



# ANÁLISIS DE SENTIMIENTO

NLP

# ¿QUÉ ES EL ANÁLISIS DE SENTIMIENTO?

- El análisis del sentimiento es una técnica utilizada para determinar el tono emocional o sentimiento expresado en un texto. Consiste en analizar las palabras y frases utilizadas en el texto para identificar el sentimiento subyacente, ya sea positivo, negativo o neutro.



## ¿PARA QUÉ SE OCUPA?

El análisis del sentimiento tiene una amplia gama de aplicaciones, como la monitorización de las redes sociales, el análisis de las opiniones de los clientes y la investigación de mercado.



## DESAFÍOS Y RETOS

Uno de los principales retos del análisis del sentimiento es la complejidad inherente al lenguaje humano. Los datos textuales suelen contener sarcasmo, ironía y otras formas de lenguaje figurado que pueden ser difíciles de interpretar con los métodos tradicionales.



## ¿CÓMO FUNCIONA?

El análisis de sentimiento utiliza tecnologías de procesamiento de lenguaje natural (PLN) y aprendizaje automático (ML) para entrenar el software informático para analizar e interpretar el texto de manera similar a los humanos. El software utiliza uno de los dos enfoques, basado en reglas o ML (o una combinación de ambos, conocida como híbrido). Cada enfoque tiene sus puntos fuertes y débiles; mientras que un enfoque basado en reglas puede ofrecer resultados casi en tiempo real, los enfoques basados en ML son más adaptables y normalmente pueden manejar escenarios más complejos.



## ANÁLISIS DE SENTIMIENTO BASADO EN REGLAS

En el enfoque basado en reglas, el software está entrenado para clasificar ciertas palabras clave en un bloque de texto basado en grupos de palabras, o léxicos, que describen la intención del autor. Por ejemplo, las palabras en un léxico positivo pueden incluir "asequible", "rápido" y "bien hecho", mientras que las palabras en un léxico negativo podrían mostrar "caro", "lento" y "mal hecho". A continuación, el software busca en el clasificador las palabras del léxico positivo o negativo y calcula una puntuación total del sentimiento basada en el volumen de palabras utilizadas y la puntuación del sentimiento de cada categoría.

## ANÁLISIS DE SENTIMIENTO MEDIANTE APRENDIZAJE AUTOMÁTICO (1/2)

Con un enfoque de aprendizaje automático (ML), se utiliza un algoritmo para entrenar al software para medir el sentimiento en un bloque de texto utilizando palabras que aparecen en el texto, así como el orden en que aparecen. Los desarrolladores utilizan algoritmos de análisis de sentimiento para enseñar al software a identificar las emociones en el texto de forma similar a como lo hacen los humanos. Los modelos de ML siguen "aprendiendo" de los datos que se les suministran, de ahí el nombre de "machine learning" (aprendizaje automático).

# ANÁLISIS DE SENTIMIENTO MEDIANTE APRENDIZAJE AUTOMÁTICO (2/2)

Estos son algunos de los algoritmos de clasificación más utilizados:

Regresión lineal: un algoritmo de estadísticas que describe un valor (Y) basado en un conjunto de características (X).

Naive Bayes: Algoritmo que utiliza el teorema de Bayes para categorizar palabras en un bloque de texto.

Máquinas de soporte vectorial (SVM): algoritmo de clasificación rápido y eficiente utilizado para resolver problemas de clasificación de dos grupos.

Aprendizaje profundo (DL): también conocido como red neuronal artificial, el aprendizaje profundo es una técnica avanzada de machine learning que vincula varios algoritmos para imitar la función cerebral humana.

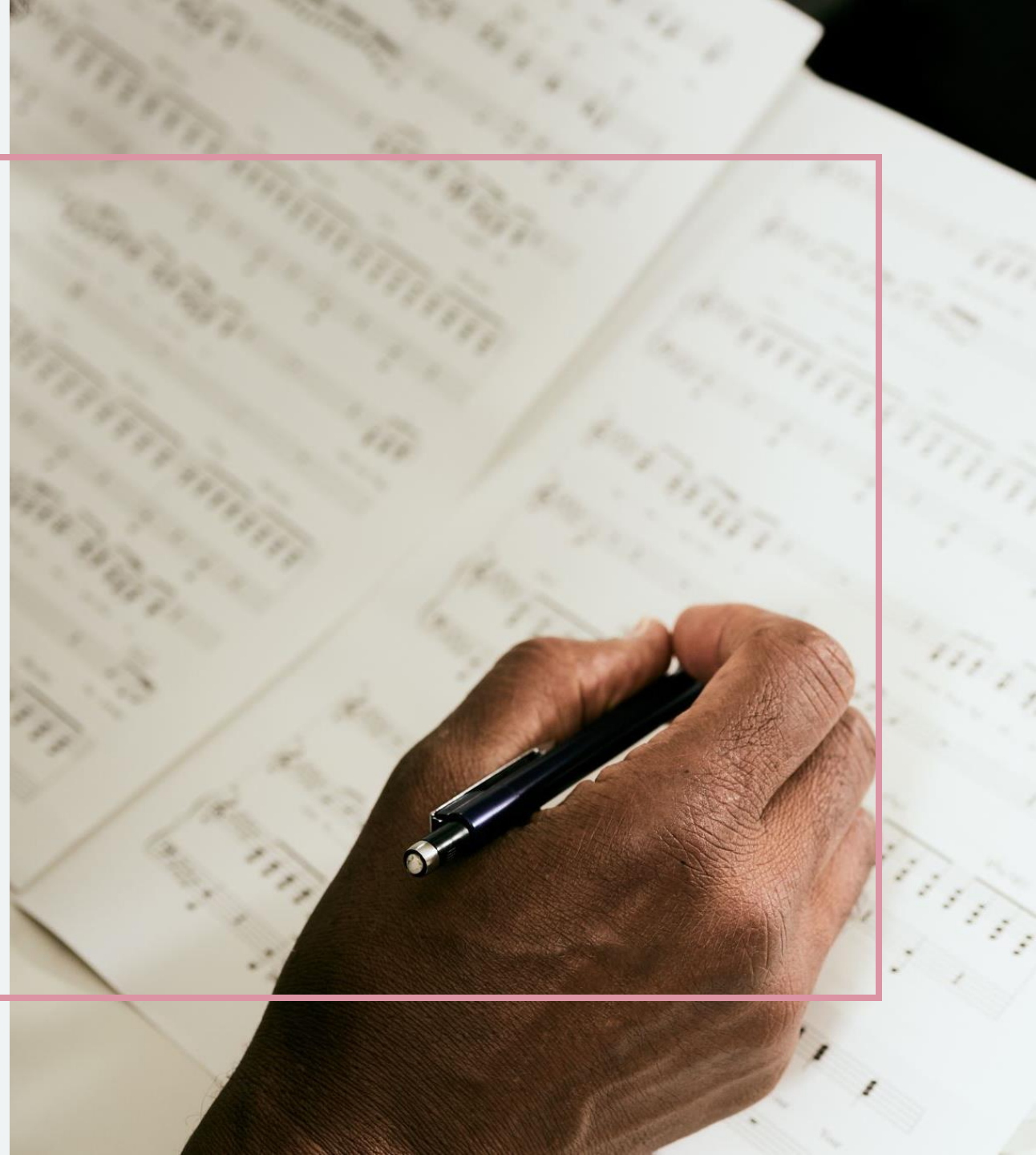


## ANÁLISIS DE SENTIMIENTO MEDIANTE ENFOQUE HÍBRIDO

Un enfoque híbrido del análisis de texto combina capacidades basadas en reglas y ML para optimizar la precisión y la velocidad. Si bien es muy preciso, este enfoque requiere más recursos, como tiempo y capacidad técnica, que los otros dos.

## ANALIZADOR DE SENTIMIENTO NLTK (1/3)

El SentimentIntensityAnalyzer de NLTK, específicamente el modelo VADER (Valence Aware Dictionary and sEntiment Reasoner), está diseñado para realizar análisis de sentimientos a nivel de oración. Este modelo es capaz de capturar matices en el texto, como la intensidad de las emociones y el manejo de negaciones y énfasis (por ejemplo, palabras en mayúsculas o con signos de exclamación).



# ANALIZADOR DE SENTIMIENTO NLTK (2/3)

El modelo devuelve un diccionario con cuatro puntuaciones clave:

Positivo (pos): La proporción de palabras en el texto que son positivas.

Negativo (neg): La proporción de palabras en el texto que son negativas.

Neutral (neu): La proporción de palabras en el texto que son neutrales.

Compuesto (compound): Una puntuación agregada que varía entre -1 (muy negativo) y 1 (muy positivo). Esta es la puntuación más importante y se calcula como una suma normalizada de las puntuaciones de sentimiento.

# ANALIZADOR DE SENTIMIENTO NLTK (3/3)

`polarity_scores(text)`

[\[source\]](#)

Return a float for sentiment strength based on the input text. Positive values are positive valence, negative value are negative valence.

## Note

Hashtags are not taken into consideration (e.g. #BAD is neutral). If you are interested in processing the text in the hashtags too, then we recommend preprocessing your data to remove the #, after which the hashtag text may be matched as if it was a normal word in the sentence.



# REFERENCIAS

<https://www.ibm.com/mx-es/topics/sentiment-analysis>

