

Análisis comparativo de métodos de *machine learning* para clasificar opiniones sobre el servicio de restaurantes peruanos en Facebook

Martín Jesús Adrianzén Torres
20160017@aloe.ulima.edu.pe

Edwin Jhonatan Escobedo Cárdenas
eescobed@ulima.edu.pe

Universidad de Lima, Perú

Recibido: 31/7/2021 Aceptado: 27/9/2021
doi: <https://doi.org/10.26439/ciis2021.5578>

RESUMEN. Las opiniones de los clientes sobre servicios en redes sociales son vitales para las empresas debido a que se pueden utilizar para mejorar y potenciar las oportunidades de negocio si los comentarios pueden analizarse a tiempo. El propósito de este trabajo es determinar los métodos de *machine learning* con mejor rendimiento para aplicar análisis de sentimientos y clasificar comentarios positivos y negativos sobre el servicio de restaurantes peruanos en Facebook. Como primera contribución en este proyecto, se crearon dos *datasets* de comentarios de publicaciones de cadenas de restaurantes peruanos en Facebook. La segunda contribución es la metodología propuesta dividida en dos etapas: en la primera etapa se aplicaron técnicas de lenguaje natural para el preprocesamiento de los comentarios; en la segunda etapa se analizó el desempeño de los algoritmos de Naïve Bayes, Random Forest y SVM con núcleos RBF y lineal para clasificar las opiniones en los *datasets*. Los resultados experimentales demostraron que el clasificador SVM obtuvo el mejor desempeño tanto en la etapa de entrenamiento como en la de pruebas con un 91,44 % y 94 % de exactitud para los *datasets* primario y secundario respectivamente, probando la viabilidad de la metodología propuesta.

PALABRAS CLAVE: análisis de sentimientos / procesamiento de lenguaje natural / *machine learning* / máquinas de soporte vectorial / Naïve Bayes / Random Forest / servicio de restaurantes

COMPARATIVE ANALYSIS OF MACHINE LEARNING METHODS TO CLASSIFY OPINIONS ABOUT THE PERUVIAN RESTAURANT SERVICE ON FACEBOOK

Abstract. Customer opinions on social media services are vital for companies because they can improve and enhance business opportunities if the comments can be analyzed in time. The purpose of this work is to determine the best performing machine learning methods to apply sentiment analysis and classify positive and negative comments about the Peruvian restaurant service on Facebook. As the first contribution to this project, two datasets of comments from Peruvian restaurant chain posts on Facebook were created. The second contribution is the proposed methodology divided into two stages: in the first stage, natural language techniques were applied for the pre-processing of comments; In the second stage, the performance of the Naïve Bayes, Random Forest, and SVM algorithms with RBF and Linear cores was analyzed to classify the opinions in the datasets. The experimental results demonstrated that the SVM classifier obtained the best performance in both the training and testing stages with 91.44% and 94% accuracy for the primary and secondary datasets, respectively, proving the viability of the proposed methodology.

KEYWORDS: sentiment analysis / natural language processin / machine learning / vector support machine / Naïve Bayes / Random Forest / restaurant service

1. INTRODUCCIÓN

El análisis de sentimientos es el estudio computacional de las opiniones de las personas, actitudes y emociones de una entidad. Autores como Ibrahim y Wang (2019) lo definen como un análisis que extrae opiniones, sentimientos y subjetividad de texto no estructurado para identificar las expresiones que indican opiniones positivas o negativas hacia un sujeto. El estudio es muy amplio y se pueden realizar diversos tipos de aplicaciones. Poornima y Priya (2020) afirman que, a través del análisis de sentimientos, se puede realizar un modelado de tópicos, que consiste en identificar palabras clave para resumir una gran cantidad de información textual; por ejemplo, He *et al.* (2016) analizaron opiniones en Facebook sobre una tienda de electrodomésticos identificando diferentes puntos de mejora en la empresa, entre ellos, el detalle de las descripciones de los productos en venta, lo que llevó a un progreso en la demanda. Por otro lado, Ibrahim y Wang (2019) recomendaron a empresas de *retail* mejorar puntos críticos en su servicio, como el tiempo de entrega y de atención, luego de analizar las opiniones de sus clientes en Twitter.

Otra modalidad en el análisis de sentimientos es usar técnicas de *machine learning* para predecir la polaridad de las oraciones o clasificarlas por aspectos o características en común; por ejemplo, en el rubro de restaurantes, el trabajo de Hossain *et al.* (2017) concluyó que el desempeño de un modelo basado en regresión logística rinde mejor con una exactitud del 79 % con comentarios colectados de una página de reseñas sobre la atención de restaurantes de Bangladés, India; en otro estudio, Krishna *et al.* (2019) demostraron que el clasificador de Máquinas de Soporte Vectorial obtuvo el mejor rendimiento con un 94,64 % de exactitud al momento de clasificar reseñas sobre la atención de restaurantes de Estados Unidos por polaridad (positiva o negativa). No obstante, Zahoor *et al.* (2020) determinaron que Random Forest es el mejor clasificador con una exactitud del 94 % para opiniones de clientes sobre el servicio de restaurantes de Karachi, Pakistán. De esta forma, se evidencia que no existe una técnica de *machine learning* general para el análisis de sentimientos aplicado en el rubro de restaurantes. Diversos factores, como la cantidad de comentarios analizados, las técnicas de procesamiento de lenguaje natural empleadas e incluso el idioma del país, son determinantes para la clasificación de los comentarios. Adicionalmente, son pocas las investigaciones enfocadas en el contexto peruano para detectar opiniones positivas o negativas sobre el servicio de venta de comida de restaurantes, sobre todo a través del análisis de comentarios en Facebook. En consecuencia, al no haber muchos trabajos en la literatura enfocados en este tema, no se pueden determinar los métodos de *machine learning* más adecuados para clasificar opiniones positivas o negativas en este contexto.

No obstante, el análisis de sentimientos aplicado en otras áreas ha demostrado ser un método viable para la identificación de aspectos relevantes en opiniones que ayuden a clasificarlas como positivas o negativas. Tradicionalmente, los diversos trabajos propuestos en la literatura comienzan con la preparación de los datos, conformados por comentarios

recolectados en redes sociales u otros medios; sobre estos se aplican técnicas de minería de texto para descubrir conocimiento que no existe explícitamente en el contenido de los datos (He *et al.*, 2016). Actualmente, existe una variedad de trabajos que abordan el análisis de sentimientos de una manera más detallada para obtener mejores resultados. Se trabaja en un preprocesamiento de texto, donde se filtran, tokenizan y derivan los comentarios. (Yulianto *et al.*, 2018). Después, se aplica un extractor de características y se usan algoritmos de *machine learning* como Naïve Bayes, Máquinas de Soporte Vectorial, Random Forest, etcétera, para clasificar las opiniones según la polaridad que se esté trabajando como, a su vez, evaluar el mejor algoritmo (Khan y Zubair, 2020).

Con base en lo mencionado, el objetivo de esta investigación es determinar los métodos de *machine learning* con mejor rendimiento para aplicar análisis de sentimientos y clasificar comentarios positivos y negativos sobre el servicio de restaurantes peruanos en Facebook. Para esto, se crearon dos *datasets* con comentarios que fueron preprocesados con técnicas de procesamiento de lenguaje natural. Posteriormente, se realizaron análisis experimentales con los clasificadores Máquinas de Soporte Vectorial con núcleos RBF y lineal, Naïve Bayes y Random Forest para determinar que técnica de *machine learning* posee un mejor desempeño en el contexto de estudio al momento de clasificar cada comentario como positivo o negativo.

Este trabajo está estructurado de la siguiente manera: en el capítulo 2 se detallará la metodología propuesta, en el capítulo 3 se presentarán los resultados obtenidos de los clasificadores y en el capítulo 4 se compartirán las conclusiones.

2. METODOLOGÍA

En la figura 1, se presenta la metodología propuesta. En primer lugar, se crearon dos *datasets* (primario y secundario) con comentarios de publicaciones de restaurantes con un extractor web. El *dataset* primario es utilizado en las etapas de entrenamiento y de pruebas, mientras que el *dataset* secundario es solo utilizado en la etapa de pruebas para verificar que el clasificador seleccionado generalice de forma adecuada. Para clasificar las opiniones de los clientes, los comentarios colectados pasaron por una etapa de preprocesamiento donde se aplican técnicas de PLN (procesamiento de lenguaje natural) y, posteriormente, se compararon tres de los clasificadores más utilizados en la literatura revisada, los cuales son: Naïve Bayes, Máquinas de Soporte Vectorial y Random Forest. A continuación, se explicará con mayor detalle la metodología propuesta.

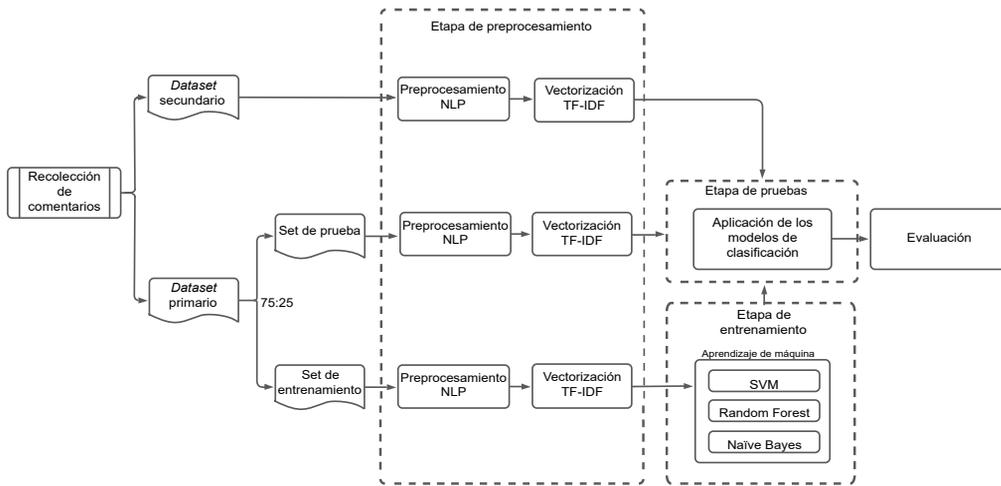


Figura 1. Flujo de metodología
Elaboración propia

2.1 Generación de dataset

Para probar la viabilidad de la metodología propuesta, se crearon dos *datasets* colectados de opiniones en páginas de Facebook pertenecientes a tres cadenas de restaurantes de Lima, Perú, con fechas de publicación desde mayo del 2020 hasta abril del 2021. Se seleccionaron publicaciones en fechas festivas como Día de la Madre, Fiestas Patrias, Navidad, Año Nuevo y otros, donde se puede encontrar una gran cantidad de comentarios con mayor facilidad. Los comentarios hacen referencia a todo el proceso de atención de los restaurantes, desde la recepción del pedido hasta su entrega, sea por medio del restaurante o de terceros (aplicaciones de *delivery*). Para la colecta de comentarios se utilizó un extractor web de la página Exportcomments, donde se coloca el enlace de la publicación y nos devuelve un archivo de extensión "*.xls" con los comentarios.

En el *dataset* primario, los comentarios pertenecen a publicaciones de acceso abierto sobre tres cadenas de restaurantes: Roky's (280 comentarios), Pardos (717 comentarios) y Don Tito (500 comentarios). A continuación, los comentarios se agruparon en un único *dataset* con 1494 comentarios en total. Después, se etiquetó de forma manual el sentimiento de cada comentario como positivo o negativo dependiendo de la expresión que representa la opinión de los clientes; por ejemplo, si el comentario expresa algún reclamo del servicio o parecido, se etiquetó como negativo; en cambio, si el comentario expresa satisfacción del servicio o parecido, se etiquetó como positivo. En total se etiquetaron 769 comentarios como negativos y 725 comentarios como positivos, formando un *dataset* con datos balanceados que ayudará a evitar problemas de sesgo en la etapa de clasificación. En los comentarios se consideraron

emoticones, jergas, palabras soeces, etcétera. En la tabla 1 se puede observar un fragmento del *dataset*.

De la misma manera, se decidió crear un *dataset* secundario con comentarios que pertenecen a publicaciones de la cadena de restaurantes Norky's, obteniéndose en total 100 comentarios, de los cuales 38 son positivos y 62 negativos. Este *dataset* fue utilizado para analizar el comportamiento del modelo seleccionado en la etapa de pruebas al clasificar un nuevo tipo de comentarios sobre el servicio de un restaurante diferente a los que se usaron para entrenar los modelos propuestos. En la tabla 2 se presenta un fragmento del *dataset*.

Tabla 1
Fragmento del dataset primario

Comentario	Sentimiento
Excelente pollería muy deliciosos y buena atención	1
Muy rico y sabroso	1
Son deliciosos pollos roky mejor comer en familia	1
Siguen dando su aji tradicional? Porque la última vez nos dieron un ají horrible	0
No responden ningún canal de atención	0

Elaboración propia

Tabla 2
Fragmento del dataset secundario

Comentario	Sentimiento
Mejor busco a la tía veneno	0
porque dicen que hay delivery hasta las 9 y el de la avenida mexico ni contesta el telefono... en realidad casi nunca contestan asi sea dia de semana	0
Quiero!!	1
Mejor el pollo es villa chicken	0
que buena promooo	1

Elaboración propia

2.2 Etapa de preprocesamiento aplicando PLN

En esta etapa, los comentarios son convertidos en un formato numérico (vector de características) para que puedan trabajar con las técnicas de *machine learning* y entrenar los modelos

predictivos. Para comenzar, se creó una función llamada *Tokenize* que recibe como parámetro un comentario y realiza las siguientes funciones:

- Permite hacer *stemming* para eliminar los sufijos de las palabras de cada comentario.
- Elimina los signos de puntuación.
- Tokeniza cada oración en unigramas, es decir, divide cada comentario palabra por palabra.

En la figura 2 podemos ver un ejemplo de su implementación paso a paso:

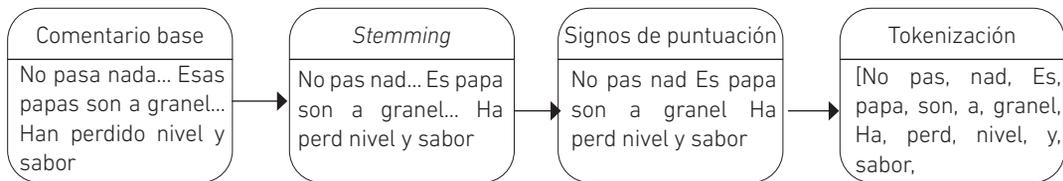


Figura 2. Implementación de *Tokenize* en un comentario
Elaboración propia

Luego, se vectorizó cada comentario para que los algoritmos de *machine learning* puedan recibirlos como parámetros de entrada; para esto se usó el vectorizador TF-IDF que convierte el nivel léxico de los comentarios a numérico dependiendo de la frecuencia de la palabra en el documento (Madasu y Elango, 2020). El vectorizador TF-IDF también considera a los emoticones y su tratamiento es similar a una palabra común. En la tabla 3 se aprecia una parte de la matriz TF-IDF de nuestro *dataset*, donde las columnas representan a las palabras de todos los comentarios ya procesados y las filas representan a cada uno de los comentarios. Si una palabra pertenece a un comentario, se le otorga un puntaje dependiendo de su frecuencia; en caso contrario, se le otorga un puntaje de cero. De la misma manera, se encarga de transformar cada palabra de los comentarios en minúsculas.

Tabla 3
Fragmento de la matriz TF-IDF

	Común	comunic	conchud	concienci	concosc	conf	confi	confirm
0	0,0	0,000000	0,0	0,0	0,0	0,0	0,000000	0,0
1	0,0	0,386243	0,0	0,0	0,0	0,0	0,000000	0,0
2	0,0	0,000000	0,0	0,0	0,0	0,0	0,000000	0,0
3	0,0	0,000000	0,0	0,0	0,0	0,0	0,000000	0,0
4	0,0	0,000000	0,0	0,0	0,0	0,0	0,538901	0,0

Elaboración propia

3. RESULTADOS EXPERIMENTALES

En este punto se presentan los resultados obtenidos de la propuesta de solución, comenzando por los protocolos que se siguieron en la experimentación, los hiperparámetros usados para los algoritmos de aprendizaje de máquina, los valores obtenidos de la validación cruzada en la etapa de entrenamiento y las métricas de rendimiento del modelo de predicción elegido en la etapa de pruebas.

3.1 Protocolo de experimentación

En esta parte se definieron algunos pasos que siguieron todos los modelos por igual al momento de su construcción. Para comenzar, el *dataset* primario se dividió de forma aleatoria en un 75 % para el set de entrenamiento y un 25 % para el set de pruebas. Después se probaron y combinaron distintos valores de hiperparámetros correspondientes a los algoritmos de Naïve Bayes, Random Forest y Máquinas de Soporte Vectorial (SVM, por sus siglas en inglés) con sus núcleos RBF y lineal; y se eligieron los que otorgaron una mejor exactitud. A continuación, en la tabla 4 se pueden apreciar los hiperparámetros utilizados en cada algoritmo y sus definiciones:

Tabla 4
Hiperparámetros de prueba

Algoritmo	Hiperparámetros	Definición
SVM lineal y RBF	C: [1,10,100,1000,10000]	Costo de clasificación errónea
SVM RBF	Gamma: [0.0001, 0.001, 0.01, 0.1, 0.2, 0.5, 0.6, 0.9]	
	Intensidad de curvatura en el límite de decisión	
Naïve Bayes	Alpha: [1, 0.1, 0.01, 0.001, 0.0001, 0.00001]	
	Valor que aborda el problema de tener una probabilidad cero	
Random Forest	Max Depth: [3, 5, 10, 25]	Número de profundidad máxima que pueden alcanzar los árboles
	Max Features: [3, 5, 10, 20]	Número de características para tener en cuenta a la hora de buscar el mejor Split
	Max Simple Slipt: [3, 5, 10]	Número máximo de muestras necesarias para dividir un nodo interno.
	N Estimators: [100, 300]	Número de árboles que se van a incluir en el modelo.

Elaboración propia

3.2 Modelos experimentales

Se aplicó el método *grid search* para obtener la mejor combinación de hiperparámetros. Para las Máquinas de Soporte Vectorial, con núcleo Lineal, el hiperparámetro seleccionado fue de $C = 1$; para las Máquinas de Soporte Vectorial, con núcleo RBF, los hiperparámetros seleccionados fueron $C = 100$ y $\gamma = 0,01$; para Naïve Bayes el hiperparámetro seleccionado fue de $\text{Alpha} = 1$ y en el caso de Random Forest los hiperparámetros seleccionados fueron de $\text{Max Depth} = 25$, $\text{Max Features} = 10$, $\text{Min Samples Split} = 5$ y $\text{N Estimators} = 100$.

3.3 Resultados en la etapa de entrenamiento

En la etapa de entrenamiento se buscó elegir el modelo más estable con un mayor porcentaje de exactitud. Se utilizó la técnica de validación cruzada en el set de entrenamiento para cada modelo predictivo con una división de $k\text{fold} = 10$. Dentro de cada división se seleccionó una pequeña parte llamada set de validación, en la cual el modelo predijo el sentimiento de los comentarios de este set. Este paso se repite en cada división y se obtuvo el promedio total de las exactitudes y la desviación estándar que nos indica la estabilidad de cada modelo.

En la tabla 5 se aprecian las exactitudes de cada modelo; para las Máquinas de Soporte Vectorial RBF, el máximo valor fue de 92,85 % y el menor fue de 88,39 %; en las Máquinas de Soporte Vectorial Lineal, el máximo valor fue de 94,64 % y el menor fue de 89,28 %; para Naïve Bayes, el máximo valor fue de 96,42 % y el menor fue de 85,71 %; en Random Forest, el máximo valor fue de 94,64 % y el menor fue de 80,35 %. Además, se pueden observar los promedios de las exactitudes y las desviaciones estándar de cada modelo. Según los resultados, el modelo predictivo de aprendizaje de máquina Máquinas de Soporte Vectorial (SVM, por sus siglas en inglés) con kernel Linear obtuvo la mayor exactitud con 91,70 % y una desviación estándar de 0,02, indicando que cuenta con buena estabilidad.

Tabla 5
Validación de los clasificadores (porcentajes)

K-Fold	SVM RBF	SVM Lineal	Naïve Bayes	Random Forest
1	91,96	90,17	95,53	82,14
2	91,07	91,07	88,39	87,50
3	92,85	92,85	84,82	80,35
4	93,75	94,64	87,50	94,64
5	91,96	94,64	88,39	92,85
6	90,17	91,07	85,71	88,39
7	89,28	90,17	96,42	82,14
8	90,17	89,28	88,39	84,82

(continúa)

(continuación)

9	88,39	91,07	88,39	90,17
10	92,85	91,96	91,07	89,28
Promedio exactitud	91,25	91,70	89,46	87,05
Desviación estándar	0,02	0,02	0,04	0,04

Elaboración propia

Podemos determinar que el modelo de las Máquinas de Soporte Vectorial lineal demuestra que puede clasificar en la mayoría de los casos con diversos tipos de comentarios con una exactitud generalizada. Según Islam *et al.* (2017), esto es debido a que las Máquinas de Soporte Vectorial, en todos sus núcleos como lineal, RBF y polinomial, cuenta con un tipo de “protección automática” ante posibles casos de *overfitting*, haciendo que su rendimiento sea el mejor en la validación cruzada en comparación de los otros modelos.

3.4 Resultados en la etapa de pruebas y evaluación

En la tabla 6 se puede apreciar que en el caso del set de pruebas del *dataset* primario, SVM Lineal obtuvo la mayor exactitud con un 91,44 % y un F1-Score de 0,91. Para el caso del *dataset* secundario, SVM Lineal y RBF obtuvieron la mayor exactitud con un 94 % y un F1-Score de 0,92 para ambos clasificadores.

Tabla 6
Reporte de clasificación

Clasificador	<i>Dataset</i>	Precisión	Exhaustividad	F1-Score	Exactitud
SVM RBF	Primario	0,91	0,91	0,91	90,90
	Secundario	0,92	0,92	0,92	94,00
SVM lineal	Primario	0,92	0,91	0,91	91,44
	Secundario	0,92	0,92	0,92	94,00
Naïve Bayes	Primario	0,94	0,87	0,90	90,64
	Secundario	0,94	0,87	0,90	93,00
Random Forest	Primario	0,92	0,85	0,88	88,50
	Secundario	0,91	0,81	0,86	90,00

Elaboración propia

En la tabla 7 se aprecia la matriz de confusión del modelo de las Máquinas de Soporte Vectorial lineal en el set de pruebas del *dataset* primario. De los 184 comentarios negativos, 14

fueron clasificados de forma incorrecta, y de los 190 comentarios positivos, 18 fueron clasificados de forma incorrecta.

Tabla 7
Matriz de confusión del set de pruebas del dataset primario

Actual	Predicción	
	SVM lineal	
	Negativos	170
	Positivos	18
		Positivos
		172

Elaboración propia

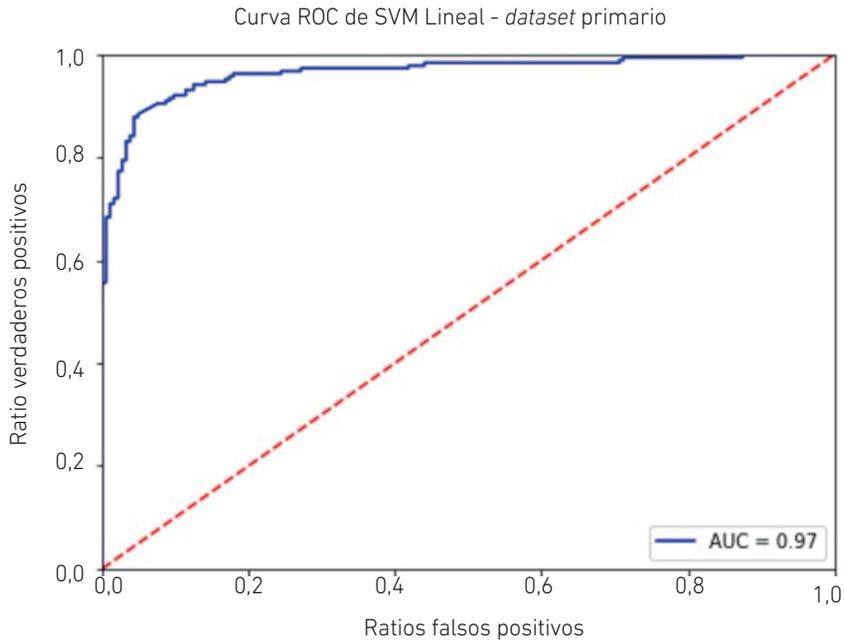
En la tabla 8 se observa la matriz de confusión del modelo de las Máquinas de Soporte Vectorial Lineal en el *dataset* secundario. De los 62 comentarios negativos, 3 fueron clasificados de forma incorrecta, y de los 38 comentarios positivos, 3 fueron clasificados de forma incorrecta.

Tabla 8
Matriz de confusión del dataset secundario

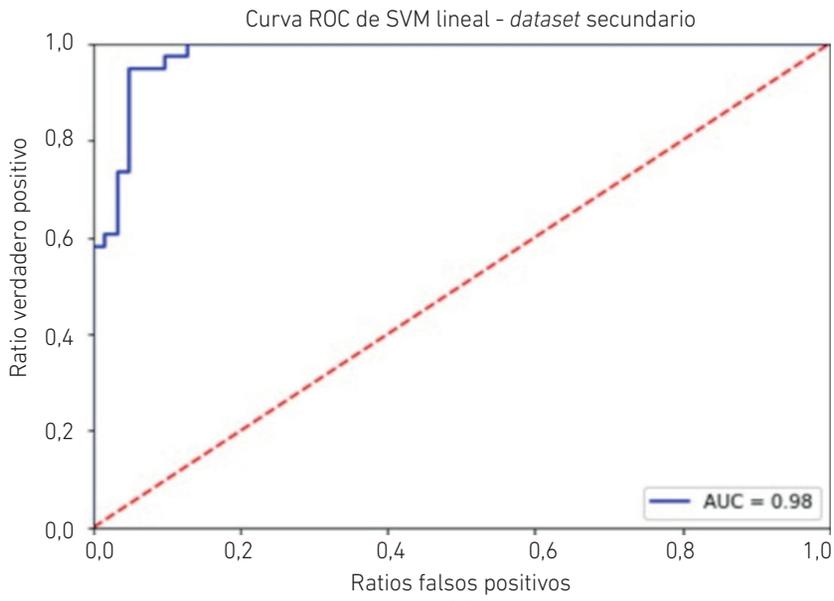
Actual	Predicción	
	SVM lineal	
	Negativos	59
	Positivos	3
		Positivos
		35

Elaboración propia

En la figura 3 se presentan las curvas ROC y AUC del modelo de las Máquinas de Soporte Vectorial con núcleo lineal en los dos escenarios. En el set de pruebas del *dataset* primario se tiene un AUC de 0,97; esto nos quiere decir que un comentario negativo seleccionado aleatoriamente tiene un mayor valor de prueba que uno seleccionado aleatoriamente de los comentarios positivos en el 97 % de las veces. Para el *dataset* secundario se obtuvo un AUC de 0,98; esto significa que un comentario negativo seleccionado aleatoriamente tiene un mayor valor de prueba que uno seleccionado aleatoriamente de los comentarios positivos en el 98 % de las veces.



(a)



(b)

Figura 3. (a) Curvas ROC y (b) AUC de los datasets

Elaboración propia

Finalmente, se analizaron los comentarios del *dataset* secundario y del set de pruebas del *dataset* primario que fueron clasificados erróneamente por el modelo de las Máquinas de Soporte Vectorial con núcleo lineal como falsos positivos; algunos comentarios fueron: “Mejor sirven en el corralito”, “Pollo Corralito ... el mejor ...”, “Mejor busco a la tía veneno”, “Mejor el pollo es villa chicken”. Como se puede apreciar, los comentarios que hacen referencia a los servicios de las competencias “Corralito” y “Villa Chicken” son clasificados de forma incorrecta; una de las razones es por la palabra “Mejor” que tiene una gran relación con varios comentarios verdaderos positivos y que fueron parte del entrenamiento del modelo.

Como falsos negativos algunos comentarios fueron: “Acabo de pedir por fchat llego al toque buenaso”, “Todas las semanas siempre compro por delivery son raudo para entrega amabilidad y cumplidos”, “Diosito el ají de pardos”, “Fanática del pollo norkys”, “Norkys es uno de los mejores por la q está en lince 21 de la Av Arequipa... felicitaciones al brasero ...”. En estos casos se puede observar que las fallas ortográficas pueden hacer que los modelos clasifiquen de forma incorrecta, como en el caso de “fchat”, “raudo”, “q”. Así como también, se puede apreciar un emoticón que tiene un significado ambiguo para el modelo debido a que aún no se realiza un tipo de análisis semántico.

A continuación, en la tabla 9 se presenta una comparación de resultados con otras investigaciones que utilizaron metodologías similares. Los trabajos de Zahoor *et al.* (2020) y Krishna *et al.* (2019) obtuvieron buenas exactitudes de sus algoritmos propuestos en sus investigaciones en restaurantes de Karachi y Estados Unidos respectivamente. Además, en el trabajo de Zahoor *et al.* (2020) el algoritmo que obtuvo una mayor exactitud fue el de Máquinas de Soporte Vectorial (SVM, por sus siglas en inglés) con un 94,56 %, mismo algoritmo que predominó en nuestra metodología propuesta, pero aplicada a opiniones sobre el servicio de restaurantes peruanos con una exactitud de 91,44 % en su núcleo Lineal.

Tabla 9
Comparación de resultados con otros autores (porcentajes)

Autores	Dataset	Algoritmos	Exactitud
Zahoor <i>et al.</i> (2020)	"SWOT's Karachi Restaurant Reviews"	Naïve Bayes	92
		SVM	92
		Random Forest	93
		Regresión Logística	95
Krishna <i>et al.</i> (2019)	Dataset de reseñas de un restaurante de Estados Unidos	Naïve Bayes	89,13
		SVM	94,56
		Random Forest	93,32

(continúa)

(continuación)

Hossain et al. (2017)	Dataset de reseñas de restaurantes de Bangladés	Naïve Bayes	77
		SVM	73
		KNN	75
		Regresión Logística	79
Metodología propuesta	Dataset propio de reseñas de restaurantes peruanos	Naïve Bayes	90,64
		SVM RBF	90,90
		Random Forest	88,50

Elaboración propia

4. CONCLUSIONES

En este trabajo se propuso un análisis comparativo de técnicas de *machine learning* para clasificar opiniones sobre el servicio de restaurantes peruanos en Facebook. Para este propósito, se recolectaron comentarios de publicaciones de restaurantes peruanos con un extractor web para elaborar dos *datasets*. Después, los textos fueron preprocesados con técnicas de procesamiento de lenguaje natural (*Tokenize*, *Stemming*, *Lower Case* y *Stop Words*). Con uno de los *datasets*, se entrenaron modelos predictivos con los algoritmos de *machine learning* de Naïve Bayes, Random Forest y Máquinas de Soporte Vectorial con núcleos RBF y Lineal. La mayor exactitud lo obtuvo el modelo de Máquinas de Soporte Vectorial lineal con un 91,44 % para el set de pruebas del *dataset* primario. De la misma manera, con el segundo *dataset* se buscó analizar el comportamiento de los anteriores modelos ante opiniones de un restaurante diferente a los del primer *dataset*, donde se presenta un empate en la exactitud en los modelos de Máquinas de Soporte Vectorial RBF y Lineal con un 94 %.

El análisis experimental nos demuestra que los algoritmos de aprendizaje de máquina de las Máquinas de Soporte Vectorial con núcleos RBF y lineal poseen mejor desempeño al momento de clasificar comentarios sobre el servicio de restaurantes peruanos en Facebook, obteniendo resultados superiores en los dos *datasets* de elaboración propia, demostrando la viabilidad de la metodología propuesta. Como trabajo futuro se puede ampliar el tamaño del *dataset* y el número de etiquetas (sentimientos) para así poder trabajar con redes neuronales artificiales, así como también comparar otros tipos modelos de aprendizaje de máquina, clasificar opiniones pertenecientes a otras redes sociales y realizar otros tipos de análisis como el modelado de tópicos.

REFERENCIAS

- He, W., Zha, S., y Li, L. (2016). Social Media Competitive Analysis and Text Mining: A Case Study in the Pizza Industry. *International Journal of Information Management*, 33(3), 464-472. <https://doi.org/10.1016/j.ijinfomgt.2013.01.001>
- Hossain, F. M. T., Hossain, M. I., y Nawshin, S. (2017). Machine Learning Based Class Level Prediction of Restaurant Reviews. *2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*, 420-423. <https://doi.org/10.1109/R10-HTC.2017.8288989>
- Ibrahim, N. F., y Wang, X. (2019). A Text Analytics Approach for Online Retailing Service Improvement: Evidence from Twitter. *Decision Support Systems*, 121, 37-50. <https://doi.org/10.1016/j.dss.2019.03.002>
- Islam, M., Jubayer, F., y Ahmed, S. (2017). A Support Vector Machine Mixed with TF-IDF Algorithm to Categorize Bengali Document. *2017 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, 191-196. <https://doi.org/10.1109/ecace.2017.7912904>
- Khan, A. H., y Zubair, M. (2020). Classification of Multi-Lingual Tweets, into Multi-Class Model Using Naïve Bayes and Semi-Supervised Learning. *Multimedia Tools and Applications*, 79(43-44), 32749-32767. <https://doi.org/10.1007/s11042-020-09512-2>
- Krishna, A., Akhilesh, V., Aich, A., y Hegde, C. (2019). Sentiment Analysis of Restaurant Reviews Using Machine Learning Techniques. En V. Sridhar, M. Padma y K. Rao (Eds.), *Emerging Research in Electronics, Computer Science and Technology* (pp. 687-696). Springer. https://doi.org/10.1007/978-981-13-5802-9_60
- Madasu, A., y Elango, S. (2020). Efficient Feature Selection Techniques for Sentiment Analysis. *Multimedia Tools and Applications*, 79(9-10), 6313-6335. <https://doi.org/10.1007/s11042-019-08409-z>
- Poornima, A., y Priya, K. S. (2020). A Comparative Sentiment Analysis of Sentence Embedding Using Machine Learning Techniques. En *2020 6th International Conference on Advanced Computing and Communication Systems, ICACCS*, 493-496. <https://doi.org/10.1109/ICACCS48705.2020.9074312>
- Yulianto, M., Girsang, A. S., y Rumagit, R. Y. (2018). Business Intelligence for Social Media Interaction in the Travel Industry in Indonesia. *Journal of Intelligence Studies in Business*, 8(2), 77-84. <https://doi.org/10.37380/JISIB.V8I2.323>
- Zahoor, K., Bawany, N. Z., y Hamid, S. (2020). Sentiment Analysis and Classification of Restaurant Reviews Using Machine Learning. *2020 21st International Arab Conference on Information Technology (ACIT)*. <https://doi.org/10.1109/ACIT50332.2020.9300098>

