

# Inteligencia Artificial para el análisis de sentimientos en textos

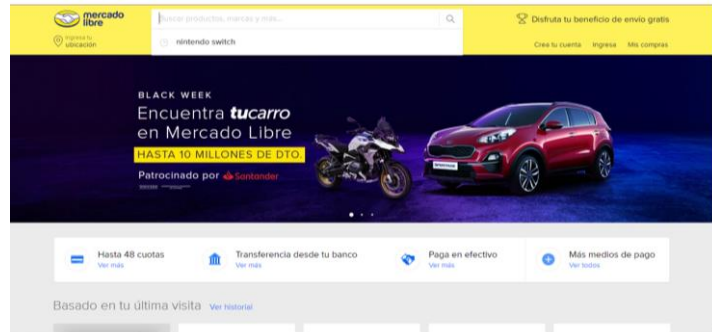
Santiago Florián Bustamante<sup>1</sup>

<sup>1</sup> Pontificia Universidad Javeriana, Cali, Colombia  
santiflo@javerianacali.edu.co

**Abstract.** El siguiente ensayo trata sobre la inteligencia aplicada al procesamiento del lenguaje natural y como este aprendizaje puede ser usado para el análisis de sentimiento en los comentarios de los productos de las plataformas de ventas virtuales para la clasificación de un buen producto. Habla del procesamiento que se le debe realizar al texto previo a la clasificación, ya que el texto presenta elementos y figuras literarias que pueden inducir al error. Al final se presenta la plataforma Driverless AI desarrollada por H2O.ai la cual permite realizar el análisis de textos.


**Keywords:** Inteligencia Artificial, Procesamiento del lenguaje natural, análisis de sentimientos en textos

Muchas empresas ofrecen una plataforma de compra y ventas a los usuarios para que puedan ofrecer productos o servicios. En la figura 1 se puede ver Mercado libre, la mejor plataforma de ventas en Latinoamérica, reconocida por brindarle a los usuarios un espacio en el cual pueden ofrecer y comprar productos que se encuentren al público. Los usuarios pueden calificar el producto dependiendo del tipo de trato recibido a la hora de adquirir el producto. Esto va desde el momento en que el usuario accede a la página hasta la entrega final del producto. En el proceso intermedio el usuario puede calificar la disposición del vendedor, así como el estado en que el producto fue recibido. Todo el proceso de compra puede ser rastreado para obtener información que califique tanto al vendedor como el producto, es muy común encontrar la calificación representada mediante una escala grafica con estrellas. En la figura 2 se puede ver la clasificación de los productos brindada por los usuarios en la plataforma de mercado libre. Donde cero estrellas significan un mal producto, hasta cinco estrellas la cual es la máxima calificación posible e indica excelente producto



**Figura 1. MercadoLibre es la mejor plataforma para vender productos que existe en Latinoamérica [1].**

### Opiniones sobre el producto

4.8   
Promedio entre 29 opiniones



Excelente

Excelente es una consola con muy buena calidad y la facilidad de llevarla a cualquier parte y jugar sin necesidad de un televisor.



Aceptable

Por los juegos, bien, la consola se calienta, los controles son algo incomodos, el precio en mercadolibre es bueno, comparado con otras tiendas.



Muy bueno

Buena la consola, sin embargo el estuche no es muy bueno y es diferente a aquel que sale en las imagenes.

**Figura 2. Comentarios de Nintendo Switch Neón Blue And Red + Estuche. Nueva Y Sellada de la página de Mercado libre [2].**

Algunas plataformas implementan algoritmos de aprendizaje automático para realizar una clasificación de que tan bueno resulta adquirir el producto que ofrece un determinado usuario, usando información obtenida por otros usuarios que han tenido la oportunidad de interactuar en el proceso de compra del producto. Las plataformas realizan encuestas a los compradores, analizan la cantidad de visitas y ventas concretadas por el producto, tiempo transcurrido desde la compra hasta la entrega del producto y lo que hablaremos más a fondo, el análisis de sentimientos textos en los comentarios de los compradores.

El análisis de sentimientos, también conocido como “opinión mining”, es un subcampo del procesamiento del lenguaje natural (PLN) que intenta identificar y extraer opiniones de un texto dado [3]. Esto representa un gran avance, debido a que las empresas tradicionalmente realizan encuestas y estudios a grupos objetivos para obtener comentarios de los consumidores. Aunque, la tecnología en conjunto con el aprendizaje automático y la inteligencia artificial han permitido analizar el texto de una amplia variedad de fuentes con mucha más precisión. Cabe resaltar que el extraer emociones del texto es una herramienta muy valiosa que tiene el potencial de mejorar el retorno de la inversión (ROI) de muchas empresas.

Con el auge de las redes sociales como medio de comunicación, las emociones pueden ser capturadas a través de los comentarios de los usuarios, lo cual termina siendo una información valiosa desde

una perspectiva comercial. Al evaluar cuidadosamente la opinión y los sentimientos de las personas, las empresas pueden averiguar razonablemente qué piensan las personas sobre un producto y, en consecuencia, incorporar retroalimentaciones.

Antes de poder aplicar el análisis de sentimientos a algún comentario realizado por un usuario se deben realizar los siguientes filtros:

- Dividir el texto en palabras.
- Eliminar caracteres especiales como comillas suspendidas (“...”), signos de admiración (!), etc. Debido a que no aportan información a la hora de evaluar un producto.
- Eliminar las palabras vacías. Las cuales no generan un valor agregado. Tales como los pronombres (Lo, la, le).
- Clasificación de las palabras restantes. Las cuales pueden estar en los conjuntos de palabras negativas, palabras neutrales, palabras positivas. Donde en las palabras negativas se le asigna un valor de menos uno (-1), palabras neutrales cero (0) y palabras positivas más uno (+1).

Al realizar la clasificación de las palabras negativas, positivas y neutrales. Se puede calcular una media del comentario realizado por un usuario. Al tener múltiples comentarios realizados por distintos usuarios se puede clasificar un producto en una escala de cero (0) a cinco (5) estrellas. El final la empresa puede tomar medidas para impulsar las ventas de un determinado producto.

El análisis de sentimientos no es una tarea fácil de realizar. Los datos de texto a menudo vienen precargados con mucho ruido. El sarcasmo es uno de esos tipos de ruido presente de forma innata en las redes sociales y las opiniones de productos que pueden interferir con los resultados. Los textos sarcásticos demuestran un comportamiento único. A diferencia de una simple negación, una oración sarcástica transmite un sentimiento negativo utilizando solo una connotación positiva de las palabras.

El análisis de sentimientos puede confundirse fácilmente con la presencia de palabras tan sarcásticas y, por lo tanto, la detección de sarcasmo es un paso de preprocesamiento vital en muchas tareas de PLN. Es útil identificar y eliminar las muestras ruidosas antes de entrenar modelos para aplicaciones de PLN.

Driverless AI es un producto de machine learning de H2O.ai. Viene equipado con guías de procesamiento del lenguaje natural (PLN) para la clasificación de texto y problemas de regresión. La plataforma admite texto independiente y texto con otros valores numéricos como características predictivas.

La plataforma convierte automáticamente las cadenas de texto en características utilizando técnicas potentes como TFIDF, CNN y GRU. Con TensorFlow, Driverless AI también puedes procesar bloques de texto más grandes y construir modelos utilizando todos los datos disponibles para resolver problemas comerciales. Driverless AI tiene capacidades de PLN de última generación para el análisis de sentimientos, y la utilizaremos para construir un clasificador de detección de sarcasmo.

Se puede concluir que el análisis de sentimientos para textos es una herramienta muy poderosa para la automatización de encuestas a los usuarios sobre un determinado producto. El análisis de sentimientos permite al dueño del producto o servicio identificar la satisfacción de sus clientes. Antes de poder realizar una clasificación se debe realizar un filtro del texto, ya que para una inteligencia artificial información de mas se traduce en ruido para una correcta clasificación. El sarcasmo como figura literaria representa un gran problema para la IA, ya que son palabras que de manera semántica tienen un cierto significado pero que en general es contrario a lo que indica la oración. Driversless IA es un producto de machin lerning desarrollado por H20.ia que permite realizar el análisis de sentimientos para textos, la cual realizar el filtrado de caracteres especiales, eliminación de palabras vacías y la detección de sarcasmos, para una correcta clasificación del texto.

## Bibliografía

[1]W. Cuban, "LAS MEJORES PLATAFORMAS DE VENTA ONLINE", *Medium*, 2019. [Online]. Available: <https://medium.com/somosacademy/las-mejores-plataformas-de-venta-on-line-b6c8a18f83af>. [Accessed: 18-Nov- 2019]

[2]W. Cuban, "LAS MEJORES PLATAFORMAS DE VENTA ONLINE", *Medium*, 2019. [Online]. Available: <https://medium.com/somosacademy/las-mejores-plataformas-de-venta-on-line-b6c8a18f83af>. [Accessed: 18-Nov- 2019]

[3]P. Pandey, "Inteligencia Artificial para el análisis de sentimientos en textos", *Medium*, 2019. [Online]. Available: <https://planetachatbot.com/inteligencia-artificial-para-an%C3%A1lisis-sentimientos-en-textos-e52782ad05e9>. [Accessed: 18- Nov- 2019]