

Aprendizaje Automático en la clasificación de Sentimientos en Twitter: Comparación de los Algoritmos de LSTM Bidireccional, Regresión Logística y Naïve Bayes

Paula Monge Valverde, Emmanuel Zúñiga Chaves

Escuela de Ciencias de la Computación e Informática, Universidad de Costa Rica

San Pedro de Montes de Oca, Costa Rica

`mariapaula.monge@ucr.ac.cr`

`emmanuel.zunigachaves@ucr.ac.cr`

Abstract—

I. INTRODUCCIÓN

El auge de las redes sociales y el uso diario que le dan los usuarios hace de interés analizar qué es lo que se está expresando por medio de estas, es por eso que surge el análisis de sentimientos en Twitter, esto se logra haciendo uso de algoritmos, los cuales permiten clasificar las tendencias de opinión del público en general, ya sea en ámbitos sociales, políticos o de la vida diaria.

Consecuentemente, la cantidad de investigaciones relacionadas al tema ha ido en aumento, por ende, existe una enorme variedad de algoritmos que se pueden utilizar para la clasificación de las cargas emocionales asociadas a dichos textos. La aplicación de estos modelos ha dado buenos resultados en un alto porcentaje de los casos, sin embargo, la mayoría de la información proporcionada y de algoritmos creados se basan en un corpus en inglés (tweets escritos en inglés). Sin embargo, cada vez son más los textos que utilizan otros idiomas. Si bien el inglés es la lengua predominante en Internet, hay otros idiomas como el español que cada vez tiene más presencia en las redes sociales, es por esto que nos lleva a preguntarnos ¿Cómo será su funcionamiento basado en un corpus de tweets en español?

Es de gran relevancia investigar y comprobar cómo se comportan varios clasificadores entrenados para determinar la polaridad de las opiniones en dicho corpus en español, puesto que, además de querer saber que se expresa en este idioma es importante medir su exactitud de clasificación cuando se cambia de lenguaje ya que se ha evidenciado que algoritmos como **Máquinas de Soporte Vectorial (SVM)** proporciona resultados fiables en textos en inglés, mientras

que en textos en español el máximo nivel de “accuracy” (indica cuántas veces el modelo fue correcto, en este caso en la clasificación) alcanzado apenas supera el 72%. Se trata de una cifra que todavía no resulta aceptable para la investigación [1].

Este artículo busca explorar y evidenciar los distintos algoritmos ampliamente utilizados para el análisis de sentimientos en Twitter basados en un corpus en español, con el fin de exponer cuáles de ellos tienen un mejor rendimiento de clasificación (basándose en el accuracy). El funcionamiento de dichos algoritmos, se centra en la clasificación de los sentimientos, a partir de un análisis de las características sintácticas y semánticas del texto en cuestión [2]. Los algoritmos seleccionados para el desarrollo del análisis serán: **LSTM Bidireccional, Regresión Logística y Naïve Bayes**; la razón de dicha selección, recae en el amplio uso que se les da en dicha área, es así como surge la necesidad de identificar el algoritmo más óptimo para la clasificación de tweets en español.

II. TRABAJOS RELACIONADOS

El internet es un amplio entorno en el cual se puede recopilar una gran cantidad de información [3], la cual enriquece la capacidad de análisis de diferentes patrones de comportamiento, opinión y sentimientos de los usuarios en diferentes plataformas. A consecuencia de esto surgen distintas alternativas en cuanto a opciones para la clasificación de sentimientos en tweets, por lo tanto, gran parte de los trabajos relacionados se enfocan en ofrecer un análisis sobre los resultados obtenidos al aplicar dichas técnicas.

El autor en [4], realiza una explicación teórica sobre cómo podría ser implementado **Naïve Bayes** en la

clasificación de sentimientos en los tweets; donde se plantea que con Naive Bayes es posible determinar cómo contribuye cada palabra al sentimiento asociado, que puede calcularse mediante la relación de la probabilidad de aparición de la palabra para la etiquetas clasificatorias. Similarmente, los autores en [5] efectúan una implementación del modelo basado en **Naive Bayes**, para ello realizan la experimentación con y sin la selección de características de información mutua (MI) para escoger las más relevantes dentro del conjunto de datos de tweets, obteniendo resultados de accuracy de entre el 96.2% y el 97.9% en el experimento implementado con MI.

Por otro lado, otras investigaciones se centran más en hacer un análisis del corpus dado un contexto. En [6] los autores realizan una implementación de un modelo de clasificación basado en **BERT**. Para ello, se recurrió a un conjunto de datos con más de un millón de tweets, clasificados según el sentimiento (positivo, negativo, neutral), obteniendo un accuracy del 92%, del cual se deduce que los modelos del lenguaje de **BERT** contribuyen a la obtención de buenos resultados en la clasificación de textos.

III. METODOLOGÍA

Cuando se quiere evaluar el rendimiento de un *programa de clasificación* se suelen usar las métricas de precision, recall, F1-Score y accuracy.

Cada una de ellas aporta distintos datos que permiten analizar el funcionamiento de la clasificación; **precision** indica el porcentaje de todas las predicciones positivas cuáles de ellas son correctas, **recall** indica el porcentaje de todos los positivos cuáles son reales (TP), **F1-Score** se trata de una mezcla de las dos métricas mencionadas anteriormente, haciéndolas funcionar de manera armónica/equilibrada sin embargo, la métrica que más interesa es **accuracy**, esta nos brinda el porcentaje de casos que el modelo ha acertado, es decir, nos indica que tan exacta fue la clasificación de los datos, si bien esta métrica es conocida por ser “engañosa”, esta característica se da cuando se poseen datos con clases desbalanceadas (una categoría con *muchísimos* mas datos que otra).

Es por esto que realizamos una integración de distintos datasets, para evitar un desbalance de las clases, se usaron 3 distintos datasets [PONER REFERENCIA] con los cuales se llega a un total de 15.000 mil tweets en español los cuales están divididos en las siguientes clases: “litigious”, “positive”, “negative” y “uncertainty”.

IV. ANÁLISIS DE RESULTADOS

V. DISCUSIONES Y CONCLUSIONES

REFERENCIAS

- [1] T. Baviera, “Técnicas para el análisis del sentimiento en Twitter: Aprendizaje Automático Supervisado y SentiStrength”, *Dígitos*, pp. (38-40), Dic, 2017
- [2] P. Mollah, “An LSTM model for Twitter Sentiment Analysis”, 2022, <https://arxiv.org/abs/2212.01791>
- [3] ON. A. Awad and A. Mahmoud, “Analyzing customer reviews on social media via applying association rule,” *Computers, Materials and Continua*, vol. 68, no. 2, pp. 1519–1530, Apr. 2021, doi: 10.32604/CMC.2021.016974.
- [4] H. Khandewal. "Sentiment Analysis of a Tweet With Naive Bayes". Medium. <https://towardsdatascience.com/sentiment-analysis-of-a-tweet-with-naive-bayes-ff9bdb2949c7>
- [5] M. A. Ulfa, B. Irmawati, and A. Y. Husodo, "Twitter Sentiment Analysis using Naive Bayes Classifier with Mutual Information Feature Selection," *J. Comput. Sci. Inform. Eng. (J-Cosine)*, vol. 2, no. 2, pp. 106-111, Dec. 2018, doi: <https://doi.org/10.29303/jcosine.v2i2.120>
- [6] A. Chiorrini, C. Diamantini, A. Mircoli, and D. Potena, "Emotion and sentiment analysis of tweets using BERT," in *Proceedings of the EDBT/ICDT Workshops*, Nicosia, Cyprus, Mar. 23-26, 2021.