
Revisión de Técnicas de Análisis de Sentimiento en Redes Sociales

Review of Sentiment Analysis Techniques in Social Networks

Brian Luis Motta Ypanaqué

brian.motta@unmsm.edu.pe

Universidad Nacional Mayor de San Marcos, Facultad de Ingeniería de Sistemas e Informática, Lima, Perú

Ana María Huayna Dueñas

ahuaynad@unmsm.edu.pe

RECIBIDO: 17/10/2023 - ACEPTADO: 25/11/2023 - PUBLICADO: 30/12/2023

RESUMEN

El presente artículo desarrolla una revisión de las técnicas usadas para el análisis de sentimiento aplicado a mensajes en redes sociales. El análisis de sentimiento es una tarea del campo de la Inteligencia Artificial conocido como Procesamiento de Lenguaje Natural y busca detectar la polaridad de sentimiento expresada por una persona en un mensaje corto o en un documento. Actualmente existen tres tipos de técnicas de las que se derivan los modelos de análisis de sentimiento propuestos: Diccionarios, aprendizaje automático tradicional y aprendizaje profundo; además de los enfoques híbridos que combinan al menos dos de estas técnicas para mejorar el rendimiento de la clasificación. El análisis de sentimiento es una tarea vigente y su importancia ha crecido con la masificación de las redes sociales que permiten la generación masiva de texto susceptible de ser clasificado.

Palabras clave: Análisis de sentimiento, redes sociales, diccionario, aprendizaje automático, aprendizaje profundo.

ABSTRACT

This article develops a review of the techniques used for sentiment analysis applied to messages on social networks. Sentiment analysis is a task from the field of Artificial Intelligence known as Natural Language Processing and seeks to detect the sentiment polarity expressed by a person in a short message or in a document. Currently there are three types of techniques from which the proposed sentiment analysis models are derived: Lexicons, traditional machine learning and deep learning – in addition to hybrid approaches that combine at least two of these techniques to improve classification performance. Sentiment analysis is a currently-valid task and its importance has grown along with the massification of social networks, which allow a massive generation of text that can be classified.

Keywords: Sentiment analysis, social networks, lexicon, machine learning, deep learning.

I. INTRODUCCIÓN

Desde la masificación de las redes sociales a inicios de la década del 2010, el análisis de sentimiento se ha convertido en uno de los campos de investigación de la Inteligencia Artificial (IA) más activos de acuerdo al creciente número de publicaciones hechos en los principales repositorios académicos y de investigación (Rodríguez-Ibáñez et al., 2023).

Plataformas como Facebook, Twitter (ahora renombrada como 'X') o Instagram se han convertido en espacios donde las personas pueden registrarse sin costo y crear contenido textual y multimedia, generando comentarios para transmitir información, pero también para expresar una opinión o actitud hacia muy diversos temas. Como resultado, se ha generado a lo largo de los años una cantidad masiva de datos susceptible de ser analizados. Esto ha hecho posible que los modelos de aprendizaje automático, muchos de los cuales fueron desarrollados de manera incipiente antes del año 2000, puedan entrenarse en diversas tareas de IA que antes del 2008 hubieran sido imposibles de realizarse de manera efectiva.

En la última década, los modelos de aprendizaje profundo, principalmente las llamadas redes neuronales, han demostrado una eficiencia en general superior en muchas tareas de la IA. De manera particular, han demostrado ser altamente efectivas para el análisis de sentimiento, tarea que en su forma más básica consiste en clasificar un texto según una polaridad de sentimiento positiva o negativa denotada por su autor hacia cierto tema. El acceso a grandes fuentes de datos reales y el surgimiento

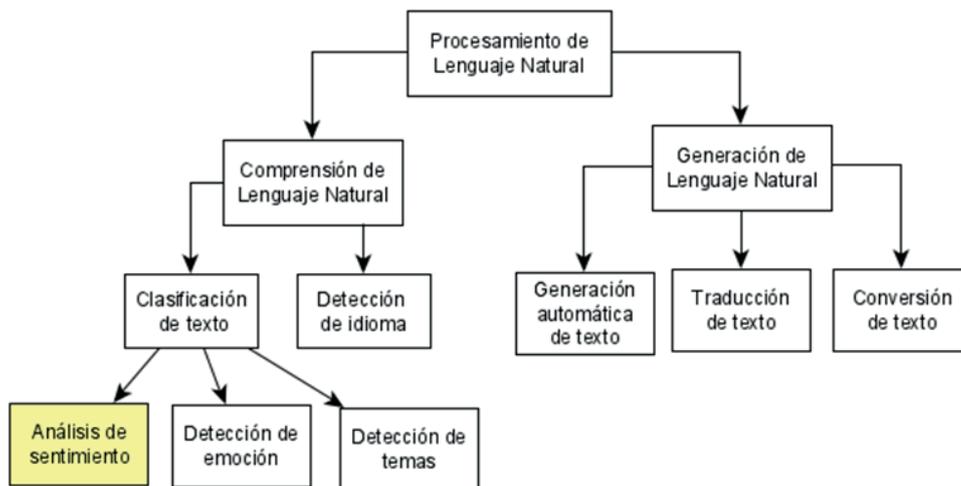
de comunidades y tecnologías para resolver problemas de IA han impulsado el desarrollo de varios modelos de redes neuronales para el análisis de sentimiento, siendo este enfoque actualmente el más usado para esta tarea.

1.1. ANÁLISIS DE SENTIMIENTO

El análisis de sentimiento es una de las varias tareas de clasificación dentro del Procesamiento del Lenguaje Natural (PLN), el dominio de la IA que estudia el lenguaje usado por los humanos para su comunicación diaria. Podemos definir el análisis de sentimiento como un proceso que “utiliza algoritmos y tecnologías de computación para sistemáticamente detectar, extraer y clasificar la información subjetiva y los estados afectivos expresados en un texto” (Lei y Liu, 2021, p. 1). El análisis de sentimiento aparece también como una aplicación de la “computación afectiva” (*affective computing*), un término que agrupa distintas técnicas para el reconocimiento de afecto en distintos tipos de datos, como una tarea de reconocimiento de grano grueso porque busca una clasificación binaria del dato ('positivo' o 'negativo'), y no se debe confundir con el *reconocimiento de emociones*, su equivalente de grano fino que busca clasificar los datos según un amplio conjunto de etiquetas que denotan las múltiples emociones del ser humano (Poria et al., 2018).

Por otro lado, dentro de la taxonomía del PLN, el análisis de sentimiento es una actividad que recae en la categoría de entendimiento de texto y, dentro de ella, en la clasificación de texto. (Arroni et al., 2023), tal como muestra la Figura 1.1. Adicionalmente

Figura 1.1
Mapa conceptual de la clasificación de las principales tareas del PLN.



Nota. Fuente: Elaboración propia en base a la revisión de la literatura realizada.

a los estados ‘positivo’ y ‘negativo’, se suele incluir una tercera categoría (‘neutral’) que denota una ausencia de sentimiento en el texto. Esta clasificación puede refinarse en escalas, que suelen ir desde la detección de un sentimiento “muy negativo” hasta uno “muy positivo”, usualmente en 5 o 7 niveles.

Como muchas tareas del PLN, el análisis de sentimiento ha cobrado importancia con la generación masiva de datos en las redes sociales y con el avance de las técnicas de procesamiento de *Big data* y de *Data mining*, además de los técnicas y algoritmos de análisis, primero con el *machine learning* (aprendizaje automático) y especialmente en los últimos años con el *deep learning* (aprendizaje profundo). El análisis de sentimiento tiene aplicaciones muy diversas que van desde la aceptación de los consumidores de un producto comercial, la medición de la opinión pública sobre campañas políticas, la percepción de servicios turísticos o la medición de indicadores de posibles problemas de salud mental (Denecke, 2023).

1.2. REDES SOCIALES Y CONTENIDO GENERADO POR USUARIOS

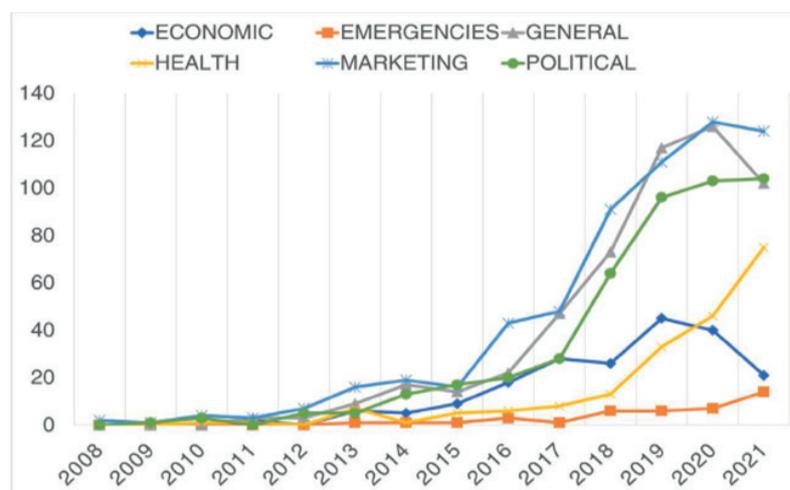
Las redes sociales son plataformas digitales que existen en Internet, en la cual una persona puede conectarse con otras hasta cierto grado (según las reglas de dicha plataforma) y donde puede generar información en distintos formatos. A finales de la década de los 2000, las redes sociales formaban parte de la vida cotidiana de gran parte de la población mundial. Las personas disponían ahora de

un medio para escribir comentarios y publicar fotos acerca de las actividades que quisieran, emitir opiniones y hacer pública la información. Este nuevo contenido creado por personas comunes y corrientes es conocido en la literatura como Contenido Generado por Usuarios (UGC) y abarca distintos formatos como imágenes, videos, textos cortos y documentos que las personas crean y hacen públicas en estas plataformas, pero incluye también el contenido generado en blogs, foros y marcadores sociales (Weitzel et al., 2016), plataformas típicas de la Web 1.0 que a día de hoy coexisten con las redes sociales y que siguen siendo fuentes de información válidas para el análisis de sentimiento.

Las redes sociales tienen actualmente un papel protagónico en el desarrollo del análisis de sentimiento de texto. Antes de su aparición, esta tarea, que tampoco contaba con potentes algoritmos de aprendizaje profundo, usaba datos disponibles en páginas de Internet. La popularización de las redes produjo un “boom” en el análisis de sentimiento debido a que una gran cantidad de usuarios podía compartir su opinión en comentarios cortos acerca de una enorme variedad de temas (Guimarães et al., 2018). La importancia de las redes se ve reflejada en el aumento de la cantidad de publicaciones de investigación sobre el análisis de sentimiento en estas plataformas desde el 2008, como muestra la Figura 1.2. Actualmente la minería de datos en redes sociales y el análisis de sentimiento siguen siendo temas de investigación populares y muy vigentes.

Figura 1.2.

Evolución de la cantidad de publicaciones sobre análisis de sentimiento en redes sociales.



Nota. Fuente: Rodríguez-Ibáñez et al. (2023). El eje vertical muestra la cantidad de publicaciones de investigación que aplican el análisis de sentimiento en redes sociales en un dominio específico.

1.3. PRE-PROCESAMIENTO DE TEXTO

La aplicación de técnicas y algoritmos al lenguaje natural requiere que el texto a procesar sea transformado antes de pasar como entrada a un modelo. Dado que los modelos de PLN, como cualquier otro modelo estadístico, solo puede recibir valores numéricos, el texto debe ser convertido antes a matrices numéricas, para lo cual se debe eliminar algunos aspectos en una fase conocida como pre-procesamiento de texto, realizada en toda tarea de PLN para evitar redundancia, eliminar el ruido y tratar solamente la información útil.

En el caso de los mensajes de redes sociales, hablamos de mensajes generalmente cortos pero que pueden presentar rasgos importantes que deben ser modificados antes de convertir el texto a matrices. Las técnicas de pre-procesamiento de texto típicas son:

A. Tokenización: división de un texto en palabras o frases que serán tratadas individualmente.

B. Stemming: convierte una palabra a la forma básica de su raíz semántica eliminando las últimas letras que forman sus sufijos.

C. Lematización: convierte una palabra en su lema. A diferencia del *stemming*, toma en cuenta el contexto y devuelve una palabra existente en el vocabulario y no la forma truncada de la raíz de la palabra.

D. Normalización: estandarización de mayúsculas, eliminación de signos de puntuación y caracteres especiales.

E. Remoción de ruido: se eliminan elementos irrelevantes para el análisis como enlaces, etiquetas HTML.

F. Remoción de stop-words: eliminación de palabras sin significado o carga semántica.

G. n-Gramas: es la tokenización de términos agrupados en series de 'n' palabras.

II. TAXONOMÍA DE LAS TÉCNICAS DE ANÁLISIS DE SENTIMIENTO

Existen dos grandes enfoques para los modelos de análisis de sentimiento tanto en mensajes como en documentos. En primer lugar, el enfoque basado en Diccionarios, donde se usa un Diccionario o Lexicón (*Lexicon*): un conjunto de palabras que denotan subjetividad positiva o negativa, etiquetadas con un valor numérico que representa esta subjetividad.

El segundo enfoque es el aprendizaje automático y reúne gran cantidad de técnicas y algoritmos usados en muchas tareas de IA en general. En este enfoque, se usa un conjunto de datos con mensajes ya etiquetados para entrenar a un modelo generalmente estadístico para que pueda clasificar nuevos datos. Dentro de este gran enfoque, se encuentran también las técnicas de aprendizaje profundo, las redes neuronales, que en los últimos años han demostrado ser generalmente las más eficientes para el análisis. Dada su complejidad y popularidad actual en el campo, se tratará a este grupo de manera separada del aprendizaje automático "tradicional".

2.1 ANÁLISIS DE SENTIMIENTO BASADO EN DICCIONARIOS

Antes de la introducción del concepto de aprendizaje automático, se usaban técnicas basadas en herramientas conocidas como "diccionarios de sentimiento", tanto para la identificación como la clasificación (Lei y Liu, 2021). Este enfoque consiste en usar un conjunto de palabras que expresan diversos sentimientos y que han sido etiquetadas manualmente por profesionales humanos. Para el análisis de sentimiento, las palabras de los diccionarios tienen asignadas una polaridad, generalmente de tres categorías: palabra positiva, palabra negativa o palabra neutral. Un diccionario de sentimiento puede tener como unidad léxica a una palabra, pero también a una frase o un concepto (Joshi et al., 2017). Actualmente, el enfoque más usado es el aprendizaje automático, que ha demostrado mejores resultados en la clasificación de sentimiento, pero los Diccionarios siguen siendo un método útil para trabajos de análisis exploratorio de datos que complementen a otras técnicas (Arroni et al., 2023) o como complemento de redes neuronales para formar modelos híbridos que mejoren el rendimiento general de la clasificación. Algunos de los diccionarios más importantes para el análisis de sentimiento son el *Valence Aware Dictionary and sEntiment Reasoner – VADER*, *SentiWordNet*, *SentiStrength*, *AFINN-111* y el diccionario de opiniones de Liu y Hu. (Al-Shaby, 2020).

2.2. ANÁLISIS DE SENTIMIENTO BASADO EN APRENDIZAJE AUTOMÁTICO TRADICIONAL

El enfoque del aprendizaje automático se diferencia de la programación clásica en que los programadores no crean las reglas para que el modelo devuelva una salida a partir de unos datos de entrada, sino que, junto con los datos, se ingresa la respuesta esperada de esos datos y el modelo procesa ambas entradas para generar las reglas necesarias que relacionen dichos datos con sus

respuestas (Chollet, 2019); es decir, tal como se muestra en la Figura 2.1, la salida de estos sistemas no son las respuestas esperadas de los datos de entrada, sino un conjunto específico de reglas que en la programación clásica el programador debía definir.

En este enfoque se empieza con un modelo matemático general con parámetros predefinidos que se van a ir modificando con el ingreso de datos para que el modelo se ajuste lo mejor posible a estos datos. A este proceso se conoce como "entrenamiento del modelo". Cuando este proceso termina, el modelo se habrá adaptado para cumplir una tarea específica y ya no será un modelo general (no podrá cumplir de manera efectiva una tarea diferente). Esa instanciación específica del modelo, ajustado según los datos particulares que hemos ingresado, también puede ser vista como "el algoritmo" que realiza esa tarea (Alpaydin, 2020)

Este enfoque ha demostrado ser capaz de resolver problemas complejos que la programación tradicional no ha conseguido o para los cuales no es eficiente, como quedó demostrado, por ejemplo, con el fracaso de los sistemas expertos en la década de 1980 (Rebala, 2019). Si bien las primeras tareas que impulsaron el aprendizaje automático estaban relacionadas con el reconocimiento de imágenes, hoy en día este enfoque encuentra aplicaciones importantes también en el PLN, siendo las redes neuronales las técnicas más populares, aunque siguen vigentes también otros modelos de aprendizaje "tradicional". Para el caso del análisis de sentimiento, podemos mencionar cinco técnicas dentro del Estado del Arte.

A. REGRESIÓN LOGÍSTICA

A pesar de su nombre, la regresión logística es una técnica estadística de clasificación binaria de datos que no es exclusiva del PLN ni del aprendizaje

automático. Es una técnica de aprendizaje supervisado y también se le conoce como modelo de "Entropía Máxima". Esta técnica determina si existe una relación entre una variable dependiente y una variable independiente. Para problemas de clasificación en IA, la variable dependiente es un dato específico que entra al modelo, y las variables independientes serán las categorías de posible pertenencia, para lo cual usa una función logística, generalmente función sigmoideal, que asigna al valor final del cálculo otro valor que determina la pertenencia del dato a una categoría.

La regresión logística es una de las dos técnicas de clasificación lineal, junto a las Máquinas de Vectores de Soporte, además de ser una de las más sencillas de aplicar y de las más antiguas, con orígenes que datan de hace más de cien años (Muller y Guido, 2017).

Podemos describir la forma simplificada de la ecuación de la siguiente forma: sean 0 y 1 los valores que representa dos categorías, la probabilidad de que una variable dependiente Y sea igual a 1 (es decir, que pertenezca a la categoría que representa el valor 1), dada una variable dependiente X, es la siguiente:

$$P(Y = 1 | X) = \frac{1}{1 + e^{-(w_0+w_1x_1+w_2x_2+\dots+w_px_p)}}$$

Donde:

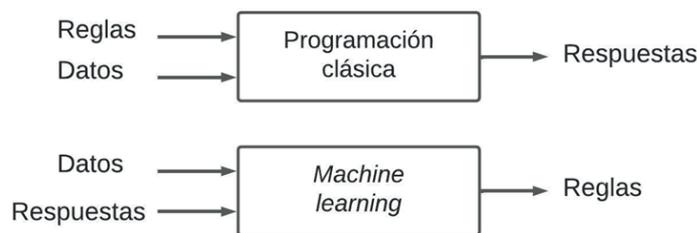
$x_i; 1 \leq i \leq p$: Característica de X

$w_i; 1 \leq i \leq p$: Parámetros del modelo que serán aprendidos en el entrenamiento

B. NAIVE BAYES

El algoritmo de Naive-Bayes es uno de los más conocidos de entre los llamados "modelos probabilísticos". Es una aplicación directa del teorema de

Figura 2.1. Esquema de programación clásica vs. machine learning



Nota. Elaboración: Francois Chollet (2019).

Bayes, que data del siglo XVIII. Es relativamente sencillo de aprender pues, en su forma básica, solo requiere conocimiento del teorema y otros principios básicos de Estadística. El nombre del algoritmo se debe a que se parte de la suposición de que los datos de entrada no son dependientes entre ellos, lo que se asume como una suposición “ingenua” o “naive” en inglés (Chollet, p. 35).

La base de este método es el teorema de Bayes, representado de la forma:

$$P(c|x) = P(x|c) \cdot P(c) / P(x)$$

Donde:

$P(c|x)$: probabilidad de c , dada la observación x

$P(x|c)$: probabilidad de x , dada la ocurrencia de c

$P(c), P(x)$: probabilidad de c y x , respectivamente

En este ejemplo genérico, podemos considerar a $P(c|x)$ como la probabilidad condicional de que el mensaje a evaluarse pertenezca a la categoría $c = 'pos'$ (mensajes positivos); $P(x|c)$ es la probabilidad condicional de que el texto x se genere dado que el mensaje pertenece a la categoría $c = 'pos'$; $P(c = 'pos')$ es la probabilidad a priori de que el mensaje pertenezca a la categoría $c = 'pos'$ y $P(x)$ la probabilidad a priori del texto x . $P('pos')$ puede calcularse como la proporción de mensajes positivos en el conjunto de entrenamiento; $P(x|'pos')$ se calcula asumiendo la independencia de los componentes del texto (las palabras, o los n -gramas) y calculando el producto de las probabilidades de cada palabra en el texto:

$$P(x|pos) = P(w_1|pos) * P(w_2|pos) * \dots * P(w_n|pos)$$

Donde:

w_i : i -ésima palabra del mensaje x

Posteriormente, se calculará la probabilidad de $P(x|neg)$, es decir, la probabilidad condicional de que el mensaje x pertenezca a la clase de mensajes negativos. El método de Naive-Bayes clasificará al mensaje según el porcentaje de pertenencia más alto obtenido.

C. MÁQUINAS DE VECTORES DE SOPORTE

Abreviadas como SVM, este algoritmo sigue siendo uno de los más populares en el campo del aprendizaje automático para diversas tareas de IA. Los orígenes del algoritmo se remontan a 1963, pero la forma actual de la SVM se formuló en los Laboratorios Bell en la década de 1990 (Chollet, 2019, p. 36).

Las SVMs son clasificadores lineales binarios que se encuentran entre los enfoques paramétrico y no paramétrico de clasificación, ya que, en el primer caso, usan parámetros para modelar, pero también usan datos de entrenamiento como en el segundo grupo. Por ello, una de las ventajas que hizo de esta técnica bastante popular y exitosa antes del avance de las redes neuronales es su capacidad de trabajar con datos escasos y de entrenarse adecuadamente con conjuntos de datos relativamente pequeños (Rebala et al., 2019).

Una SVM crea un hiperplano de $N-1$ dimensiones para vectores de características de dimensión n . En su caso más simple, cuando $N=2$, el hiperplano es una línea recta que divide el espacio en dos clases diferentes. Este clasificador lineal puede ser representado de la forma:

$$y = w \cdot f(x) + b$$

Donde:

$f(x)$: vector de características

w : peso asignado al vector de características

b : sesgo

En nuestro ejemplo de dos dimensiones, los valores y que quedan por encima de la recta son clasificados en una primera categoría (por ejemplo: “mensajes positivos”), y los que quedan por debajo se clasifican en la segunda (por ejemplo: “mensajes negativos”). Para casos en los que los datos no pueden separarse de manera lineal, se puede usar un *kernel*, una función matemática de mapeo que sustituye puntos de datos por vectores de características de dimensión más alta que sean linealmente separables, y que además puede descomponerse en un producto punto (Rebala, p. 65). La elección de la función *kernel* adecuada dependerá del problema a resolver y del tipo de datos disponibles. Esta función puede ser representada como K tal que:

$$K(x, y) = \langle f(x), f(y) \rangle$$

Donde:

x, y : entradas n -dimensionales

$f(x), f(y)$: funciones de mapeo de espacio dimensional n a espacio m

$\langle a, b \rangle$: producto punto de los vectores a y b

Debido a esto, las SVM es también clasificado por algunos autores dentro del grupo de las máquinas

de *kernel*, que son métodos de margen máximo que permiten ver al modelo como “la suma de las influencias de un subconjunto de entrenamiento” (Alpaydin, 2020, p. 514).

D. BOSQUES ALEATORIOS

Los Bosques Aleatorios (BAs) son una técnica que forman parte de los llamados métodos de *bagging*, en los que a partir de un conjunto de entrenamiento D , se realizan M muestreos con reemplazo uniforme para obtener M nuevos subconjuntos, cada uno de los cuales se usará de manera independiente para entrenar M modelos diferentes, o M instancias distintas de un mismo modelo, para finalmente combinar los resultados obtenidos en un único modelo según alguna técnica de promedio o elección (Jiang, 2022). Estos métodos a su vez forman parte de los métodos de *ensemble learning* o aprendizaje conjunto.

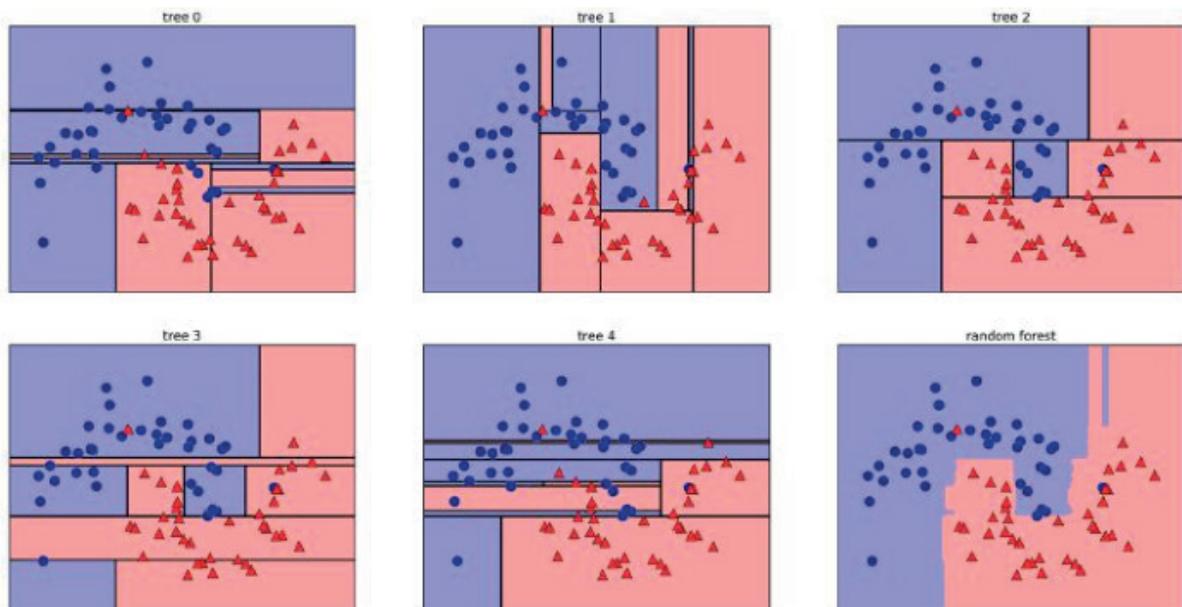
En el caso de los BA, el modelo base usado son árboles de decisión (AD). Se aplican sucesivamente ADs con pesos aleatorios. Los ADs son árboles binarios en el que cada nodo representa una decisión a tomar según el valor de una característica única que entra al árbol en el nodo raíz. La técnica del BA consiste en generar una gran cantidad de ADs para

características específicas de los datos y después ponderar los parámetros de los árboles para obtener un único clasificador eficiente y general, como se puede apreciar en la Figura 2.2. Este algoritmo demostró ser bastante eficiente para varios problemas de aprendizaje automático y casi siempre fue el segundo mejor para cada problema de este tipo antes del *boom* de las redes neuronales (Chollet, p. 38), además de ser un método bastante intuitivo ya que nos permite ver con claridad qué características contribuyen a la clasificación y a determinar la importancia de las variables que influyen en la decisión final. (Rebala, 2019). El éxito de este clasificador, como de cualquier método de *bagging*, depende de la diversidad de sus modelos base, sus árboles de decisión.

Una ventaja de este método es la mitigación del sobreajuste al usar modelos base muy variados. Sin embargo, incluso con esta característica, la tendencia al sobreajuste vuelve a aparecer si las categorías de clasificación son demasiadas. Por otro lado, la independencia de los modelos base para el entrenamiento hace posible usar el procesamiento paralelo para la construcción del modelo general, pero se debe tener en cuenta que esa velocidad puede reducirse cuando la profundidad de los árboles es muy alta.

Figura 2.2.

Representación gráfica de las fronteras de decisión de cinco árboles de decisión con pesos aleatorios y la frontera media del Bosque Aleatorio resultante



Nota. Autor: Müller y Guido (2017).

E. AUMENTO DEL GRADIENTE

Es otra técnica de aprendizaje conjunto, también basado en ADs. A diferencia de los BAs, el aumento del gradiente es una técnica de *boosting* (de ahí su nombre, *gradient boosting machine*), las cuales consisten en generar modelos base de aprendizaje de manera secuencial, cada uno aprendiendo de los errores cometidos por el modelo anterior. Es decir, a diferencia de los BA que generan árboles con parámetros aleatorios de forma simultánea e independiente, el aumento del gradiente trabaja generando árboles consecutivos, donde los pesos de un árbol están influenciados por los resultados del árbol anterior y sus parámetros (Alpaydin, 2020). La diferencia entre las predicciones obtenidas y el valor real de los datos será el error o “residuo”. Según Pirayonesi et al. (2020), esta técnica suele ser más eficiente que los Bosques Aleatorios.

La cuestión esencial de esta técnica es escoger un modelo base en cada paso que nos asegure que el modelo conjunto está mejorando su rendimiento. Podemos describir esto como la búsqueda de un modelo $F(x)$ que minimice la función de pérdida $L(f(x), y)$, donde x es un vector de características e y es la variable de respuesta. Entonces, el gradiente de L es:

$$\frac{\partial L(f(x), y)}{\partial f}$$

Este gradiente es una nueva función en el espacio de funciones que apunta en dirección del mayor incremento de $L(f(x), y)$ en f . Entonces, el aumento del gradiente busca estimar el nuevo modelo base en dirección al negativo del gradiente del actual modelo conjunto $F_{m-1}(x)$, representado por:

$$-\nabla L(F_{m-1}(x)) \triangleq -\left. \frac{\partial L(f(x), y)}{\partial f} \right|_{f = F_{m-1}}$$

Pero este gradiente puede salirse del espacio H donde vive F , por lo que se debe buscar una función dentro del espacio H que se aproxime más al gradiente especificado (Jiang, 2022).

2.3 ANÁLISIS DE SENTIMIENTO BASADO EN APRENDIZAJE PROFUNDO

A. REDES NEURONALES CONVOLUCIONALES

Las redes neuronales convolucionales (CNN) surgieron para el problema de la clasificación de imágenes. Estas redes se componen de varias capas

que capturan los distintos niveles de características o patrones presentes en una imagen, siendo que cada capa captura patrones más complejos que la anterior. Debido a la necesidad de usar varios niveles de capas, fueron el primer modelo del llamado *aprendizaje profundo* (Goyal et al., 2018). A partir del éxito de las CNN en el procesamiento y clasificación de imágenes, se empezó a usar este modelo para tareas de PLN. Para esto, el texto se convierte a una representación numérica, generalmente vectores, y estos pasan a una capa de convolución donde distintos *kernels* extraen las características de los vectores. Luego, las capas ocultas extraen distintos niveles de características textuales que serán reducidos en una capa de *pooling* que reduce el tamaño vectorial de las características y usa un algoritmo de actualización para mejorar los pesos de las capas.

B. REDES NEURONALES RECURRENTE SIMPLES

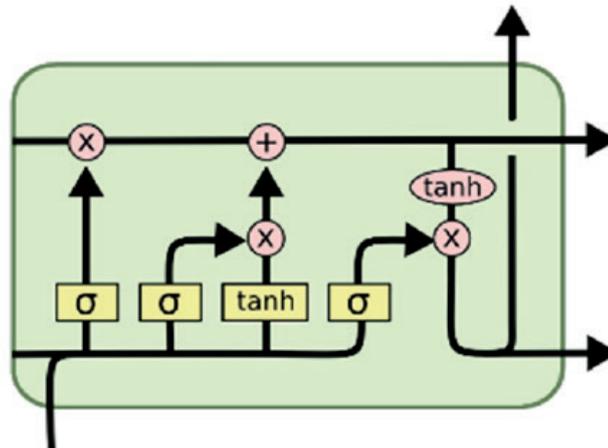
Este tipo de redes buscan subsanar la deficiencia de las CNN de no poder extraer características de grupos de palabras que se encuentran muy separadas en el texto. Este modelo es usado cuando los patrones de datos cambian con el paso del tiempo (Goyal et al., p. 47). En el caso de un texto, este es considerado una secuencia de palabras en el que la posición de cada una es una característica importante. Las RNN tienen bucles de retroalimentación, a diferencia de las redes *feed-forward*, en donde los datos son procesados capa por capa de manera secuencial y únicamente hacia adelante, como en las CNN. En estas redes, cada neurona es considerada una “celda de memoria”.

C. REDES DE MEMORIA DE CORTO Y LARGO PLAZO

Son una versión modificada de las RNN convencionales que usan un mecanismo para “almacenar” información de las primeras palabras del texto, que se suelen perder en otros modelos al ingresar sucesivamente nuevos términos. Esta pérdida se conoce formalmente como pérdida del gradiente y sucede cuando un modelo actualiza las características de una secuencia de datos y las características de las primeras entradas se van perdiendo. Para solucionar este problema, las redes *Long-Short Term Memory* (LSTM) tienen en sus celdas de memoria compuertas que le permiten almacenar o desechar información, tal como muestra la Figura 2.3. Estas compuertas se activan con diversas funciones, principalmente sigmooidal o tangencial hiperbólica, y típicamente son:

Figura 2.3.

Estructura interna de una celda de memoria de una LSTM



Nota. Los símbolos " σ " y " \tanh " representan las funciones sigmoide y tangente hiperbólica. Los símbolos "+" y "x" representan operaciones de suma y multiplicación vectorial. Tomado de "A Sentiment Analysis Method of Capsule Network Based on BiLSTM" (p. 37017), por Y. Dong et al., 2019, *IEEE Access*, 8.

- Compuerta de olvido: determina la información que la celda debe olvidar. Se activa con una función sigmoide.
- Compuerta de entrada: determina la nueva información que ingresa a la celda y que se va a almacenar. Usa funciones sigmoide y tangente hiperbólica.
- Compuerta de salida: controla qué información saldrá de una celda hacia la siguiente celda. Usa funciones sigmoide y tangente hiperbólica.

D. TRANSFORMADORES

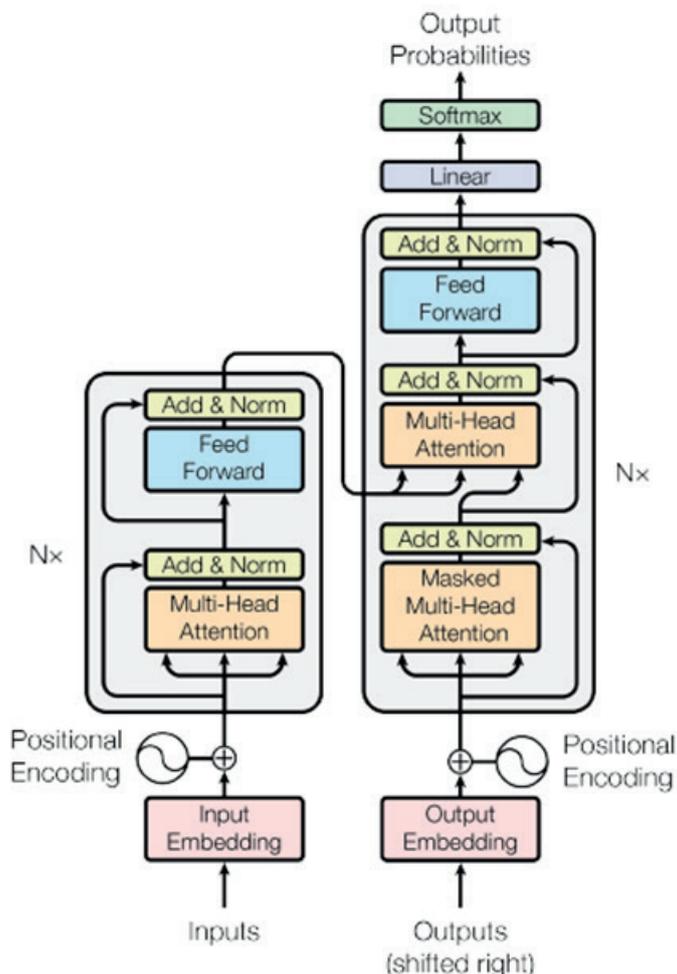
Modelo de red neuronal basado únicamente en mecanismos de atención. Tiene la ventaja de procesar en paralelo las distintas palabras o términos de un texto, a diferencia del procesamiento secuencial de las RNN. Gracias a este diseño, ha sido posible entrenar a estos modelos con gigantescos volúmenes de datos en tiempos razonables. Son la arquitectura usada por los grandes modelos de lenguaje actuales (*Large Language Models*). La Figura 2.4 muestra la arquitectura básica de este modelo tal como fue propuesto en el 2017.

Desde la publicación del trabajo de Vaswani y sus colaboradores en Google, los modelos basados en Transformadores han superado a las redes neuronales predecesoras en casi todas las tareas de PLN (Tunstall et al., 2022). Para entender los

Transformadores, es necesario revisar tres conceptos clave:

- A. Marco codificador – decodificador: Diseño en la arquitectura del Transformador en la que un bloque codificador recibe la entrada y la transforma en un vector numérico, llamado *último estado oculto*, que representa un concepto abstracto, el cual será transformado a un texto de salida en el bloque decodificador. El cuello de botella generado por el codificador al representar el significado de toda la secuencia de entrada (el texto completo) es resuelto con el mecanismo de atención.
- B. Mecanismo de atención: Asigna pesos a las diferentes partes de la secuencia de entrada según su importancia en la tarea a realizar. Para subsanar el tiempo de procesamiento que implica el carácter secuencial de las RNN, se propuso el mecanismo de "auto-atención", para que la atención pueda usarse en todas las capas de la red, permitiendo un rápido entrenamiento de muchos más datos.
- C. Aprendizaje por transferencia: Este enfoque de aprendizaje automático consiste en usar un modelo ya entrenado y validado con muchos datos para resolver una tarea similar. Una vez que se dispone del modelo, se hace un ajuste fino (*fine-tuning*) para adaptarlo a nuestra tarea.

Figura 2.4.
 Arquitectura original del modelo de Transformador propuesto por Vaswani et al.



Nota. Tomado de "Attention is all you need" (p. 3), por A. Vaswani et al., 2017, 31st Conference on Neural Information Processing Systems (NIPS 2017).

Según Tunstall et al. (2022), estas tres características esenciales de los Transformadores son los que han facilitado que se produjera una revolución dentro del aprendizaje automático y, por tanto, de la IA. Actualmente existen Grandes Modelos de Lenguaje (*Large Language Models*, LLMs), que son Transformadores entrenados con millones de datos para realizar tareas complejas a una velocidad instantánea, siendo los más célebres el Generative Pre-trained Transformer (GPT) de OpenAI, y el Bidirectional Encoder Representations from Transformers (BERT) de Google. Ambos modelos, como muchos otros disponibles actualmente, pueden ajustarse para realizar tareas de traducción de texto, detección de temas, generación de respuestas y análisis de sentimiento. BERT se encuentra disponible para ajustarse en la plataforma HuggingFace, un repositorio abierto de

modelos de IA, que ha acelerado el desarrollo y la construcción de modelos y aplicaciones basados en PLN, facilitando el trabajo para muchos desarrolladores (Jain, 2022).

III. TÉCNICAS HÍBRIDAS

Adicionalmente a las técnicas básicas de análisis de sentimiento, investigadores han propuesto modelos que combinan dos o más de estas técnicas, dando como resultado modelos híbridos que subsanan o atenúan dificultades propias del problema específico del caso, o que delegan a una técnica alguna tarea previa como la representación vectorial o la ingeniería de características. A continuación, se muestran las más comunes según la revisión de la literatura realizada:

3.1. Diccionarios y Aprendizaje Automático

Chiny et al. (2021) proponen un modelo que usa el diccionario VADER con una LSTM y algunos algoritmos de aprendizaje automático tradicional para análisis de mensajes cortos generales con el fin de reducir el costo de implementación sin reducir la exactitud. La propuesta de Mohamed et al. (2023) llamada LexDeep usa también el diccionario VADER y evalúa el rendimiento de los modelos CNN, LSTM, CNN+LSTM y LSTM+CNN para analizar el sentimiento de mensajes en Twitter sobre discusiones relacionadas al desempleo en el contexto de la Covid-19. Por otro lado, Mohamad Sham y Mohamed (2022) usaron siete diccionarios distintos (SentiWordNet, TextBlob, VADER, SentiStrength, Hu & Liu, MPQA y WKWSCI) y tres clasificadores de aprendizaje automático (Naive-Bayes, SVM y Regresión Logística) para aplicar el análisis de sentimiento a mensajes relacionados con el cambio climático, un dominio poco explorado en este campo del PLN, según los autores.

3.2. CNN-RNN y CNN-LSTM

Dong et al. (2020) propusieron un método de análisis basado en una arquitectura de cápsula que combina una capa convolucional para la extracción de n-gramas y un módulo LSTM bidireccional (Bi-LSTM) para la clasificación general de mensajes cortos. Por su parte, el modelo CNN-LSTM de Rehman et al. (2019) usa tres capas convolucionales para extraer las características de texto y un módulo LSTM básico para la clasificación, aplicada a reseñas de películas. Behera et al. (2019) presentan un modelo Co-LSTM de clasificación libre de dominio para mensajes en redes sociales combinando CNN y LSTM, buscando flexibilidad y escalabilidad en la clasificación. Otras propuestas resaltantes son el modelo basado en búsqueda en rejilla de Priyadarshini y Cotton (2021), que usa además el algoritmo de k-vecinos más cercanos; el modelo para análisis de sentimiento de reseñas de productos en redes sociales de Bhuvaneshwari et al. (2022) que usa CNN, Bi-LSTM y mecanismo de auto atención; y el modelo híbrido estándar CNN-LSTM para evaluar reseñas de restaurantes de Hossain et al. (2020).

3.3. BERT-LSTM y RoBERTa-LSTM

Cai et al. (2020) usan métodos estadísticos para clasificar mensajes de inversores en el mercado energético en redes sociales chinas y un modelo que integra BERT y BiLSTM para detectar el sentimiento de estos mensajes, además de predecir tendencias en este mercado. Por su parte, Tan et

al. (2022) proponen un modelo que combina las ventajas de la arquitectura de Transformador y mitiga las limitaciones del modelo secuencial combinando el modelo pre-entrenado RoBERTa para el mapeo de palabras a un espacio de *word embeddings*, y una red LSTM para capturar la semántica contextual entre palabras distanciadas, aplicado a mensajes cortos y generales. También destaca la propuesta Agrawal et al. (2021), quienes afinan el modelo BERT con una red Bi-LSTM, incorporando una capa de tarea específica para analizar el sentimiento de mensajes cortos generales.

IV. CONCLUSIONES

El presente trabajo ha hecho una revisión de las principales técnicas vigentes usadas para el análisis de sentimiento en mensajes cortos de redes sociales. Las técnicas son: a) Diccionarios, que engloba distintos conjuntos de palabras etiquetados por expertos; dentro de las técnicas de aprendizaje automático encontramos: b) la Regresión Logística, c) Naive-Bayes, d) Máquinas de Vectores de Soporte, e) Bosques Aleatorios y f) Aumento del Gradiente; las técnicas de aprendizaje profundo: g) las Redes Neuronales Convolucionales, h) las Redes Neuronales Recurrentes básicas, i) las Redes de Memoria de Corto y Largo Plazo, una variación mejorada de la anterior; y finalmente, j) los Transformadores, una técnica de mención especial por haber dado paso a los Grandes Modelos de Lenguaje que realizan tareas textuales muy sofisticadas en un tiempo muy corto. Adicionalmente, se revisaron las principales técnicas híbridas, que combinan por lo menos dos de las técnicas mencionadas.

Por el alcance de este trabajo, solo se ha presentado estas técnicas de manera general, con versiones simplificadas de sus ecuaciones. Se debe mencionar que las bases teóricas de muchas de ellas requieren ser explicado de manera más extensa. Finalmente, podemos mencionar que, si bien los Transformadores han demostrado una mejor eficiencia para el análisis de sentimiento, la diversidad de dominios, plataformas, vocabularios, cantidad de mensajes, el tamaño de los textos y otros factores hacen que en la práctica sea suficiente, y a veces mejor, la elección de otras técnicas.

REFERENCIAS BIBLIOGRÁFICAS

- [1] Al-Shabi, M. (2020). Evaluating the performance of the most important Lexicons used to Sentiment analysis and opinions Mining. *IJCSNS*, 20(1), 1.

- [2] Agrawal, S., Dutta, S. y Patra, B.K. (2021). "Sentiment Analysis of Short Informal Text by Tuning BERT - Bi-LSTM Model," IEEE EUROCON 2021 - 19th International Conference on Smart Technologies, Lviv, Ukraine, 2021, pp. 98-102, doi: 10.1109/EUROCON52738.2021.9535535.
- [3] Alpaydin, E. (2020). *Introduction to machine learning*. MIT press.
- [4] Arroni, S., Galán, Y., Guzmán-Guzmán, X., Núñez-Valdez, E. R., y Gómez, A. (2023). Sentiment Analysis and Classification of Hotel Opinions in Twitter With the Transformer Architecture.
- [5] Behera, R. K., Jena, M., Rath, S. K., & Misra, S. (2021). Co-LSTM: Convolutional LSTM model for sentiment analysis in social big data. *Information Processing & Management*, 58(1), 102435.
- [6] Cai, R., Qin, B., Chen, Y., Zhang, L., Yang, R., Chen, S., y Wang, W. (2020). Sentiment analysis about investors and consumers in energy market based on BERT-BiLSTM. *IEEE access*, 8, 171408-171415.
- [7] Cambria, E., Das, D., Bandyopadhyay, S., y Feraco, A. (2017). Affective computing and sentiment analysis. *A practical guide to sentiment analysis*, 1-10.
- [8] Chiny, M., Chihab, M., Bencharef, O., & Chihab, Y. (2021). LSTM, VADER and TF-IDF based hybrid sentiment analysis model. *International Journal of Advanced Computer Science and Applications*, 12(7).
- [9] Chollet, F. (2021). *Deep learning with Python*. Simon and Schuster.
- [10] Denecke, K. (2023). *Sentiment Analysis in the Medical Domain*. Springer Nature.
- [11] Dong, Y., Fu, Y., Wang, L., Chen, Y., Dong, Y., & Li, J. (2020). A sentiment analysis method of capsule network based on BiLSTM. *IEEE Access*, 8, 37014–37020. <https://doi.org/10.1109/ACCESS.2020.2973711>
- [12] Guimarães, N., Torgo, L., y Figueira, Á. (2018). Twitter as a source for time-and domain-dependent sentiment lexicons. *Social Network Based Big Data Analysis and Applications*, 1-19.
- [13] Goyal, P., Pandey, S., & Jain, K. (2018). Deep learning for natural language processing. *New York: Apress*.
- [14] Hossain, N., Bhuiyan, M. R., Tumpa, Z. N., & Hossain, S. A. (2020, July). Sentiment analysis of restaurant reviews using combined CNN-LSTM. In *2020 11th International conference on computing, communication and networking technologies (ICCCNT)* (pp. 1-5). IEEE.
- [15] Jain, S. M. (2022). Introduction to Transformers. In *Introduction to Transformers for NLP: With the Hugging Face Library and Models to Solve Problems*. Berkeley, CA: Apress.
- [16] Joshi, A., Bhattacharyya, P., & Ahire, S. (2017). Sentiment resources: Lexicons and datasets. *A Practical Guide to Sentiment Analysis*, 85-106.
- [17] Lei, L. y Liu, D. (2021). *Conducting Sentiment Analysis*. Cambridge University Press.
- [18] Mohamad Sham, N., & Mohamed, A. (2022). Climate change sentiment analysis using lexicon, machine learning and hybrid approaches. *Sustainability*, 14(8), 4723.
- [19] Mohamed, A., Zain, Z. M., Shaiba, H., Alturki, N., Aldehim, G., Sakri, S., ... & Zain, J. M. (2023). Lexdeep: Hybrid lexicon and deep learning sentiment analysis using Twitter for unemployment-related discussions during covid-19. *Computers, Materials & Continua*, 75(1), 1577-1601.
- [20] Müller, A. C., y Guido, S. (2017). *Introduction to machine learning with Python*. O'Reilly.
- [21] Piryonesi, S.M. y El-Diraby, Tamer E. (2020). Data Analytics in Asset Management: Cost-Effective Prediction of the Pavement Condition Index. *Journal of Infrastructure Systems*. 26 (1): 04019036. doi:10.1061/(ASCE)IS.1943-555X.0000512
- [22] Poria, S., Hussain, A., y Cambria, E. (2018). *Multimodal sentiment analysis*. Cham: Springer International Publishing.
- [23] Priyadarshini, I., & Cotton, C. (2021). A novel LSTM-CNN-grid search-based deep neural network for sentiment analysis. *The Journal of Supercomputing*, 77(12), 13911-13932.
- [24] Rebala, G., Ravi, A., y Churiwala, S. (2019). *An introduction to machine learning*. Springer.
- [25] Rehman, A. U., Malik, A. K., Raza, B., & Ali, W. (2019). A hybrid CNN-LSTM model for improving accuracy of movie reviews sentiment analysis. *Multimedia Tools and Applications*, 78, 26597-26613.

- [26] Rodríguez-Ibáñez, M., Casáñez-Ventura, A., Castejón-Mateos, F., y Cuenca-Jiménez, P.-M. (2023). A review on sentiment analysis from social media platforms. *Expert Systems with Applications*, 223(August 2022), 119862. <https://doi.org/10.1016/j.eswa.2023.119862>
- [27] Tan, K. L., Lee, C. P., Anbananthen, K. S. M., y Lim, K. M. (2022). RoBERTa-LSTM: a hybrid model for sentiment analysis with transformer and recurrent neural network. *IEEE Access*, 10, 21517-21525.
- [28] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... y Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- [29] Weitzel, L., Prati, R. C., y Aguiar, R. F. (2016). The comprehension of figurative language: What is the influence of irony and sarcasm on NLP techniques?. *Sentiment Analysis and Ontology Engineering: An Environment of Computational Intelligence*, 49-74.

Financiamiento

Propia.

Conflictos de interés

Los autores declaran no tener conflictos de interés.

Contribuciones de autoría

Todos los autores participaron en las diferentes actividades para la elaboración del artículo.