



UNIVERSIDAD CÉSAR VALLEJO

FACULTAD DE INGENIERÍA Y ARQUITECTURA

ESCUELA PROFESIONAL INGENIERÍA DE SISTEMAS

Mejora del Servicio de publicidad de empresas MYPE aplicando
minería de texto basadas en redes sociales. Caso Propuesto en
CyR PUBLICIDAD

TESIS PARA OBTENER EL TÍTULO PROFESIONAL DE:
INGENIERO DE SISTEMAS

AUTORES:

Cueva Ramos Anggelo Fabrizio (ORCID: 0000-0002-7602-3968)

Lopez Vara Luis Antonio (ORCID: 0000-0002-7188-9690)

ASESOR:

Mg.Ing. Carranza Barrena Wilfredo Eduardo (ORCID: 0000-0003-0845-1984)

LÍNEA DE INVESTIGACIÓN:

Sistemas de Información y Comunicaciones

LIMA – PERÚ

2021

Dedicatoria

Dedico esta tesis a Dios, que me guardo en este tiempo me dio sabiduría y no dejo que caiga en desgano ni conformidad, a nuestros padres que, con su fe, su amor y su fuerza siempre están para incentivarlos a seguir adelante.

Agradecimiento

A mis docentes y compañeros por permitirnos compartir una vida de experiencias, aprendizaje, a mi familia y a todos los que aportaron en esta etapa de mi vida.

Índice de contenidos

Carátula.....	i
Dedicatoria.....	ii
Agradecimiento.....	iii
Índice de contenidos.....	iv
Índice de tablas.....	v
Índice de figuras y gráficos.....	vi
Resumen.....	vii
Abstract.....	viii
I. INTRODUCCIÓN	11
II. MARCO TEÓRICO	14
III. METODOLOGÍA.....	22
3.1 TIPO Y DISEÑO DE INVESTIGACIÓN.....	22
3.1.1 Tipo de Investigación.....	22
3.1.2 Diseño de Investigación.....	22
3.2 METODOLOGÍA CRIS-DM	23
3.3 VARIABLE Y OPERACIONALIZACIÓN	23
3.4 POBLACIÓN, MUESTRA Y MUESTREO.....	28
3.5 TÉCNICAS E INSTRUMENTOS DE RECOLECCIÓN DE DATOS	28
3.5.1 Técnica para recolección de datos	28
3.5.1.1 La observación.....	28
3.5.2 Instrumentos para recolección y análisis de datos	29
3.5.2.1 Extracción de data en redes sociales	29
3.5.2.2 Facepager	29
3.5.2.3 Power BI desktop	30
3.6 PROCEDIMIENTOS.....	31
3.6.1 Ciclo de vida Metodología CRIS-DM	31
3.7 MÉTODO DE ANÁLISIS DE DATOS	36

3.7.2.7 Power BI Desktop.....	44
3.7.2.8 Power BI Service.....	45
3.7.2.9 Power BI Mobile	45
3.7.2.10 Power BI - Líder en Plataformas de Inteligencia de Negocios	45
3.7.2.11 Power BI - Fortalezas.....	46
3.7.2.12 Power Bi - Cuadro Comparativo de herramientas de inteligencia de negocios de autoservicio.....	47
3.7.3 Text Analytics	48
3.7.3.1 Análisis de los sentimientos.....	49
3.7.3.2 Extracción de frases clave.....	49
3.7.3.3 Analizar texto con el servicio Text Analytics	49
3.7.4 Algoritmos de aprendizaje.....	50
3.7.4.1 Comparación de algoritmos de aprendizaje.....	58
3.7.5 Machine Learning	59
3.7.6 Procesamiento de lenguaje natural.....	59
3.8 ASPECTOS ÉTICOS	60
3.8.1 Costos de servicios.....	60
IV. RESULTADOS	63
V. DISCUSIÓN	71
VI. CONCLUSIONES	73
VII. RECOMENDACIONES.....	74
REFERENCIAS	75

ÍNDICE DE TABLAS

TABLA N° 01: OPERACIÓN DE VARIABLES.....	24
TABLA N° 02: MATRIZ DE CONSISTENCIA.....	26
TABLA N° 03: TOTAL DE POST Y OPINIONES	32
TABLA N° 04: CRONOGRAMA.....	62
TABLA N° 05: RESULTADOS DEL EXPERIMENTO 1	63
TABLA N° 06: RESULTADOS DEL EXPERIMENTO 2.....	65
TABLA N° 07: RESULTADOS DEL EXPERIMENTO 3.....	67
TABLA N° 08: RESULTADOS DEL EXPERIMENTO 4.....	69

ÍNDICE DE FIGURAS

FIGURA N° 01: METODOLOGIA CRIS-DM	23
FIGURA N° 02: EXTRACCIÓN DE REDES	29
FIGURA N° 03: PROGRAMA FACEPAGER.....	30
FIGURA N° 04: PROGRAMA PROGRAMA POWER BI	30
FIGURA N° 05: CICLO DE VIDA DEL CRISP-DM	31
FIGURA N° 06: ARCHIVOS CSV	33
FIGURA N° 07: CAMPOS DE LOS ARCHIVOS CSV	33
FIGURA N° 08: CUADROS ESTADISTICOS.....	35
FIGURA N° 09: HERRAMIENTA POWER BI.....	39
FIGURA N° 10: ESTILOS DE POWER BI.....	42
FIGURA N° 11: POWER BI OPERACIONAL.....	43
FIGURA N° 12: FLUJO DE TRABAJO DE POWER BI.....	44
FUENTE N° 13: CUADRANTE MÁGICO DE GARTNER 2016.	46
FIGURA N° 14: CUADRO COMPARATIVO HERRAMIENTAS DE BI DE AUTOSERVICIO.	47
FIGURA N° 15: ALGORITMOS DE MÁQUINAS DE VECTORES DE SOPORTE (SVM)50	
FIGURA N° 16: K-MEDIAS O K-MEANS (MÉTODO DE AGRUPAMIENTO).....	51
FIGURA N° 17: REGRESIÓN LINEAL Y REGRESIÓN LOGÍSTICA	52
FIGURA N° 18: CLASIFICADOR BAYESIANO INGENUO.....	54
FIGURA N° 18: K-VECINOS MÁS PRÓXIMOS	55
FIGURA N° 19: K-VECINOS MÁS PRÓXIMOS	56
FIGURA N° 20: VECTOR DE CUANTIFICACIÓN VECTORIAL (LVQ).....	57
FIGURA N° 21: COMPARACIÓN DE ALGORITMOS DE APRENDIZAJE	58
FIGURA N° 21: COSTOS DE SERVICIOS MICROSOFT AZURE.....	60
FIGURA N° 22: COSTOS DE SERVICIOS MICROSOFT AZURE.....	61
FIGURA N° 23: IMAGEN DE EXPERIMENTO 1	63

FIGURA N° 24: PALABRAS MÁS UTILIZADAS	64
FIGURA N° 25: SCORE DE SENTIMIENTO DEL EXPERIMENTO 1.....	64
FIGURA N° 26: IMAGEN DEL EXPERIMENTO 2.....	65
FIGURA N° 27: PALABRAS MÁS UTILIZADAS	66
FIGURA N° 28: SCORE DE SENTIMIENTO DEL EXPERIMENTO 2.....	66
FIGURA N° 29: IMAGEN DEL EXPERIMENTO 3.....	67
FIGURA N° 30: PALABRAS MÁS UTILIZADAS	68
FIGURA N° 31: SCORE DE SENTIMIENTO DEL EXPERIMENTO 3.....	68
FIGURA N° 32: IMAGEN DEL EXPERIMENTO 4.....	69
FIGURA N° 33: PALABRAS MÁS UTILIZADAS	70
FIGURA N° 34: SCORE DE SENTIMIENTO DEL EXPERIMENTO 4.....	70

RESUMEN

La presente tesis detalla Mejora del Servicio de publicidad de empresas MYPE aplicando minería de texto basadas en redes sociales. La situación anterior presentaba el cómo saber si un servicio o producto lanzado en las redes sociales es acepto o no por el cliente. En este sentido el objetivo del estudio fue, determinar la manera en que apoya la minería de textos para extraer opiniones en redes sociales sobre el uso de un producto o servicio, la metodología utilizada en esta investigación fue Crisp-DM, siguiendo las fases de esta.

El tipo de investigación a realizarse fue de tipo aplicada y explicativa con diseño de investigación experimental, en el cual tuvo como población los usuarios la red social Facebook y la muestra estuvo conformada por las personas de la zona de Lima desde 18 a 50 años. Se empleo la herramienta Facepager y Power Bi para la extracción de datos y análisis de sentimientos. Obteniendo como resultados del primer experimento que el 18% de comentarios son positivos, 56% indiferentes y 26% negativos, del segundo experimento que el 23,08% de comentarios son positivos, 50% indiferentes y 26,92% negativos, del tercer experimento muestra que el 34,71% de comentarios son positivos, 38,84% indiferentes y 26,45% negativos, y por último el cuarto experimento muestra que el 18,18% son comentarios positivos, 44,16% indiferentes y 37,66% negativos. En conclusión, la herramienta Power BI muestran los sentimientos de cada usuario que opina sobre un producto o servicio de una empresa, manteniendo la fidelidad del cliente.

Palabras clave: Minería de Textos, Marketing, redes sociales, crisp-DM, análisis de sentimiento.

ABSTRACT

The present thesis details the improvement of the advertising service of MSE companies by applying text mining based on social networks. The previous situation presented the question of how to know if a service or product launched in social networks is accepted or not by the customer. In this sense, the objective of the study was to determine the way in which text mining supports to extract opinions in social networks about the use of a product or service, the methodology used in this research was Crisp-DM, following its phases.

The type of research to be carried out was applied and explanatory with experimental research design, in which the population was the users of the social network Facebook and the sample consisted of people in the Lima area from 18 to 50 years old. The Facepager and Power Bi tools were used for data extraction and sentiment analysis. The results of the first experiment showed that 18% of comments were positive, 56% indifferent and 26% negative; the second experiment showed that 23.08% of comments were positive, 50% indifferent and 26.92% negative; the third experiment showed that 34.71% of comments were positive, 38.84% indifferent and 26.45% negative and finally, the fourth experiment showed that 18.18% of comments were positive, 44.16% indifferent and 37.66% negative. In conclusion, these tools show the feelings of each user who gives an opinion about a product or service of a company.

Keywords: Text mining, marketing, social networks, crisp-DM, sentiment analysis.