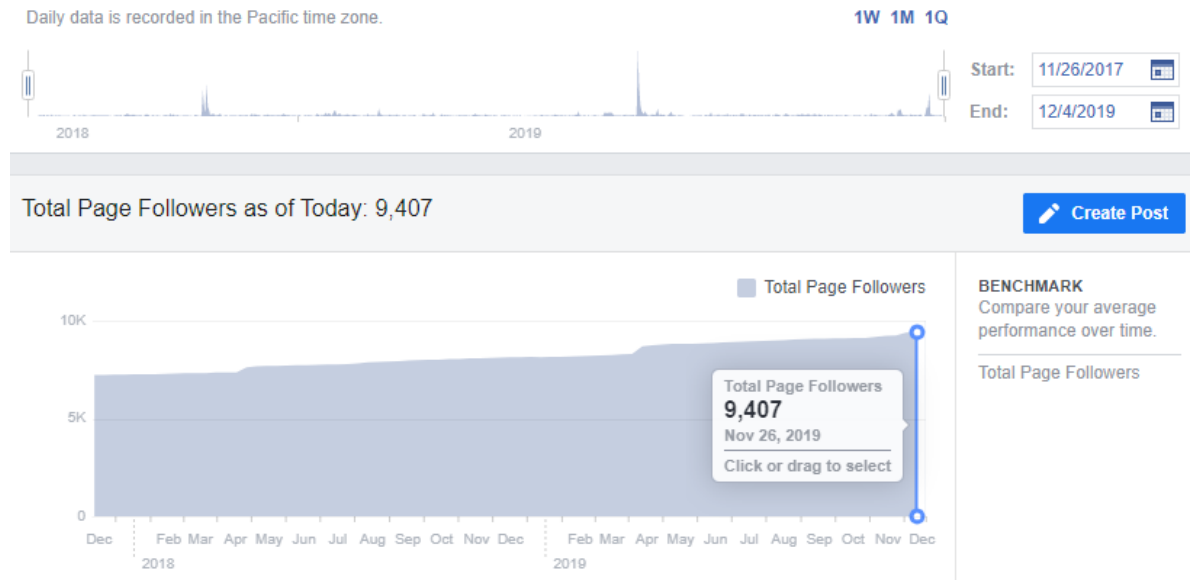


Ilustración 51.- Número de likes en 2019

Posteriormente, se muestra la gráfica de seguidores (*followers*) y se puede observar que tuvo puntos altos en los meses de abril, mayo y noviembre del 2019. Meses que coinciden con la implementación de tácticas y actividades en línea para generar *engagement*.

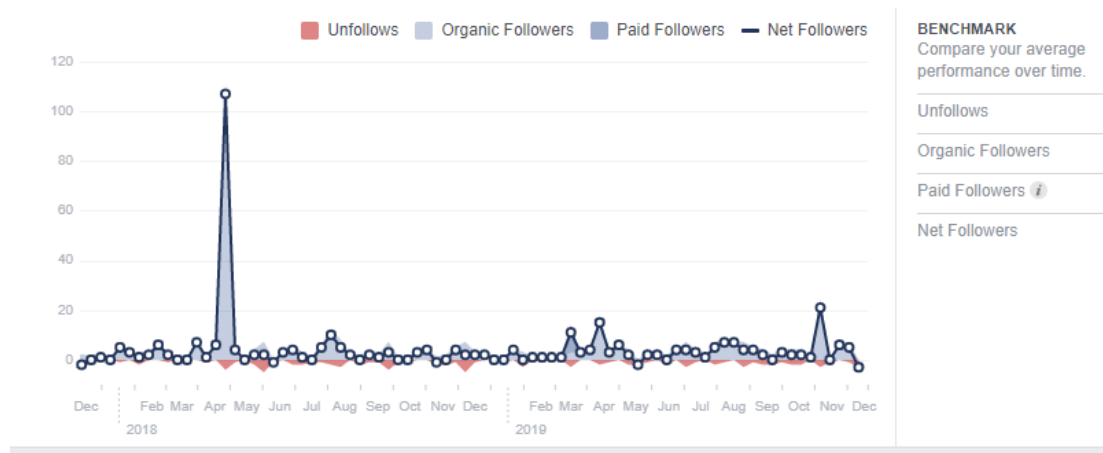


Tabla 23.- Gráfica de followers

Lo anterior, muestra de manera gráfica el crecimiento en la red social Facebook del grupo educativo en el último año, se debe de reconocer que planificación de las publicaciones, actividades y tácticas a implementar resultaron ser favorables para el crecimiento de la plataforma y el *engagement*. Por su parte, la Minería de Opiniones ayudo a redireccionar las estrategias cada dos meses e innovar conforme a las actividades a realizar para seguir manteniendo estas métricas.

- El análisis de opiniones conforme a identificar polaridades es un campo en el que solo habían estado inmersas organizaciones con giros a fines a la clasificación de datos o en su defecto, que su comercio se realice directamente en plataformas digitales; dejando a un lado la oportunidad para las empresas de servicios el tomar estas técnicas para su conveniencia, en el caso específico de los centros educativos se debe aprovechar el uso de su información de una manera que pueda satisfacer su planificación estratégica de mercadotecnia tradicional y digital a corto, mediano y largo plazo y con una retroalimentación directa de sus consumidores y clientes.

5.3 Aportaciones

- En este trabajo de investigación se presentaron los fundamentos enfocados para enfrentar la problemática planteada, la clasificación automática de textos de opinión a través de su polaridad.
- El avance tecnológico en el que se ven conjuntas las empresas exhorta al consumidor a expresarse a través de los medios sociales digitales, permitiendo a través de las redes sociales como se sienten respecto a cierto producto o servicio y si este es agradable, bueno, malo o si se sienten disconformes puesto que, para una organización es de suma importancia el rescatar esta información ya que, esta permite tomar decisiones conforme a campañas tradicionales y sobre todo digitales en las que está invirtiendo sus recursos la organización.
- La clasificación automática es una herramienta potente, que ha ido facilitando el proceso de la extracción y clasificación manual, disminuyendo costos y tiempos para las organizaciones en la que es utilizada.

- El proceso para la clasificación es fundamental que se encuentre bien estructurado, para que así los algoritmos puedan funcionar de manera correcta, por ello es que se trabajó con una mayor cantidad de datos como sea posible para lograr mejores resultados y también se probaron diferentes tipos de algoritmos.
- Realizar un modelo de Minería de Opiniones permite que tanto en la etapa de pre-procesamiento, al igual que en el proceso de Minería de Datos formen un papel fundamental para la obtención de buenos resultados. Transformar las opiniones, realizar limpieza a los datos, convertirlos para que el clasificador entienda el texto es un punto fundamental para evitar errores. Sin embargo, se puede concluir en este punto que es necesario tener un conocimiento significativo sobre disciplinas como las ciencias computacionales, lingüística y estadística para obtener mejores resultados.

5.4 Trabajos futuros

Tanto la Minería de Web Social como la Minería de Opinión son áreas que tienen muchas problemáticas por resolver puesto que, hay muchas aplicaciones donde se puede integrar. Es necesario estudiar los diferentes fenómenos lingüísticos que interfieren o se encuentran presentes en la Minería de Opinión de manera específica, los que se extraen de redes sociales.

Como trabajo futuro bajo el contexto de esta investigación se propone lo siguiente: Profundizar en temáticas relacionadas con fenómenos lingüísticos que tengan relación directamente con la polaridad como por ejemplo, la neutralidad y el sarcasmo ya que son aspectos que el lenguaje presente y que en cada idioma se ven integradas y puede cambiar de manera drástica el análisis y los resultados de una investigación es por ello, que se recomienda colaborar de manera sinérgica con diferentes lingüistas para hacer un estudios de los distintos idiomas, el comportamiento que presenten y los niveles de afectación que pueden tener en documentos de opinión.

Los trabajos futuros que se pueden dar a partir de la comprensión y aplicación de un modelo basado en la Minería de Opiniones tiene un horizonte muy amplio, puesto que, no sólo se puede utilizar para investigaciones con fines de Marketing Digital y comercio electrónico sino también, se puede realizar sinergia y aplicación académica,

administración estratégica, recurso humano, logística, seguridad, biomedicina, entre otras áreas

5.5 Vinculación

Es necesario mencionar que como parte de la elaboración de esta tesis y con referencia a las temáticas implementadas: Minería de la Web Social y Análisis de Sentimientos se realizó una estancia de investigación en Lakehead University, Thunder Bay, Canada en el laboratorio DataLab del centro de Estudios Superiores de Ingeniería Y Ciencia (Centre for Advanced Studies in Engineering and Scienc) bajo la supervisión del Dr. Vijay Mago y del Dr. Eduardo Roldán Reyes durante el otoño del 2019 (Septiembre 15, 2019 a Noviembre 15, 2019).

En esta estancia académica se desarrolló un método que sea capaz de pre-procesar el texto y también reconocer que tipo de algoritmo se adapta mejor al texto de las opiniones proporcionadas por los usuarios de una campaña digital del Grupo Educativo Univo. Es decir, a través del Análisis de Sentimientos catalogar las opiniones como positivas y negativas.

La metodología con la que se trabajó conto con los siguientes pasos:

- Revisión de literatura: Se revisaron diferentes investigaciones afines para tener un mejor entendimiento de las temáticas mencionadas anteriormente.
- Construcción del modelo: Se construyó en lenguaje Python un modelo de Minería de Opiniones.
- Implementación del modelo: Se implementó el modelo y a continuación se muestran las secciones del código:

A continuación se muestra parte del código que se desarrolló:

1.- Importación de módulos y librerías.

```

import pandas as pd
import numpy as np
import csv
import nltk
import string
import matplotlib.pyplot as plt
%matplotlib inline
import seaborn as sns
import random

from sklearn.pipeline import Pipeline
from sklearn.model_selection import train_test_split
from sklearn.metrics import classification_report, confusion_matrix, accuracy_score
from sklearn.naive_bayes import MultinomialNB
from sklearn.model_selection import KFold, cross_val_score
from sklearn.ensemble import RandomForestClassifier
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.feature_extraction.text import TfidfTransformer
from sklearn.model_selection import GridSearchCV
from sklearn.externals import joblib
from nltk.corpus import stopwords
from nltk.tokenize import TweetTokenizer
from nltk.stem.wordnet import WordNetLemmatizer

```

2.- Cargar archivo:

```

# Load train data
data = pd.read_csv('trivia_comentarios_base_de_datos.csv',
                  error_bad_lines=False)
data.columns = ['id', 'label', 'source', 'text']
data

```

3.- Mostrar texto y etiquetas:

```

In [3]: # get text and matching label columns
data = data.drop(['id', 'source'], axis=1)
data.head(10)

```

```

Out[3]:

```

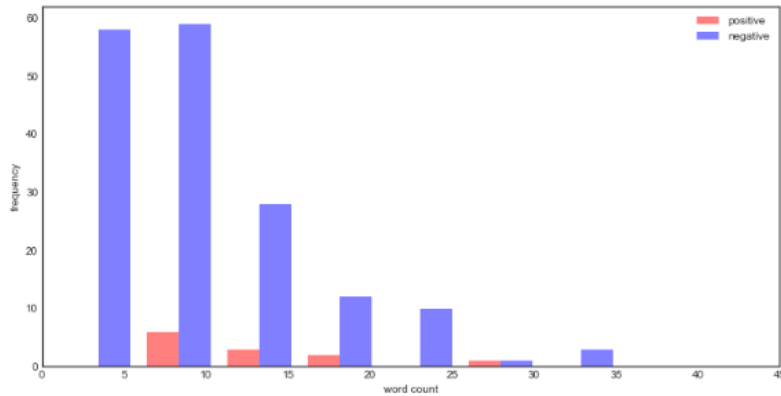
	label	text
0	0	Excellent
1	0	Very good trivia
2	0	Great trivia good tribute to a classic game
3	1	The video was going too fast you could barely see the options but the content is fine
4	0	Very funny
5	0	Very good
6	0	Hahaha I thought they would be more complicated questions
7	0	UniVo always surprising me
8	0	Interesting dynamics
9	0	Good dynamics

- Resultados: Los resultados que se obtuvieron fueron del 20% (275 comentarios) de la base de datos original y a continuación se muestra como se comportaban los mismo. Es fundamental mencionar que este comportamiento no es definitivo y que los resultados cambian dependiendo de la implementación del algoritmo en función. Como conclusión general de los mismos resultados es que los datos tienen una tendencia primordialmente positiva. Se muestran algunas gráficas de los mismos enseguida:

4.-Muestra de positivos y negativos de manera gráfica:

```
In [7]: # plot word count distribution for both positive and negative sentiments
x = data['word count'][data.label == 1]
y = data['word count'][data.label == 0]
plt.figure(figsize=(12,6))
plt.xlim(0,45)
plt.xlabel('word count')
plt.ylabel('frequency')
g = plt.hist([x, y], color=['r','b'], alpha=0.5, label=['positive','negative'])
plt.legend(loc='upper right')
```

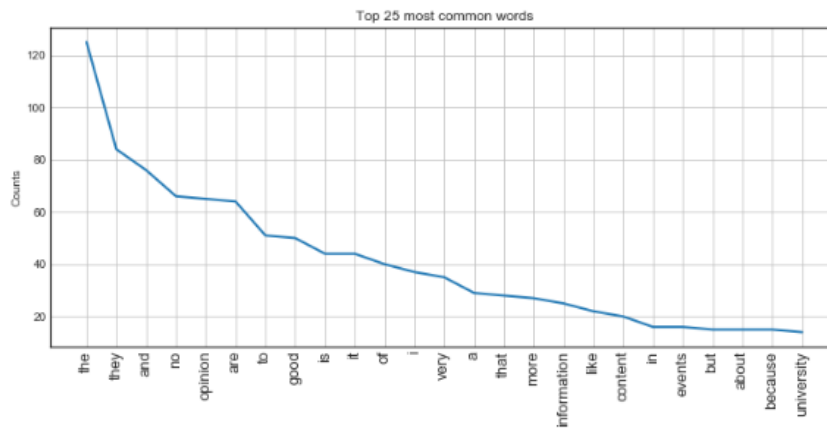
Out[7]: <matplotlib.legend.Legend at 0x2310550df98>



5.- Gráfica de distribución de frecuencias de las palabras más utilizadas

```
# plot word frequency distribution of first few words
plt.figure(figsize=(12,5))
plt.title('Top 25 most common words')
plt.xticks(fontsize=13, rotation=90)
fd = nltk.FreqDist(all_words)
fd.plot(25,cumulative=False)

# Log-Log plot
word_counts = sorted(Counter(all_words).values(), reverse=True)
plt.figure(figsize=(12,5))
plt.loglog(word_counts, linestyle='-', linewidth=1.5)
plt.ylabel("Freq")
plt.xlabel("Word Rank")
plt.title('log-log plot of words frequency')
```



De la estancia académica se obtuvo el siguiente conocimiento:

A. Convertir datos de texto en características.

B. Limpiar, etiquetar y entrenar datos.

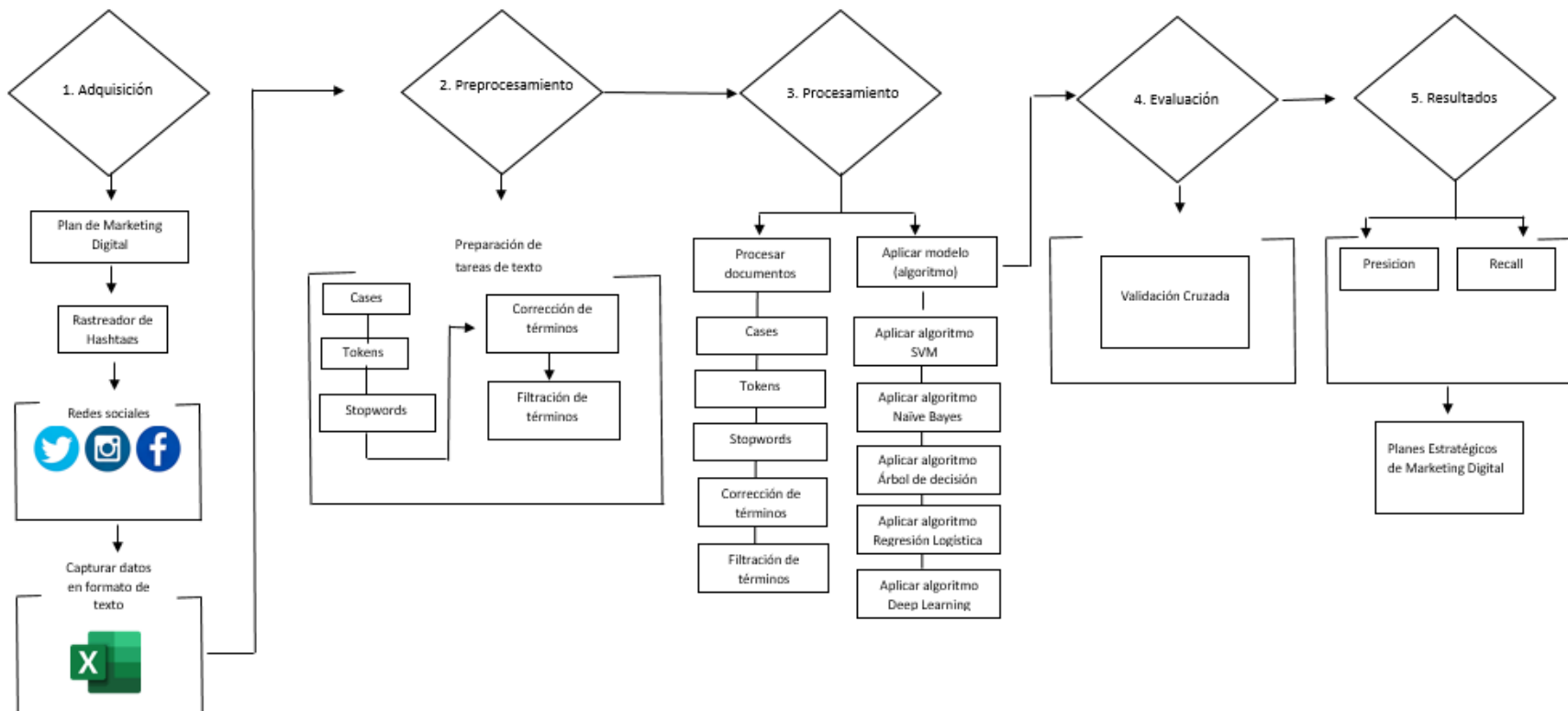
C. Compara el análisis de sentimientos con diferentes algoritmos.

Sin duda alguna, se obtuvo conocimientos enfocados en programación (especialmente en Python). Esto me permite analizar los datos de la investigación con mayor flexibilidad que usando un software, mismo que puede algunas veces puede limitar la funcionalidad.

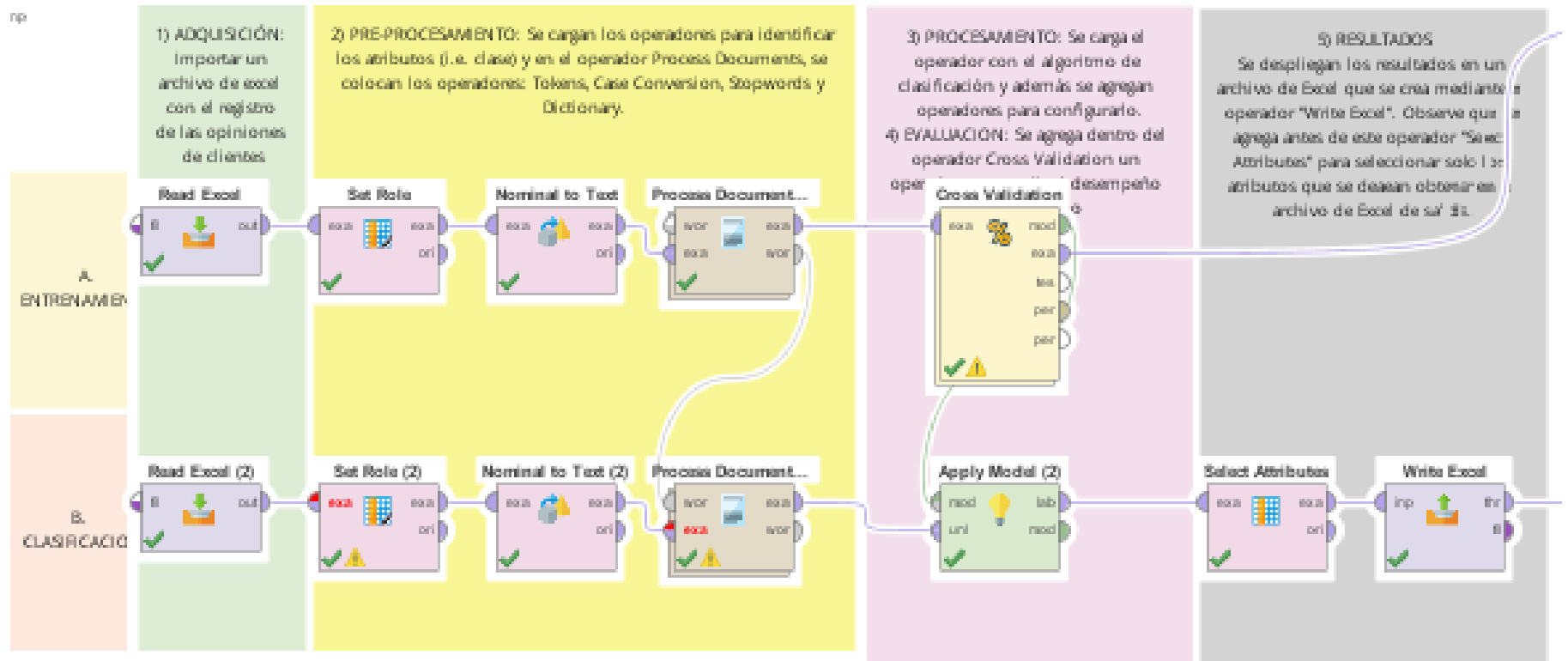
Así mismo, se pudo analizar la diferencia entre los resultados obtenidos mediante el uso de código de programación y el uso del software RapidMiner

Anexos

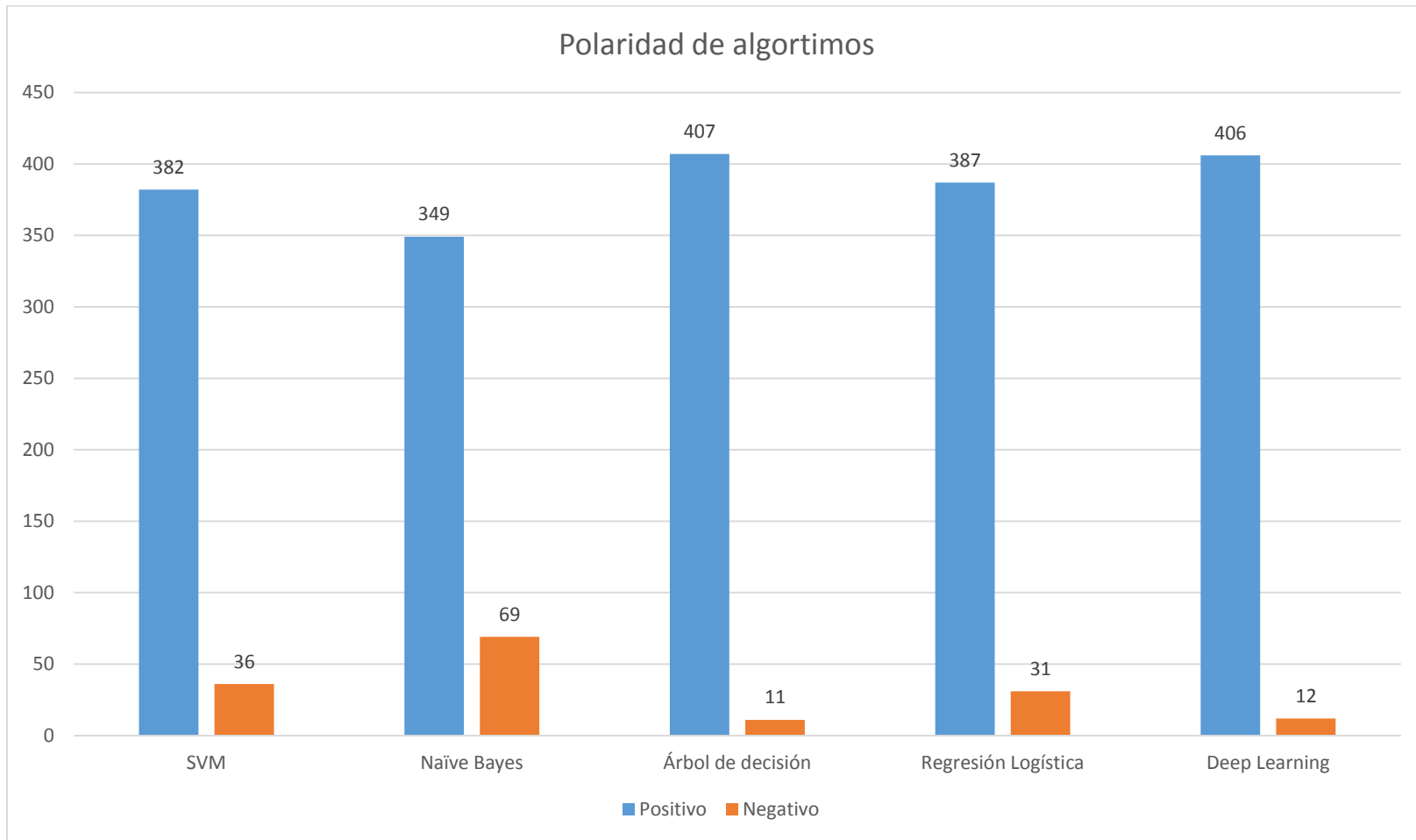
Anexo 1.- Diagrama de Minería de Opinión para analizar los comentarios.



Anexo 2.- Modelo de Minería de Opinión en RapidMiner Studio 9.1



Anexo 3.- Modelo de Minería de Opinión en RapidMiner Studio 9.1



Anexo 4.- Extracto de la base de datos de opiniones y su polaridad

Opinión	Confianza (positivo)	Confianza (negativo)	Predicción (Sentimiento)
excelente	.8	.2	positivo
buena trivía	.7	.3	positivo
gran trivía buen homenaje un juego clásico	.7	.3	positivo
iba demasiado rápido vídeo apenas podían ver opciones contenido está	.7	.3	positivo
divertida	.7	.3	positivo
	.7	.3	positivo
jajaja creí serian preguntas más complicadas v	.7	.3	positivo
univo sorprendiéndome	.7	.3	positivo
interesante dinámica	.7	.3	positivo
buena dinámica	.7	.3	positivo
buena trivía conocedores	.7	.3	positivo
una dinámica fácil	.7	.3	positivo
buena estrategia activar mentalmente niños les pidieron estudiaran	.7	.3	positivo
bonita forma poner prueba niños saben	.7	.3	positivo
una dinámica interesante seamos externos escuela participa	.8	.2	positivo
pa ganen un juego mesa	.7	.3	positivo
buena trivía	.7	.3	positivo
una dinámica interesante seamos externos escuela participar	.8	.2	positivo
excelente centro educativo	.8	.2	positivo
nivel educativo excelente	.8	.2	positivo
excelente servicio atención e instalación	.8	.2	positivo
una excelente escuela excelente catedráticos	.8	.2	positivo
universidad dado gran satisfacción prepararme ayudó descubrir potencial hay	.7	.3	positivo
excelente escuela toda extensión palabra orgullosa hija estudie escuela mil felicitaciones	.8	.2	positivo
todos laboran			
una opción educativa valiosa esmero conjunto personal univo formar profesionistas exitosos	.7	.3	positivo
mejor institución educativa región orgullosamente soygenteunivo	.7	.3	positivo

carreras ofrece jóvenes estudian proyección sus alumnos extranjero profesionalismo sus maestros sus instalaciones	.7	.3	positivo
sido una experiencia aprendizaje genuinos crecimiento personal profesional orgullosa pertenecer Grupo Educativo Univo	.8	.2	positivo
irresponsabilidad avisar momento carreras afecta moral ética debería poseer adultos profesionistas afecta imagen personal labora institución	.5	.5	negativo
bilingual education is taken seriously with very professional staff	.7	.3	positivo
alto nivel educativo docentes actualizados planes estudios pensados retos futuro	.7	.3	positivo
perfecta enseñanza parte maestros soygenteunivo agradable	.7	.3	positivo
excelente universidad	.8	.2	positivo
calidad calidez prestigio profesionalismo todo excelencia académica	.8	.2	positivo
mejor universidad aprende ke carera kieras escojer	.7	.3	positivo
gustaría alguna vez pudieran conocer visitar grupo educativo	.7	.3	positivo
universidad prestigio además fomentan valores	.7	.3	positivo
una gran universidad sus valores visión compromiso social recuerdo gratitud eterna sus fundadores ing roberto lòpez arano qepd don mario martinez dector ing antonio gutièrrez hernàndez	.7	.3	positivo
buena información acerca acontecimientos escuela gustaría ver más fotos instalaciones	.7	.3	positivo
plataforma facebook buena presentación dinámica	.7	.3	positivo
publicaciones atractivasm deberian mas dinámicas	.7	.3	positivo
instalaciones buenas	.7	.3	positivo
página facebook dinamica da conocer información requiere conocer escuela manera digital	.7	.3	positivo
página facebook llamativa mucha información	.7	.3	positivo
publicaciones parecen llamativas divertidas	.7	.3	positivo
buena vinculación una buena comunicación usuarios	.7	.3	positivo
página parece completa interesante contenido bueno cuanto información brindada	.8	.2	positivo
gusta página facebook parece buena gusta actualizada fotos alumnos así actividades realizan institución	.4	.6	negativo
página bonita podrían poner más publicaciones exista más interacción volverla más dinámica información general	.7	.3	positivo
cuentan redes sociales dinámicas muchas fotos muchos videos	.3	.7	negativo
contenido hace indiferente	.7	.3	positivo

página padre sería mejor agrupara aspectos	.8	.2	positivo
gusta mucha actividad dejan abandonada	.4	.6	negativo
página cuenta mucha atracción captar atención consumidor	.7	.3	positivo
una página dinámica información completa buena difusión cuentas	.7	.3	positivo
interesante entrar página información dinámica actualizada	.7	.3	positivo
plataforma estructurada ve formal actualiza actividades universidad	.7	.3	positivo
publican constantemente actividades realiza institución	.7	.3	positivo
página buena información	.7	.3	positivo
impresión da acerca página buena información correcta difusión aparte dinámica	.7	.3	positivo
red social hace útil viable difusión información dado caso dinámica	.7	.3	positivo
contenido un poco aburrido sus publicaciones actividades apropiadas	.7	.3	positivo
página facebook contiene mucha información colores utiliza poco llamativos	.7	.3	positivo
interés	.7	.3	positivo
interés	.7	.3	positivo
redes sociales funcionales necesita una mayor participación alumnado	.4	.6	negativo
útiles	.7	.3	positivo

Referencias bibliográficas

- Balazr Thenot, J. A. (2015). Diseño, desarrollo e implementación de una aplicación de web opinion mining para indentificar el sentimiento de usuarios de Twitter con respecto a una compañía retail.
- Bayo-Moriones, A., & Lera-López, F. (2007). A firm-level analysis of determinants of ICT adoption in Spain. *Technovation*, 27(6-7), 352-366. <https://doi.org/10.1016/j.technovation.2007.01.003>
- Cenamor, J., Parida, V., & Wincent, J. (2019). How entrepreneurial SMEs compete through digital platforms: The roles of digital platform capability, network capability and ambidexterity. *Journal of Business Research*, 100, 196-206. <https://doi.org/10.1016/j.jbusres.2019.03.035>
- Cenamor, Rönnerberg Sjödin, D., & Parida, V. (2017). Adopting a platform approach in servitization: Leveraging the value of digitalization. *International Journal of Production Economics*, 192, 54-65. <https://doi.org/10.1016/j.ijpe.2016.12.033>
- Chaffey, D. (2011). *E-business & e-commerce management: Strategy, implementation and practice* (5th ed). Harlow, England ; New York: Pearson/Financial Times Prentice Hall.
- Chaffey, D., & Ellis-Chadwick, F. (2014). *Marketing Digital: Estrategia, implementación y práctica*. 632.
- Chaffey, D., & Smith, P. R. (2013). *EMarketing excellence: Planning and optimizing your digital Marketing* (4th ed). London ; New York: Routledge.
- Combe, C. (2006). Chapter 3—E-business markets and models. En C. Combe (Ed.), *Introduction to e-Business* (pp. 53-86). <https://doi.org/10.1016/B978-0-7506-6731-9.50007-7>

- Conde Sanchez, S. (2017). Aplicación de la Minería de Opiniones para la extracción del conocimiento significativo durante la creación de planes estratégicos de las PYMES.
- Deshmukh, J. S., & Tripathy, A. K. (2018). Entropy based classifier for cross-domain opinion mining. *Applied Computing and Informatics*, 14(1), 55-64.
<https://doi.org/10.1016/j.aci.2017.03.001>
- Echeverría Gustavo. (2009). *Marketing en Internet. Professional Tools*.
- Espinosa Roberto. (2014). *Marketing Mix: Las 4P's*. Recuperado de <https://robertoespinosa.es/2014/05/06/Marketing-mix-las-4ps-2/>
- Figuroa Rojas, S. C. (2015). *Estrategia de Marketing Digital en redes sociales para Pymes mexicanas: Aplicación al Centro Universitario de América (CUAM)*. Universidad Autónoma de México, Ciudad de México.
- Giuffra, J. (2019). *Marketing Digital*.
- Kotler, P. (2016). *Marketing Digital: Tendencias En Su Apoyo Al E-Commerce Y Sugerencias De Implementación*. 121.
- Kotler, P., Karajaya, H., & Setiawan, I. (2016). *Marketing 4.0: From Traditional to Digital*. Recuperado de <https://www.casadellibro.com/libro-Marketing-40/9788416894840/6459182>
- Lee, C. (2001). An analytical framework for evaluating e-commerce business models and strategies. *Internet Research*, 11(4), 349-359.
<https://doi.org/10.1108/10662240110402803>
- Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. 168.
- Mars, A., & Gouider, M. S. (2017). Big data analysis to Features Opinions Extraction of customer. *Procedia Computer Science*, 112, 906-916.
<https://doi.org/10.1016/j.procs.2017.08.114>

- Miranda, C. H., Guzmán, J., & Salcedo, D. (2016). Minería de Opiniones basado en la adaptación al español de ANEW sobre opiniones acerca de hoteles. 8.
- Montesinos García, L. (2014). Análisis de Sentimientos y Predicción de Eventos en Twitter.
- Othman, R., Belkaroui, R., & Faiz, R. (2017). Extracting Product Features for Opinion Mining Using Public Conversations in Twitter. *Procedia Computer Science*, 112, 927-935. <https://doi.org/10.1016/j.procs.2017.08.122>
- Parker, G., Van Alstyne, M., & Choudary, S. P. (2016). *Platform revolution: How networked markets are transforming the economy - and how to make them work for you* (First edition). New York London: W.W. Norton & Company.
- Ramesh, S., Dursun, D., & Turban, E. (2015). *Business intelligence and analytics: Systems for decision support* (Tenth edition). Boston: Pearson.
- Rogers, K. (2014). Where are our digital ads really going? Presentado en Where are our digital ads really going?
- Tabuenca, S. (2018). 6 etapas de la metodología SOSTAC® para hacer un Plan de Marketing. Recuperado 9 de junio de 2019, de <https://www.marketinet.com/blog/etapas-de-metodología-sostac-para-hacer-plan-de-Marketing>
- Tavakoli, M., Zhao, L., Heydari, A., & Nenadić, G. (2018). Extracting useful software development information from mobile application reviews: A survey of intelligent mining techniques and tools. *Expert Systems with Applications*, 113, 186-199. <https://doi.org/10.1016/j.eswa.2018.05.037>
- Thomas J. Russell, & W. Ronald Lane. (2005). *Kleppner: Publicity* (16ª ED). Prentice Hall.
- Tiago, M. T. P. M. B., & Veríssimo, J. M. C. (2014). Digital Marketing and social media: Why bother? *Business Horizons*, 57(6), 703-708. <https://doi.org/10.1016/j.bushor.2014.07.002>

- Tjahyanto, A., & Sisepaputra, B. (2017). The Utilization of Filter on Object-based Opinion Mining in Tourism Product Reviews. *Procedia Computer Science*, 124, 38-45. <https://doi.org/10.1016/j.procs.2017.12.127>
- Turban, E., Sharda, R. E., & Dursun, D. (2011). *Decision Support and Business Intelligence Systems*. Recuperado de <https://www.pearson.com/us/higher-education/product/Turban-Decision-Support-and-Business-Intelligence-Systems-9th-Edition/9780136107293.html>
- Valdebenito, F. I. O. (2014). MINERÍA DE OPINIÓN Y ANÁLISIS DE SENTIMIENTOS.
- Vásquez Rojas, C. (2014). "Desarrollo de un Modelo de Afinidad Utilizando Minería de Textos para el Análisis del Entorno y la Generación de Proyectos Estratégicos".
- Zurita Guerrero. (2017). El Marketing Digital y su influencia en el posicionamiento de los colegios privados del distrito de Trujillo, La Libertad-2017. Recuperado de http://repositorio.ucv.edu.pe/bitstream/handle/UCV/11792/zurita_gm.pdf?sequence=1