



Pistas sentimentales en las redes sociales

Por Asdrúbal López Chau



Ilustración: Luis Ángel Velázquez

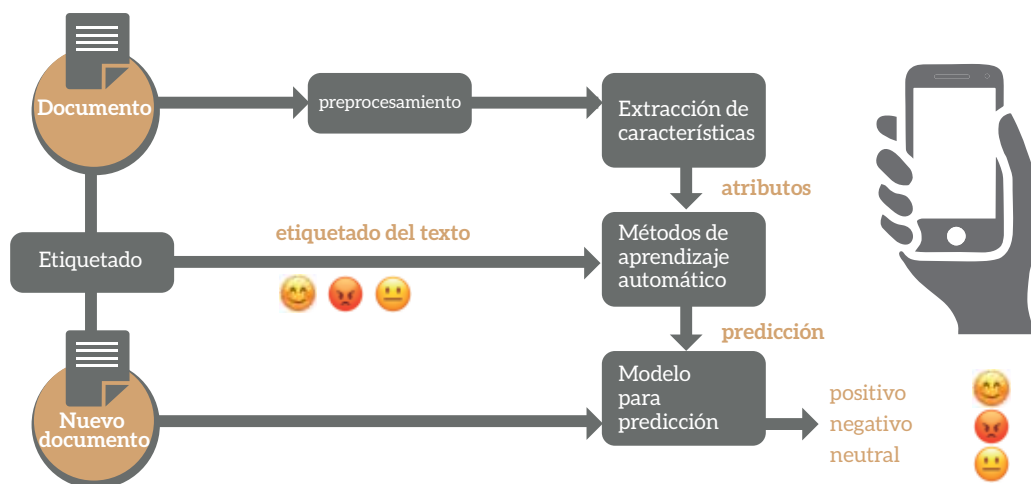
El análisis de sentimientos es un proceso computacional para identificar si una opinión expresada en un documento es positiva, negativa o neutral respecto a algo o alguien. Gracias a las redes sociales, este tipo de estudios ha llegado a ser muy importante para las empresas y otras organizaciones, ya que les permite ver de una manera casi inmediata lo que opinan los usuarios de sus productos o servicios.

El método más utilizado actualmente para realizar este análisis se basa en inteligencia artificial o, en específico, en

métodos de aprendizaje automático. El enfoque general se muestra en la figura 1. El documento puede ser cualquier texto, como una publicación en alguna red social.

Para generar modelos que predigan la clase (positivo, negativo o neutral) de un documento se crean atributos que pueden ser utilizados por métodos de aprendizaje supervisado, como las redes neuronales, máquinas de vectores, soporte o árboles de decisión. Entre las características más comunes que se extraen de documentos se encuentran los siguientes:

Figura 1. Proceso general simplificado del análisis de sentimientos



Fuente: Elaboración propia.

1. Etiquetado gramatical (*POS-tagging*): identifica cada tipo de palabra en el documento (sujeto, verbo, adjetivo, femenino, masculino, plural, singular, etcétera).
2. N-gramas: vocablos que con frecuencia se escriben juntos, como *buena idea, mala idea, por ejemplo, es grato informarle*. Se clasifican en uni-grama, bi-grama y más.
3. Frecuencia inversa del término: es un hecho que las palabras encontradas con menos frecuencia en un documento ofrecen más información que las que aparecen más frecuentemente.
4. Word2vect: método que convierte cada palabra en un vector único y multidimensional, usando redes neuronales de dos capas. Estos vectores son obtenidos de los pesos sinápticos de una red neuronal. En ese espacio, las palabras relacionadas corresponden a vectores similares; por ejemplo, cereal y desayuno serían vectores muy cercanos entre ellos, pero alejados de los vectores correspondientes a las palabras *vértice* o *cuántico*.

5. Bolsa de palabras (*bag of words*): analiza todas las palabras de acuerdo con el orden de aparición.

Estas transformaciones de vocablos en atributos son esenciales para que las computadoras entiendan lo que escribimos y determinar los sentimientos que escribimos en nuestras redes sociales. 🌐

Referencia

Pasupa, K. y Thititorn Seneewong Na Ayutthaya, T. (2019) "Thai sentiment analysis with deep learning techniques: A comparative study based on word embedding, POS-tag, and sentic features", en *Sustainable Cities and Society*. < doi: 10.1016/j.scs.2019.101615 >.



Asdrúbal López Chau es doctor en Ciencias en Computación por el CINVESTAV-IPN, miembro del Sistema Nacional de Investigadores Nivel I desde el 2015. Actualmente es profesor de tiempo completo del Programa Educativo de Ingeniero en Computación, en el Centro Universitario UAEM Zumpango.

