



2019 10 al 13 de septiembre - Cartagena de Indias, Colombia

RETOS EN LA FORMACIÓN
DE INGENIEROS EN LA
ERA DIGITAL

ANÁLISIS DE SENTIMIENTOS, UNA HERRAMIENTA PARA VALORAR LA ACTITUD DEL ESTUDIANTE FRENTE A UN CURSO

Sandra Patricia Castillo Landínez, Pablo Eduardo Caicedo Rodríguez

**Corporación Universitaria Autónoma del Cauca
Popayán, Colombia**

Resumen

Una categoría especial del análisis de texto y el procesamiento de lenguaje natural es el análisis de sentimientos (Sentiment Analysis), enfocado en examinar el tono emocional (positivo, negativo o neutro) que las personas plasman en los textos que escriben sobre un tema particular en medios digitales como Facebook, Twitter, LinkedIn, foros, microblogs, entre otros. El tema comúnmente se asocia con las redes sociales, pero su campo de aplicación es mucho más amplio; para las organizaciones tiene un valor significativo hacer seguimiento a las opiniones y comentarios de sus clientes y usuarios sobre el grado de satisfacción en productos y servicios adquiridos, a fin de establecer planes y estrategias de mejora.

En el ámbito académico es importante analizar el desempeño de los estudiantes paralelo a la exploración de estrategias que permitan subsanar los inconvenientes generados durante el proceso enseñanza-aprendizaje. En tal sentido, el análisis de sentimientos constituye una interesante herramienta para estudiar las emociones de los educandos expresadas en los comentarios publicados en los blogs de sus compañeros y la evaluación cualitativa que realizan al finalizar un determinado curso. Los sentimientos negativos pueden ser indicadores de posibles deserciones, conflictos de índole personal con los compañeros o los docentes o dificultades personales, situaciones que identificadas a tiempo pueden ser tratadas por expertos profesionales en las unidades de Bienestar Universitario.

Se realizó una prueba piloto en cuatro cursos del componente de investigación en la Facultad de Ingeniería de la Corporación Universitaria Autónoma del Cauca; los comentarios de los estudiantes publicados en los blogs de sus compañeros fueron analizados con una herramienta de minería de datos que permitió comparar los textos con un léxico de libre distribución a fin de determinar el tono emocional. Las conclusiones de esta exploración permiten alimentar los

procesos de autoevaluación y cuantificar la evolución del estudiante, además apoyan las estrategias de permanencia estudiantil impulsadas por la Facultad.

Palabras clave: análisis de sentimientos; blog; herramienta de aprendizaje

Abstract

A particular category of text analysis and natural language processing is sentiment analysis, focused on examining the emotional tone (positive, negative or neutral) that people express in texts they write about a particular topic in digital media such as Facebook, Twitter, LinkedIn, forums, microblogs, among others. The topic is commonly associated with social networks, but its field of application is much broader; for organizations, it is of significant value to follow up on the opinions and comments of their clients and users regarding the degree of satisfaction with products and services acquired, in order to establish plans and strategies for improvement.

In the academic field, it is crucial to analyze the students' performance parallel to the exploration of strategies to overcome the inconveniences generated during the teaching-learning process. In this sense, the analysis of feelings is a tool to study the emotions of the students expressed in the comments published in the blogs of their peers and the qualitative evaluation they perform at the end of a course. Negative feelings can be indicators of possible desertions, conflicts of a personal nature with colleagues or teachers or personal difficulties, situations that identified in time can be treated by professional experts in the University Welfare Units.

A pilot test was carried out in four courses of the research component in the Faculty of Engineering of the Autonomous University Corporation of Cauca; The students' comments published on their colleagues' blogs were analyzed with a data mining tool that allowed them to compare the texts with a freely distributed lexicon in order to determine the emotional tone. The conclusions of this exploration allow to feed the self-evaluation processes and quantify the evolution of the student, besides they support the strategies of student permanence promoted by the Faculty.

Keywords: sentiment analysis; blog; learning tool

1. Introducción

La incursión de la Web 2.0 sumado a las actuales condiciones tecnológicas permiten a las personas mantenerse conectadas 24/7 y en consecuencia ser muy activas en plataformas (como chat, conferencias, transacciones en línea, comercio electrónico, etc.) y redes sociales, expresando y compartiendo sus opiniones respecto a diversos temas o simplemente publicando sus emociones, gustos, expectativas o desavenencias sobre sus actividades cotidianas o un asunto particular (Ravi & Ravi, 2015).

Estos canales pasaron de ser un simple mecanismo de interacción de las nuevas generaciones a convertirse en un poderoso agente influenciador en ámbitos tan disímiles como los negocios, el

turismo, el comercio, la industria, la economía e incluso la política; en sus redes las personas tienen tanto poder como para difundir rumores que afecten (de manera positiva o negativa) las finanzas de una organización, así como destruir la reputación y el buen nombre de terceros.

Su relevancia radica en que hoy en día, los usuarios al momento de comprar o suscribir un producto o un servicio, toman sus decisiones influenciados por los comentarios que han publicado los clientes de una organización en su sitio web; de otra parte, estas reseñas también son aprovechadas por las empresas para instaurar políticas orientadas a mejorar la calidad y los estándares de sus productos, servicios y atención al cliente.

El análisis de sentimiento (AS) también conocido como minería de opinión, se enfoca en detectar, extraer, analizar y clasificar actitudes, opiniones y comentarios publicados en diferentes plataformas digitales (Facebook, Twitter, LinkedIn, foros, microblogs, entre muchos otros) respecto a temas variados; estos datos se estudian utilizando una combinación de algoritmos de machine learning, técnicas de minería de datos, minería web y minería de textos (Agarwal & Mittal, 2016) con el fin de establecer el tono emocional (positivo, negativo o neutro) que encierra el mensaje (Haddi, Liu, & Shi, 2013; Serrano-Guerrero, Olivas, Romero, & Herrera-Viedma, 2015)

Un punto interesante de estudio es ¿Cuál es la efectividad de los resultados de este tipo de análisis?, por más sofisticados que resulten ser los desarrollos en este campo, no hay una herramienta que sea 100% precisa, entre otras razones porque indagar los sentimientos es una tarea altamente subjetiva, incluso para los mismos humanos.

Las tareas básicas del análisis de sentimientos son el reconocimiento de la emoción y la detección de la polaridad; la primera se enfoca en extraer un conjunto de etiquetas de emoción y la segunda, generalmente es una labor de clasificación binaria con resultados tales como "positivo" y "negativo"; se consideran tres enfoques principales para realizar el AS (Cambria, 2016):

- Técnicas basadas en conocimiento: el texto se clasifica en categorías haciendo uso de Léxicos preexistentes u otras bases de conocimiento probabilístico entrenadas a partir de corpus lingüísticos.
- Métodos estadísticos: emplea técnicas como las máquinas de soporte vectorial y el aprendizaje profundo, son ampliamente usados para la clasificación de los textos
- Enfoques híbridos: combina los dos anteriores

Existen diferentes autores que plantean el uso de AS en el ámbito educativo, Kechaou y colaboradores proponen un modelo basado en el tono emocional de las palabras y lo utilizan para generar un clasificador basado en máquinas de soporte vectorial; sin embargo, sus datasets de trabajo están basados en almacenes de conocimiento descargables de internet (Kechaou, Ammar, & Alimi A, 2011). Otro caso se puede encontrar en el trabajo de Barrón-Estrada y colaboradores (Barrón-Estrada, Zatarain-Cabada, Ramírez-Ávila, Oramas-Bustillos, & Graff Guerrero, 2018) donde se desarrolla un analizador de sentimiento basado en diferentes tipos de clasificadores, pero sin analizar el tono emocional a nivel de palabra. En resumen, el análisis de sentimiento es usado para evaluar el desempeño del estudiante y así evitar emociones negativas

que disminuyan la tasa de abandono (Wen, Yang, & Rosé, 2014), el desempeño de los docentes (Balahadia, Fernando, & Juanatas, 2016), el desagaste de los estudiantes (Chaplot, Rhim, & Kim, 2015), detección de estereotipos (Arceo-Gomez & Campos-Vazquez, 2019).

En este trabajo se analizaron los mensajes publicados en blogs, sitios web donde el autor publica de manera periódica contenidos de diversa índole temática (entradas o posts), permitiendo a sus lectores participar a través de comentarios reflexiones u opiniones (Molina Aventosa, Valenciano Valcárcel, & Valencia-Peris, 2015). En el ámbito educativo constituyen una interesante herramienta para fomentar la comunicación, el trabajo colaborativo y la generación de redes de aprendizaje; apoyan los procesos de enseñanza-aprendizaje ya que el estudiante se convierte en un sujeto activo responsable de su propio aprendizaje (Molina Alventosa, Valencia Peris, & Suárez Guerrero, 2015; Pedrero García & Morón Marchena, 2015).

2. Análisis de sentimientos en blogs

Durante tres semestres los estudiantes matriculados en la asignatura Taller de Investigación del programa de Ingeniería de Sistemas Informáticos (Corporación Universitaria Autónoma del Cauca), construyeron un blog en el cual publicaron los ítems básicos de una propuesta de investigación (problema de investigación, justificación, los objetivos, etc.); cada entrada recibió comentarios de sus compañeros y se buscó propiciar el intercambio respetuosos de opiniones, la participación activa, la generación de un ambiente de crítica constructiva y de autocrítica y finalmente el enriquecimiento de la idea original de investigación.

El proyecto utilizó los comentarios publicados en los diferentes blogs (Figura 1) para realizar AS y establecer si las opiniones de los estudiantes respecto al trabajo realizado por sus compañeros ofrecían un tono positivo o negativo; se usó la técnica basada en conocimiento, descrito por medio del Léxico afectivo de Finn Årup Nielsen conocido como AFINN, que maneja una escala de tono emocional entre -5 (negativa) y +5 (positiva). Se realizó una adaptación para incluir términos y expresiones propios del ámbito y la cultura regional.

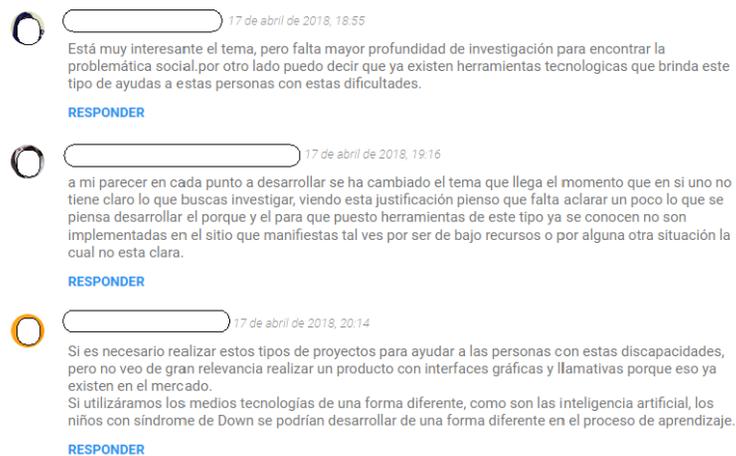


Figura 1. Comentarios generados después de una publicación

Cada palabra de un comentario fue comparada con el Léxico y valorada de acuerdo a la escala, a fin de establecer la connotación positiva o negativa del mensaje. Con la estimación cuantitativa se procedió a realizar una evaluación estadística que permitió cuantificar el desempeño del estudiante respecto a la herramienta.

3. Resultados

La versión final del Léxico Afinn estaba compuesta por 2783 términos, de los cuales 1088 son vocablos positivos y 1695 negativos. Después de eliminar las expresiones que no aportaban información relevante al texto (artículos, preposiciones, pronombres) se obtuvieron 360 palabras únicas.

En las Figuras 2 y 3 se observa la frecuencia de aparición de las 10 palabras positivas y negativas más usadas más por los estudiantes en sus comentarios. Aunque el término “problema” tiene en el Léxico una connotación negativa, es importante aclarar que, en el contexto del curso hace referencia a observaciones hechas sobre el problema de investigación. También se advierte que en casi 200 ocasiones se usó la palabra “interesante” como un calificativo del trabajo desarrollado por los compañeros, y en un número similar de oportunidades se inició el comentario mediante el saludo “hola”.

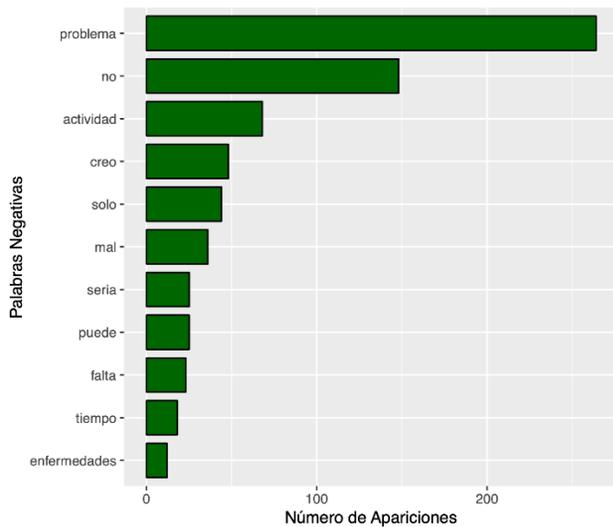


Figura 2. Palabras negativas más usadas. Fuente propia

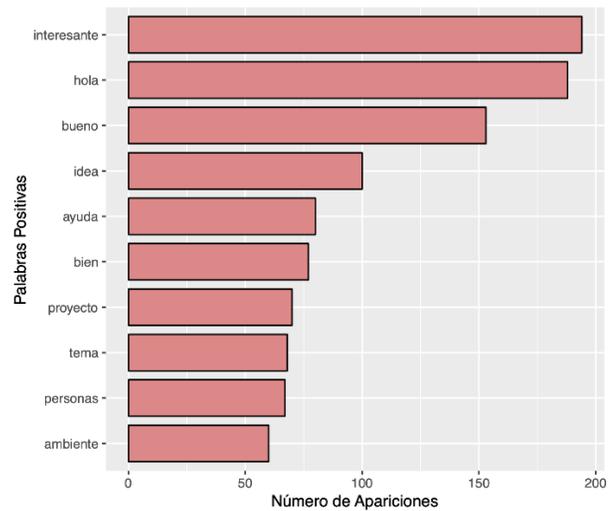


Figura 3. Palabras positivas más usadas. Fuente propia

Las nubes de palabras (Figuras 4 y 5) son representaciones gráficas de los términos más usados, los tamaños y colores de letra representan pesos de frecuencias relativas dentro de los comentarios.



Figura 4. Nube de palabras negativas. Fuente propia



Figura 5. Nube de palabras positivas. Fuente propia

Finalmente, mediante histogramas de frecuencia (Figura 6) se puede sintetizar el valor emocional de los comentarios publicados en el blog, se observa un sesgo importante hacia la positividad, lo cual corrobora los resultados obtenidos anteriormente en la clasificación de términos.

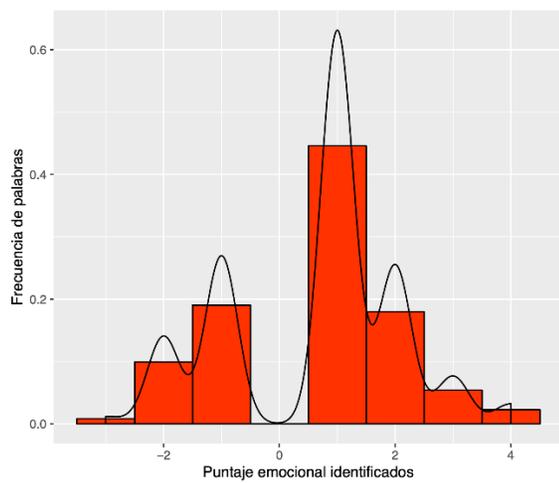
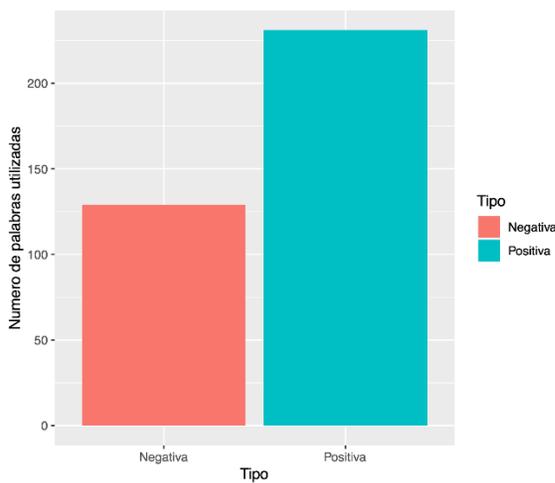


Figura 6. Histogramas de frecuencia de para cada tipo de valor emocional. Fuente propia

4. Conclusiones

- El análisis de sentimiento resulta una técnica de alta complejidad, tanto por las implicaciones de procesamiento de datos que involucra (Procesamiento de Lenguaje Natural-NLP, limpieza y filtrado, construcción de clasificadores, evaluación de resultados) cómo por la naturaleza de los resultados que se pueden obtener y su nivel de precisión dada la ambigüedad que puede contener un estado de ánimo.
- El lenguaje natural es complejo y a la vez muy rico en expresiones coloquiales y múltiples dialectos lo que añade mayor complejidad al análisis de sentimiento, ¿Hasta qué punto un

algoritmo puede identificar ironías, sarcasmo, símbolos o siglas (usadas en muchas tribus urbanas), jergas o variaciones culturales?

- Como usos alternos, es posible realizar análisis de sentimientos en diferentes instrumentos de valoración de la satisfacción estudiantil permitiendo efectuar una eficiente asignación de recursos a programas o facultades dentro de las Instituciones Educativas.

4. Referencias

Artículos de Revistas

- Arceo-Gomez, E. O., & Campos-Vazquez, R. M. (2019). Gender stereotypes: The case of MisProfesores.com in Mexico. *Economics of Education Review*, 72, 55–65. <https://doi.org/10.1016/j.econedurev.2019.05.007>
- Balahadia, F. F., Fernando, M. C. G., & Juanatas, I. C. (2016). Teacher's performance evaluation tool using opinion mining with sentiment analysis. In 2016 IEEE Region 10 Symposium (TENSYP) (pp. 95–98). IEEE. <https://doi.org/10.1109/TENCONSpring.2016.7519384>
- Barrón-Estrada, M. L., Zatarain-Cabada, R., Ramírez-Ávila, S. L., Oramas-Bustillos, R., & Graff Guerrero, M. (2018). Uso de analizador de emociones en sistemas educativos inteligentes. *Research in Computing Science*, 147(6), 179–188. Recuperado de [http://www.rcs.cic.ipn.mx/rcs/2018_147_6/Use de analizador de emociones en sistemas educativos inteligentes.pdf](http://www.rcs.cic.ipn.mx/rcs/2018_147_6/Use%20de%20analizador%20de%20emociones%20en%20sistemas%20educativos%20inteligentes.pdf)
- Cambria, E. (2016). Affective Computing and Sentiment Analysis. *IEEE Intelligent Systems*, 31(2), 102–107. <https://doi.org/10.1109/MIS.2016.31>
- Chaplot, D., Rhim, E., & Kim, J. (2015). Predicting Student Attrition in MOOCs using Sentiment Analysis and Neural Networks. *AIED Workshops*.
- Haddi, E., Liu, X., & Shi, Y. (2013). The Role of Text Pre-processing in Sentiment Analysis. *Procedia Computer Science*, 17, 26–32. <https://doi.org/10.1016/j.procs.2013.05.005>
- Kechaou, Z., Ben Ammar, M., & Alimi, A. (2011). Improving e-learning with sentiment analysis of users' opinions. In *2011 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1032–1038). IEEE. <https://doi.org/10.1109/EDUCON.2011.5773275>
- Molina Alventosa, J. P., Valencia Peris, A., & Suárez Guerrero, C. (2015). Percepción de los estudiantes de una experiencia de uso didáctico de blog docente en Educación Superior. *Educación XX1*, 19(1). <https://doi.org/10.5944/educxx1.15579>
- Molina Aventosa, P., Valenciano Valcárcel, J., & Valencia-Peris, A. (2015). Los blogs como entornos virtuales de enseñanza y aprendizaje en Educación Superior. *Revista Complutense de Educación*, 26. https://doi.org/10.5209/rev_RCED.2015.v26.43791
- Pedrero García, E., & Morón Marchena, J. (2015). Experiencia universitaria con blogs en educación para la salud. *IJERI: International Journal of Educational Research and Innovation*, 5, 150–159. Recuperado de <https://www.upo.es/revistas/index.php/IJERI/article/view/1607>
- Ravi, K., & Ravi, V. (2015). A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. *Knowledge-Based Systems*, 89, 14–46. <https://doi.org/10.1016/j.knosys.2015.06.015>

- Serrano-Guerrero, J., Olivas, J. A., Romero, F. P., & Herrera-Viedma, E. (2015). Sentiment analysis: A review and comparative analysis of web services. *Information Sciences*, 311, 18–38. <https://doi.org/10.1016/j.ins.2015.03.040>
- Wen, M., Yang, D., & Rosé, C. P. (2014). Sentiment analysis in MOOC discussion forums: What does it tell us? *Proceedings of Educational Data Mining*, 130–137.

Libros

- Agarwal, B., & Mittal, N. (2016). Machine Learning Approach for Sentiment Analysis. In *Prominent Feature Extraction for Sentiment Analysis* (pp. 21–45). Springer, Cham. https://doi.org/10.1007/978-3-319-25343-5_3

Sobre los autores

- **Sandra Patricia Castillo Landínez:** Ingeniera de Sistemas (Universidad Nacional de Colombia), Especialista en Administración de la Información y Bases de Datos (Colegio Mayor del Cauca), Certified Big Data Professional, Certified Big Data Scientist. Docente de la Facultad de Ingeniería, investigadora adscrita al Grupo de Investigación en Tecnología y Ambiente (GITA), coordinadora de la línea de Investigación en Ingeniería de Software y líder del Semillero de Investigación en Minería de Datos (SIMD). sandra.castillo.l@uni-autonoma.edu.co
- **Pablo Eduardo Caicedo Rodríguez:** Ingeniero en Electrónica y Telecomunicaciones (Universidad del Cauca), Magister en Ingeniería Electrónica y Telecomunicaciones, Certified Big Data Professional, Certified Big Data Scientist, Candidato a Doctor en Ciencias de la Electrónica, Docente de la Facultad de Ingeniería, investigador adscrito al Grupo de Investigación en Tecnología y Ambiente (GITA). pablo.caicedo.r@uni-autonoma.edu.co

Los puntos de vista expresados en este artículo no reflejan necesariamente la opinión de la Asociación Colombiana de Facultades de Ingeniería.

Copyright © 2019 Asociación Colombiana de Facultades de Ingeniería (ACOFI)