



CIENCIAS TÉCNICAS

Artículo original de investigación

Contribuciones al análisis de sentimientos en opiniones de usuario aplicando técnicas y algoritmos de inteligencia artificial

Alfredo Javier Simón Cuevas ^{1*} <https://orcid.org/0000-0002-6776-9434>
Orlando Grabiél Toledano López ² <https://orcid.org/0000-0001-8263-0425>
Julio Madera Quintana ³ <https://orcid.org/0000-0001-5551-690X>
Hector Raúl González Díez ² <https://orcid.org/0000-0002-7601-4201>
Alejandro Ariel Ramón Hernández ⁴ <https://orcid.org/0000-0003-0749-5977>
María Matilde García Lorenzo ⁴ <https://orcid.org/0000-0002-1663-5794>
Jesús Serrano Guerrero ⁵ <https://orcid.org/0000-0002-6177-8188>
José Ángel Olivas Varela ⁵ <https://orcid.org/0000-0003-4172-4729>
Leticia Arco García ⁷ <https://orcid.org/0000-0002-5154-4441>
José Manuel Perea Ortega ⁶ <https://orcid.org/0000-0002-7929-3963>
Francisco P. Romero ⁵ <https://orcid.org/0000-0002-6993-2434>
Oscar Miguel González Parets ¹ <https://orcid.org/0009-0009-0530-9211>
Patricia Montañez Castelo ¹ <https://orcid.org/0000-0003-4506-5054>

¹ Universidad Tecnológica de La Habana José Antonio Echeverría. La Habana, Cuba

² Universidad de las Ciencias Informáticas. La Habana, Cuba

³ Universidad de Camagüey Ignacio Agramonte Loynaz. Camagüey, Cuba

⁴ Universidad Central Martha Abreu de las Villas. Santa Clara, Cuba

⁵ Universidad de Castilla La Mancha, Ciudad Real, España

⁶ Universidad de Extremadura, Badajoz. Extremadura, España

⁷ Universidad Libre de Bruselas. Bruselas, Holanda

* Autor para la correspondencia: asimon@ceis.cujae.edu.cu

Editor

Lisset González Navarro
Academia de Ciencias de Cuba.
La Habana, Cuba

Traductor

Darwin A. Arduengo García
Academia de Ciencias de Cuba.
La Habana, Cuba

RESUMEN

Introducción: La minería de opinión resulta de gran interés y utilidad para la sociedad y sectores importantes de la economía. Los textos de las opiniones son datos complejos caracterizados por complicadas relaciones semánticas y sintácticas del idioma, lo que implica un desafío a considerar desde el punto de vista computacional para la toma de decisiones.

Objetivos: Incrementar la eficacia en el entrenamiento neuronal artificial en la clasificación de textos. Desarrollar nuevos métodos que combinan recursos lingüísticos y modelos neuronales de lenguaje para la detección de polaridad y análisis de sentimientos. **Métodos:** Se proponen un conjunto de métodos para la clasificación del sentimiento (polaridad positiva o negativa) en opiniones de usuarios con enfoques no supervisados y supervisados. Adicionalmente se introduce un nuevo método híbrido para la detección de emociones en opiniones de usuarios con enfoque semántico de selección de características. **Resultados:** Se evaluaron los métodos desarrollados en diferentes colecciones de prueba con opiniones de diferentes ámbitos. Se usaron colecciones utilizadas en desafíos científicos promovidos por comunidades especializadas en la temática: SemEval2016, SemEval2018 y TASS-2020; colecciones

provenientes de sitios web de noticias y sus comentarios, a saber: opiniones provenientes de la British Broadcasting Corporation (BBC), de Youtube y de Cubadebate; y opiniones sobre lugares turísticos provenientes de TripAdvisor. **Conclusiones:** Los métodos desarrollados representan contribuciones que arrojan resultados prometedores en el estado del arte para el análisis de opinión, siendo muy competitivos y en determinados escenarios superiores.

Palabras clave: minería de opinión; análisis de sentimientos; clasificación supervisada de textos; clasificación no supervisada de textos

Contributions to sentiment analysis in user opinions applying artificial intelligence techniques and algorithms

ABSTRACT

Introduction: Opinion mining is of great interest and usefulness to society and important sectors of the economy. Opinion texts are complex data where complex semantic and syntactic relationships occur due to language, which imply a challenge to consider from the computational point of view for decision-making. **Objectives:** To increase the effectiveness of artificial neural training in text classification. To develop new methods that combine linguistic resources and neural models of language for polarity detection and sentiment analysis. **Methods:** They are proposed a set of methods for sentiment classification (positive or negative polarity) in user opinions with unsupervised and supervised approaches. Additionally, it is introduced a new hybrid method for emotion detection in user feedback with semantic feature selection approach. **Result:** We evaluated the methods on different databases for opinion analysis. The data come from news websites and their comments, British Broadcasting Corporation public databases, video comments from the Youtube platform and those proposed by SemeEval. We perform a polarity detection study on TripAdvisor comments based on the most representative places of Mexican and Cuban tourism. **Conclusions:** The methods developed represent contributions that yield promising results in the state of the art for opinion analysis, being very competitive and superior in certain scenarios.

Keywords: opinion mining; sentiment analysis; supervised text classification; unsupervised text classification

INTRODUCCIÓN

La clasificación de textos muestra utilidad en diferentes sectores de la sociedad, como el sector de la tecnología mediante la clasificación de comentarios y opiniones de clientes en línea, lo que permite a las empresas comprender mejor las necesidades y preferencias de sus clientes. ^(1,2,3) También se utiliza en la detección de spam, en la clasificación de contenido en sitios web y redes sociales. ^(4,5,6) Otra de las aplicaciones es el análisis de sentimientos en el ámbito del mercado, lo que permite a los inversores tomar decisiones más informadas y estratégicas. ⁽⁷⁾ Por otro lado, en la publicidad permite la personalización de anuncios y la segmentación de audiencias a partir de búsquedas en línea, interacciones en redes sociales

e historiales de navegación en sitios web. ^(8,9,10) Esto permite a las empresas dirigirse a audiencias específicas de manera más efectiva.

Recientes investigaciones en procesamiento de lenguaje natural (PLN) apuntan al sector turístico en aras de mejorar la calidad de los servicios y atraer visitantes de otras regiones. ^(11,12) Algunas plataformas en la web, como TripAdvisor, permiten a los turistas escribir su opinión sobre los lugares y atracciones turísticas que visitan. Las opiniones recogidas constituyen una fuente de datos que ayuda a identificar problemas a partir del análisis de los aspectos semánticos del contenido de las opiniones emitidas por los visitantes. ⁽¹³⁾ A priori la revisión de las opiniones puede proporcionar al visitante una idea sobre los servicios ofrecidos

en el lugar, su comodidad y si han superado las expectativas de los clientes. De este modo los destinos turísticos utilizan esta información para mejorar los servicios al cliente, ya que se sabe que las opiniones emitidas sobre un lugar influyen en la futura selección del destino a visitar. ^(13,14)

En análisis de sentimientos orientado a tópicos y aspectos significan un nivel de granularidad mayor en el procesamiento automático de las opiniones y mayor nivel de detalle en los resultados, dado que la polaridad no se determina sobre todo el contenido de la opinión, sino asociado a los tópicos o aspectos que se identifiquen. ⁽¹⁵⁾ Mediante el análisis de sentimientos basado en temas el contenido de la opinión se suele dividir en fragmentos de texto, en palabras o frases, luego se agrupan en temas y calcula los sentimientos de cada uno de ellos. Los aspectos comúnmente se refieren a atributos, características, cualidades de un producto, entidad o servicio, por tanto, aporta un mayor nivel de detalles, pero está más orientado a opiniones sobre entidades, productos, o servicios.

El proceso de análisis automático de sentimientos en textos de opinión es complejo, y se ha abordado a partir de diferentes tareas o problemas específicos. ⁽¹⁶⁾ Las investigaciones desarrolladas tratan algunas de esas tareas, a saber: detección de polaridad, detección de emociones, construcción de resúmenes automáticos de opiniones y la extracción de aspectos y su polaridad. En cada una de estas direcciones se reportan avances científicos importantes, aunque la mayor parte asociados a soluciones computacionales para opiniones en idioma inglés, por tanto, se necesita un mayor desarrollo de métodos dirigidos al idioma español, por la utilidad de estas tecnologías para nuestro país. Por otro lado, los resultados de eficacia obtenidos hasta ahora demandan ser mejorados. En tal sentido, se abordaron los siguientes retos:

- Mejora de la eficacia en la detección de polaridad del sentimiento de las opiniones desde diferentes perspectivas;
- construcción automática de resúmenes de opiniones sensibles al contexto. Mejora de la eficacia en la extracción automática de aspectos y en la predicción de la polaridad asociada;
- mejora del entrenamiento de redes neuronales profundas para la clasificación de textos y detección de polaridad;
- mejora de la eficacia en la detección de emociones en opiniones.

Con lo anterior el objetivo de este trabajo es exponer las principales contribuciones en el área de la minería de opinión, el análisis de sentimientos orientado a aspectos y análisis de polaridad en tareas de clasificación, identificación de tópicos y generación de resúmenes haciendo uso de las redes neuronales profundas con mecanismos eficientes de entrenamiento.

MÉTODOS

Método de análisis de sentimiento orientado a tópicos con construcción de resúmenes de opiniones

El método está orientado al modelado de tópicos en opiniones y la posterior construcción de resúmenes extractivos de las opiniones con un enfoque multidocumento, por tanto, los resúmenes se construyen con las sentencias (oraciones) más relevantes del conjunto de opiniones (documentos) y se obtiene un único resumen del conjunto de opiniones. ⁽¹⁷⁾ Se concibió para ser utilizado en el contexto de una plataforma de noticias, como Cubadebate, donde las opiniones se emiten en relación a una noticia que se publica. El método propuesto combina el modelado de tópicos y el contenido de las noticias, para determinar la “relevancia” de las oraciones de opinión e incluye el análisis del sentimiento para determinar la polaridad de las oraciones. ⁽¹⁸⁾ El mapeo oración-tópico y la contextualización temática permite alinear las oraciones con temas de opinión identificados y determinar los temas más relevantes asociados a la noticia. ⁽¹⁷⁾ Los temas menos relevantes se descartan y seguidamente se realiza un proceso de *ranking* de las oraciones, para finalmente construir el resumen, con las oraciones más relevantes

Método de análisis de sentimientos orientado a aspectos basado en aprendizaje profundo con transformers

Se ha desarrollado un nuevo método de análisis de sentimientos orientado a aspectos para el idioma español basado en el uso de aprendizaje profundo con transformers en ambas tareas (extraer aspectos y clasificar el sentimiento), lo que constituye una contribución en esta área. El método propuesto se concibió en 3 fases: a) preprocesamiento de las opiniones, b) extracción automática de aspectos y c) clasificar el sentimiento de la opinión hacia los aspectos extraídos. ⁽¹⁹⁾ En la extracción de aspectos y la determinación de la polaridad se aplicaron modelos preentrenados de aprendizaje profundo con transformers, específicamente se utilizó BETO para la extracción de aspectos y BERTIN para determinar la polaridad. ^(20,21)

Método para la detección de polaridad basado en tecnologías de aprendizaje profundo usando transformers con algoritmos de estimación de distribución

Se propone como principal contribución un nuevo método de entrenamiento para redes neuronales profundas (RNP) orientado fundamentalmente a mejorar la detección de pola-

ridad en opiniones. Este nuevo proceso de entrenamiento se ejecuta en 2 fases: a) aprendizaje de características mediante métodos basados en el gradiente descendente, y b) optimización del bloque de capas densas mediante los algoritmos de estimación de distribución (EDA).⁽²²⁾ El objetivo del entrenamiento consiste en encontrar una configuración de parámetros de la RNP que minimizan la función de pérdida *cross-entropy loss*. Mediante los EDA, la exploración y explotación de la superficie de optimización se lleva a cabo a partir del muestreo sobre un subconjunto de soluciones de la población ordenadas por la función de costo ascendentemente. A partir de lo anterior, se obtiene un conjunto intermedio o modelo de soluciones candidatas, luego se estima la distribución en el modelo de soluciones candidatas para generar nuevas soluciones que representan los pesos en la última capa.

En este sentido, se desarrollaron 2 arquitecturas de RNP. La primera se basa en una CNN (convolutional neural network) con un bloque de capas densas y como salida la función Softmax. La segunda, consiste en la adaptación de un transformer utilizando la parte del encoder preentrenada, y se le acopla a la salida del encoder un bloque de capas densas también con salida Softmax.

En la primera fase del entrenamiento se aplican los métodos SGD a y Adam⁽²³⁾ para entrenar la CNN, mientras para la arquitectura basada en transformer se emplea AdamW.⁽²⁴⁾ Por otro lado, en la segunda fase se aplicarán diferentes estrategias evolutivas: UMDAc, b EMNAC y CMA-ES. d La arquitectura basada en CNN recibe como entrada la representación de un documento mediante word-embedding usando GloVe.⁽²⁵⁾ Se utilizan 3 filtros convolucionales secuenciales, cada uno de los cuales utiliza en la salida una capa de agrupación máxima y ReLU como función de activación. El último filtro de la CNN se conecta con una capa totalmente conectada (FC) y la función Softmax como salida en la última capa. La segunda arquitectura consiste en una adaptación de un Transformer para el proceso de clasificación, que recibe una secuencia de palabras de entrada. Del Transformer, se utiliza la parte del encoder y la salida es acoplada a un bloque FC en la parte superior. Como modelos preentrenados en la parte del encoder, se utilizarán los modelos mT5 y RoBERTa.^(26,27)

Método híbrido para la detección de emociones en opiniones de usuario con enfoque semántico de selección de características

En esta investigación se abordó la detección de emociones desde un enfoque híbrido, donde se integran el uso de lexicones, con un esquema tradicional de clasificación

a SGD: gradiente descendente estocástico.

b UMDAc: algoritmo de distribución marginal univariado.

c EMNAC: algoritmo de estimación normal multivariado.

d CMA-ES: estrategia evolutiva de la adaptación de la matriz de covarianza.

supervisada de emociones en textos, la utilidad de este tipo de enfoque también referente al uso de conocimiento afectivo en un modelo de clasificación supervisada se demuestra también en Serrano-Guerrero J (parte de esta propuesta).^(30,28,29) En este sentido se concibió un método en 4 fases: a) preprocesamiento; b) evaluación de la orientación emocional; c) selección de características; y d) clasificador supervisado (predicción de la emoción).^(28,29)

Tomando como referencia el vocabulario de cada emoción del lexicón se determina el grado de afinidad semántica que tiene el contenido de la opinión con cada emoción, a través del cómputo de medidas de relaciones semánticas. La relación semántica se computa con la combinación de 4 métricas de relación semántica mediante un operador de agregación difuso compensatorio, específicamente el reportado en Zimmermann y Zysno (ecuación 1), donde γ es el grado de compensación proporcionado y se calcula según ecuación 2, en la que $T(e_1, e_2, \dots, e_n)$ es una función t-norma que se calcula según la ecuación 3, donde e_i el valor promedio de la relación semántica computada entre cada término de la opinión y el vocabulario de la emoción en el lexicón.⁽³¹⁾

El operador de agregación permite obtener un valor representativo de los valores obtenidos por cada una de las 4 métricas empleadas. Las métricas de relación semántica usadas son: *Pointwise Mutual Information* (PMI), *Chi-square* (CHI), las basadas en *WordNet-Resnik Similarity* y JCN.⁽³²⁾ Como resultado del proceso de evaluación de la orientación emocional se construye una matriz de afinidad, a partir de la cual se computa un grado de afinidad semántica (SAD, *Semantic Affinity Degree*) de cada opinión (o) con respecto a cada una de las emociones (E_j), según la ecuación 4.

Luego se selecciona la clase de emoción con la cual la opinión tiene mayor afinidad, siendo esta la de mayor valor de SAD(o, E_j), y a partir de esa emoción es que se construye el vector de características afectivas de cada opinión mediante un esquema binario (mejor resultado obtuvo en experimentos parciales). Por último, se pasa a la fase de predicción de la emoción a un clasificador supervisado, el cual recibe como entrada la matriz de características de la opinión. En este sen-

$$fuzz_sem_agg(e_1, e_2, \dots, e_n) = \left(\prod_{i=1}^n e_i \right)^{1-\gamma} * \left(1 - \prod_{i=1}^n (1 - e_i) \right)^{\gamma} \quad (1)$$

$$\gamma = \frac{T(e_1, e_2, \dots, e_n)}{T(e_1, e_2, \dots, e_n) + T(1 - e_1, \dots, 1 - e_n)} \quad (2)$$

$$T(e_1, e_2, \dots, e_n) = \prod_{i=1}^n e_i \quad (3)$$

$$SAD(o, E_j) = \frac{\sum_{i=1}^n w_i \in o O(w_i, E_j)}{|o|} \quad (4)$$

tido se consideraron los clasificadores: Support Vector Machine (SVM), Logistic Regression (LR), Multilayer Perceptron (MLP) y Naïve Bayes (NB), cada uno de los cuales fueron sometidos a procesos de entrenamiento y también fueron evaluados de forma independiente en los experimentos.

RESULTADOS Y DISCUSIÓN

En esta sección se esboza la evaluación de los 4 métodos desarrollados. En general, los datos son recopilados de sitios *web* de noticias y sus comentarios, bases de datos públicas de la BBC, comentarios de videos de la plataforma *Youtube* y las propuestas por el *SemeEval* en el año 2016 y 2018. Se realiza un estudio de detección de polaridad en comentarios de *TripAdvisor* a partir de lugares más representativos del turismo mexicano. Los métodos evaluados muestran la aplicación de modelos neuronales de lenguaje, el uso de recursos lingüísticos como lexicón para el español, algoritmos de aprendizaje automático y técnicas basadas en algoritmos evolutivos para el entrenamiento neuronal.

Evaluación del análisis de sentimiento orientado a tópicos con construcción de resúmenes de opiniones

La evaluación de esta solución se realizó utilizando información proveniente de Cubadebate, referente a noticias y sus opiniones, con lo que se conformaron 2 colecciones de prueba: TelecomServ, con 80 noticias de la empresa ETECSA publicadas del 2018-2020 y COVID-19 con 85 noticias publicadas entre marzo-agosto 2020. La calidad de los resultados se midió a través de la divergencia Jensen-Shannon (JSD) en la forma propuesta en Saggion y otros.⁽³³⁾ En este proceso de evaluación no se pudieron aplicar medidas de calidad comúnmente utilizadas en la evaluación de resúmenes automáticos, dado que necesitaban resúmenes elaborados por humanos para comparar. El

objetivo de esta evaluación fue comprobar si efectivamente los resúmenes construidos representaban información relevante del conjunto de opiniones y al mismo tiempo tenían alta relación con el contenido de la noticia, y la medida de divergencia JSD permitió medir ese aspecto. La tabla 1 muestran los resultados y la mejor combinación para cada base de datos. , los cuales se reportan por Hernández et al. (2020).⁽¹⁷⁾

Evaluación del análisis de sentimiento orientado a aspectos basado en aprendizaje profundo con transformers

La evaluación de la solución propuesta fue dividida en 2 objetivos: a) evaluar la calidad de la extracción de aspectos, y b) evaluar la calidad de la clasificación de polaridad; en correspondencia con el esquema empleado en la tarea 5 de *SemEval-2016*.⁽³⁴⁾ Los experimentos se ejecutaron sobre la colección de evaluación de opiniones en español de la temática de restaurantes usada en esa tarea.⁽³⁴⁾ En la experimentación se incluyeron 10 modelos de lenguaje basado en aprendizaje profundo con transformers: BETO, BERT multilingüe, ELECTRA (versión para español), BERTIN, RoBERTa-base, RoBERTa-large, GPT-2-spanish, y AIBETO en 3 tamaños diferentes.^(20,35,21,36) La calidad de los resultados se midió calculando precisión, Recall y F1 para la extracción de aspectos y además exactitud (Acc.) para clasificación de polaridad. Cada modelo fue entrenado durante 10 épocas y el tamaño del *batch* o lote, se definió igual a 8. Por otro lado, en la evaluación de la clasificación de sentimientos (polaridad), se midió la eficacia de los 2 métodos para definir el contexto de sentimiento de los aspectos: a) ventana de palabras, y b) dependencias sintácticas.

En la tabla 2 se comparan los resultados de la mejor alternativa (BETO) con los obtenidos por otros trabajos reportados en la literatura. Por otro lado, en la tabla 3, se muestran los resultados en la detección de polaridad de los aspectos.

Según se aprecia en la tabla 3, los resultados obtenidos por el método propuesto para ambas tareas son superiores a los

Tabla 1. Resultados usando representación basada en word2vec

Soluciones comparadas	TelecomServ		COVID-19	
	JSD _{Opiniones}	JSD _{news}	JSD _{Opiniones}	JSD _{news}
we1	0,278	0,487	0,314	0,453
we2	0,403	0,479	0,420	0,474
we4	0,388	0,388	0,392	0,404
we5	0,424	0,473	0,445	0,474
we6	0,416	0,459	0,439	0,460
TextRank (baseline: we3)	0,370	0,457	0,411	0,458

Tabla 2. Comparación de los resultados de la extracción de aspectos con otras soluciones

Soluciones reportadas	Precisión	Recall	F1
Henríquez Miranda & Buelvas (37)	-	-	0,730
GTI/C*	-	-	0,685
GTI/U*	-	-	0,683
IIT-T/U*	-	-	0,643
TGB/C*	-	-	0,557
SemEval-2016 - baseline	-	-	0,519
Solución propuesta (<i>bert-base-spanish-cased+Batch Size 8</i>)	0,736	0,815	0,767

Nota: * Soluciones en SemEval-2016 Task 5: ABSA (resultados en Slot2 español: OTE) ⁽³⁴⁾

reportados por otras soluciones (inclusive Martínez-Seis *et al.* del 2022), lo cual constituye una importante contribución en este campo, dada la ausencia de soluciones de este tipo para español, que incluyan aprendizaje profundo en ambas tareas. ⁽³⁸⁾ Además de los resultados a nivel experimental, la solución presentada forma parte de los resultados preliminares de un proyecto con la empresa de alta tecnología CITMATEL, cuyo objetivo es aplicar estas soluciones de análisis de sentimientos en las plataformas de comercio electrónico que posee la empresa, lo que se fundamenta en Mar Rodríguez *et al.* ⁽³⁹⁾

Evaluación de la detección de polaridad basado en tecnologías de aprendizaje profundo usando transformers con EDA

El método de entrenamiento propuesto se evaluó en dos escenarios, uno general enmarcado en la clasificación de textos que no son opiniones (noticias, *spam*, etc.) ⁽²²⁾, y otro más específico orientado a clasificar la polaridad en opiniones de

usuario. ^(40,41) Se incluyó el primero de los escenarios para evaluar el rendimiento de la solución propuesta dada sus potencialidades para ser generalizado a otras tareas de clasificación de texto, para ello se usaron las colecciones: BBC Sport (BBCS), BBC News (BBCN), y YouTube Spam Collection (YSC). ⁽²²⁾

En la evaluación para el problema de clasificación de textos se combinaron los métodos basados en gradiente SGD y Adam con diferentes variantes EDA: EMNA, UMDAc y CMA-ES. Se consideró como línea de base el promedio de los resultados reportados por las soluciones del estado de arte. En el estudio se incluyeron soluciones que aplican CNN, LSTM, la combinación de ambos, algoritmos basados en agrupamiento, así como enfoques de clasificación tradicionales: MLP, NB, LR, y RF, entre otros. Los resultados se encuentran publicados en Toledano-López *et al.* ^(22,42,43,44,45,46,47,48,49,50)

En la experimentación, se aprecia cómo las variantes propuestas incrementan la exactitud con respecto a los métodos

Tabla 3. Comparación de los resultados de la detección de polaridad con otras soluciones

Soluciones reportadas	Precisión	Recall	F1	Acc,
IIT-T/U*	-	-	-	0,835
TGB/C*	-	-	-	0,820
UWB/C*	-	-	-	0,813
INSIG/C*	-	-	-	0,795
Henríquez Miranda & Buelvas	-	-	-	0,848
Martínez-Seis <i>et al.</i> ,	-	-	-	0,796
SemEval-2016 - baseline	-	-	-	0,778
Solución Propuesta (RoBERTa + ventana_palabra)	0,836	0,864	0,848	0,866

Nota: * Soluciones en SemEval-2016 Task 5: ABSA (resultados en Slot3: Sentiment Polarity) ⁽³⁴⁾

basados en el gradiente de primer orden, lo que indica el efecto de los EDA para el refinamiento del bloque de capas densas en ambas arquitecturas neuronales. Desde el punto de vista de la arquitectura, con el *transformer* se logran mejores resultados en comparación con la CNN y se puede apreciar en las bases de datos BBCN y YSC. Los resultados sugieren la influencia de los mecanismos de autoatención para el aprendizaje de relaciones complejas entre las palabras del texto.

La evaluación de la solución propuesta en la detección de polaridad se realizó en el marco del Foro de Evaluación de las Lenguas Ibéricas-IberLEF 2022 (*Iberian Languages Evaluation Forum*), específicamente, en la tarea *Rest-Mex:Recommendation System, Sentiment Analysis and Covid Semaphore Prediction for Mexican Tourist Texts*.⁽⁵¹⁾ El conjunto de datos empleado está formado por una colección de opiniones de TripAdvisor, entre 2002 y 2021, sobre los lugares más representativos de México. La clase de cada opinión es un número entero entre 1 y 5, donde 1 representa la polaridad más negativa y 5 la más positiva. En la tabla 4 se muestra una comparación entre variantes del método propuesto usando RoBERTa y mT5, como arquitectura de redes profundas, con las 3 soluciones que mejores resultados obtuvieron en esa competición, de 26 alternativas evaluadas.⁽⁵¹⁾ Cabe destacar que la variante del método que combina AdamW con CMA-ES para entrenar el modelo RoBERTa alcanza el mejor resultado en la predicción de polaridad a partir de las métricas MAE y la exactitud.

Los resultados en la clasificación de opiniones muestran que la optimización de la última capa de la arquitectura basada en *transformer* mediante los EDA, incrementa la eficacia con respecto a AdamW. En ambos modelos preentrenados (mT5 y RoBERTa), se puede notar que los EDA multivariados (EMNA y CMA-ES), que consideran múltiples dependencias entre los parámetros, logran mejores resultados en comparación con UMDAc, el cual no asume dependencias entre los parámetros al construir un modelo de probabilístico univariado.⁽⁴¹⁾

Evaluación de la detección de emociones en opiniones de usuario con enfoque semántico de selección de características

La solución desarrollada fue evaluada en la clasificación de 7 emociones: enojo, miedo, tristeza, alegría, sorpresa, y repulsión, para ello se utilizan las colecciones de prueba *AIT-SemEval2018* y *TASS-2020*.^(52,53) En la tabla 5 se muestran los resultados obtenidos con la colección *AIT-SemEval2018*, a partir de diferentes variantes del método desarrolladas y estudiadas, que incluyen estrategias de poda. Se resaltan en negritas los mejores resultados para cada estrategia de poda.

Según se aprecia en la tabla 5 el mejor resultado se obtuvo aplicando un enfoque pesimista en el tratamiento del lexicón, la agregación difusa de las métricas de relación semántica para obtener las características afectivas, y el clasificador LR. Los resultados de esta variante superan de forma significativa el *baseline*. En sentido general, se demuestra con esta colección el aporte que representa la combinación de métricas con el operador de agregación difusa para tratar la semántica asociada al vocabulario afectivo del lexicón en la selección de características, respecto al uso de métricas de relación semántica individuales.

Se puede concluir que constituye una contribución científica el uso de conocimiento afectivo en el proceso de selección de características dentro de un esquema tradicional de clasificación, y que se mejoran los resultados cuando se aplican métricas de relación semántica, combinadas con un operador de agregación difusos, para tratar la semántica de ese vocabulario.

Conclusiones

Los métodos desarrollados representan contribuciones que arrojan resultados prometedores en el estado del arte para el análisis de opinión, siendo muy competitivos y en determinados escenarios superiores. Desde el punto de vista práctico, se brindan herramientas para la toma de decisiones

Tabla 4. Comparación de soluciones propuestas con resultados de Rest-Mex 2022

	Equipos	Final Rank	MAE	Exactitud
Rest-Mex 2022	UMU-Team-Run-1	0,8924 (I)	0,2584 (III)	75,98
	UC3M-Run1	0,8908 (III)	0,2607	76,25 (III)
	CIMAT MTY-GTO-Run1	0,8899	0,2643	75,78
Variantes propuestas	RoBERTa+AdamW+CMA-ES	0,8911 (II)	0,2568 (I)	76,70 (I)
	RoBERTa+AdamW	0,8906	0,2578 (II)	76,64 (II)
	mT5+AdamW+CMA-ES-Run2 (40)	0,8722	0,3050	74,53
	mT5+AdamW-Run1 (40)	0,8692	0,3053	73,69

Tabla 5. Síntesis de los resultados obtenidos sobre la colección *AIT-SemEval2018*

		Medidas	Clasificador	P	R	F1	Acc,
Enfoques de Poda	Optimista	WN-RES	MLP	0,817	0,775	0,779	0,775
		PMI	MLP	0,955	0,950	0,953	0,950
		WN-JCN	LR	0,821	0,780	0,784	0,780
		CHI	MLP	0,958	0,956	0,956	0,956
		Operador difuso	MLP	0,959	0,957	0,957	0,957
	Pesimista	WN-RES	LR	0,817	0,775	0,780	0,775
		PMI	MLP	0,950	0,945	0,945	0,950
		WN-JCN	LR	0,821	0,781	0,785	0,781
		CHI	LR	0,958	0,955	0,955	0,955
		Operador difuso	LR	0,960	0,958	0,958	0,958
	Fuzzy	WN-RES	LR	0,818	0,777	0,781	0,777
		PMI	MLP	0,948	0,945	0,945	0,950
		WN-JCN	MLP	0,825	0,780	0,785	0,780
		CHI	LR	0,958	0,956	0,956	0,956
		Operador difuso	MLP	0,959	0,957	0,957	0,956
	Sin podar	WN-RES	LR	0,824	0,782	0,786	0,782
		PMI	LR	0,956	0,953	0,953	
WN-JCN		MLP	0,819	0,780	0,784	0,780	
CHI		MLP	0,957	0,955	0,955	0,955	
Operador difuso		MLP	0,959	0,957	0,957	0,957	
TF-IDF + SVM (baseline)				0,33	0,31	0,31	0,34

con mayor eficacia para la detección de polaridad y sentimiento a partir de textos en lenguaje natural en un contexto donde cada día las personas usan las redes para expresar sus ideas. El desarrollo de métodos de entrenamiento para RNP mediante los EDA incrementa la capacidad de generalización de estos modelos en la clasificación de textos y opiniones. Se logran aplicaciones con alto valor agregado y potencialmente aplicables al estudio de las opiniones de la población, así como la detección de aspectos que más inciden en el estado de opinión.

REFERENCIAS BIBLIOGRÁFICAS

1. Quba GY, Al Qaisi H, Althunibat A, Alzu'Bi S. Software Requirements Classification using Machine Learning algorithm's. In: 2021 International Conference on Information Technology, ICIT 2021-Proceedings. 2021. 685-90 p.
2. Al Kilani N, Tailakh R, Hanani A. Automatic Classification of Apps Reviews for Requirement Engineering: Exploring the Customers Need from Healthcare Applications. In: 2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS). 2019. 541-8 p.
3. Kici D, Malik G, Cevik M, Parikh D, Basar A. A BERT-based transfer learning approach to text classification on software requirements specifications. Proc Can Conf Artif Intell. 2021;1-13.
4. Alias N, Foozy CFM, Ramli SN. Video spam comment features selection using machine learning techniques. Indones J Electr Eng Comput Sci. 2019;15(2):1046-53.
5. Ezpeleta E, Iturbe M, Garitano I, de Mendizabal IV, Zurutuza U. A mood analysis on youtube comments and a method for improved social spam detection. In: Lecture Notes in Computer Science. 2018. 514-25 p.
6. Gupta A, Bhatia R. Ensemble approach for web page classification. Multimed Tools Appl. 2021 jul 15;80(16):25219-40.
7. Mishev K, Gjorgjevikj A, Vodenska I, Chitkushev LT, Trajanov D. Evaluation of Sentiment Analysis in Finance: From Lexicons to Transformers. IEEE Access. 2020;8:131662-82.
8. Glazkova A, Egorov Y, Glazkov M. A Comparative Study of Feature Types for Age-Based Text Classification. In: van der Aalst WMP, Batagelj V, Ignatov DI, Khachay M, Koltsova O, Kutzov A, et al., editors. Analysis of Images, Social Networks and Texts. Cham: Springer International Publishing; 2021. 120-34 p.
9. Esteves B, Fraser K, Kulkarni S, Conlan O, Rodríguez-Doncel V. Extracting and Understanding Call-to-actions of Push-Notifica-

- tions. In: Rosso P, Basile V, Martínez R, Métais E, Meziane F, editors. *Natural Language Processing and Information Systems*. Cham: Springer International Publishing; 2022. 147-59 p.
10. Greco F, Polli A. Emotional Text Mining: Customer profiling in brand management. *Int J Inf Manage*. 2020;51.
 11. Guerrero-Rodríguez R, Álvarez-Carmona MÁ, Aranda R, López-Monroy AP. Studying Online Travel Reviews related to tourist attractions using NLP methods: the case of Guanajuato, Mexico. *Curr Issues Tour*. 2023 Jan 17;26(2):289-304.
 12. Grey EM. Opinion Mining With Hotel Review using Latent Dirichlet Allocation-Fuzzy C-Means Clustering (LDA-FCM). *Turkish J Comput Math*. 2021;12(11):3001-7.
 13. Frouzandeh S, Berahmand K, Nasiri E, Rostami, Mehrdad. A Hotel Recommender System for Tourists Using the Artificial Bee Colony Algorithm and Fuzzy TOPSIS Model: A Case Study of TripAdvisor. *Int J Inf Technol Decis Mak*. 2021;20(01):399-429.
 14. Zuheros C, Martínez-Cámara E, Herrera-Viedma E, Herrera F. Sentiment Analysis based Multi-Person Multi-criteria Decision Making methodology using natural language processing and deep learning for smarter decision aid. Case study of restaurant choice using Trip Advisor reviews. *Inf Fusion*. 2021;68(November 2020):22-36.
 15. Do HH, Prasad, Penatiyana WC, Maag A, Alsadoon A. Deep learning for aspect-based sentiment analysis: a comparative review. *Expert Syst Appl*. 2019;118:272-99.
 16. Ravi K, Ravi V. A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. *Knowledge-Based Syst*. 2015;89:14-46.
 17. Hernández AR, Simón-Cuevas A, Lorenzo MMG, Arco L, Serrano-Guerrero J. Towards Context-Aware Opinion Summarization for Monitoring Social Impact of News. *Information*. 2020;11(11).
 18. Hernández AR, Lorenzo MMG, Simón-Cuevas A, Arco L, Serrano-Guerrero J. A semantic approach for topic-based polarity detection: a case study in the Spanish language. *Procedia Comput Sci*. 2019;162:849-56.
 19. Montañez Castelo P, Simón-Cuevas A, Olivás JA, Romero FP. Enhancing Spanish Aspect-Based Sentiment Analysis Through Deep Learning Approach. In: *Lecture Notes in Computer Science (LNCS)* [Internet]. 2024. 215-24 p. Disponible en: https://link.springer.com/10.1007/978-3-031-49552-6_19
 20. Cañete J, Chaperon G, Fuentes R, Ho JH, Kang H, Pérez J. Spanish Pre-trained BERT Model and Evaluation Data. In: *Practical ML for Developing Countries Workshop at ICLR 2020*. 20AD. 1-9 p.
 21. De la Rosa J, Ponferrada EG, Villegas P, De Prado Salas PG, Romero M, Grandury M. BERTIN: Efficient Pre-Training of a Spanish Language Model using Perplexity Sampling. *Proces del Leng Nat*. 2022;68:13-23.
 22. Toledano-López OG, Madera J, González H, Simón-Cuevas A. A hybrid method based on estimation of distribution algorithms to train convolutional neural networks for text categorization. *Pattern Recognit Lett* [Internet]. 2022 ago;160:105-11. Disponible en: <https://linkinghub.elsevier.com/retrieve/pii/S0167865522002021>
 23. Kingma DP, Ba JL. Adam: A method for stochastic optimization. In: *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*. 2015. p. 1-15.
 24. Loshchilov I, Hutter F. Decoupled weight decay regularization. In: *7th International Conference on Learning Representations, ICLR 2019*. 2019. 1-8 p.
 25. Pennington J, Socher R, Manning CD. GloVe: Global Vectors for Word Representation. In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics; 2014. 1532-43 p.
 26. Xue L, Constant N, Roberts A, Kale M, Al-Rfou R, Siddhant A, et al. mT5: A Massively Multilingual Pre-trained Text-to-Text Transformer. In: *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 2021. 483-98 p.
 27. Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D, et al. RoBERTa: A Robustly Optimized BERT Pretraining Approach. In: *arXiv preprint*. 2019.
 28. Serrano-Guerrero J, Alshouha B, Romero FP, Olivás JA. Affective Knowledge-enhanced Emotion Detection in Arabic Language: A Comparative Study. *J Univers Comput Sci*. 2022;28(7):733-57.
 29. González Parets OM, Simón-Cuevas A, Olivás JA, Perea-Ortega JM. A Hybrid Approach for Spanish Emotion Recognition Applying Fuzzy Semantic Processing. In: *Lecture Notes in Computer Science (LNCS)* [Internet]. 2024. p. 225-36. Disponible en: https://link.springer.com/10.1007/978-3-031-49552-6_20
 30. González-Guerra H, Simón-Cuevas A, Perea Ortega JM, Olivás JA. Un enfoque semántico en la selección de características basadas en léxico para la detección de emociones. *Proces del Leng Nat*. 2021;67:115-26.
 31. Zimmermann HJ, Zysno P. Latent connectives in human decision making. *Fuzzy Sets Syst*. 1980;4(1):37-51.
 32. Pedersen T, Patwardhan S, Michelizzi J. WordNet: Similarity - measuring the relatedness of concepts. In: *HLT-NAACL 2004 - Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics, Demonstrations*. 2004. p. 38-41.
 33. Saggion H, Torres-Moreno JM, da Cunha I, SanJuan E, Velázquez-Morales P. Multilingual Summarization Evaluation without Human Models. In: *Coling 2010: Posters*. Beijing, China: Coling 2010 Organizing Committee; 2010. p. 1059-67.
 34. Pontiki M, Galanis D, Papageorgiou H, Androutsopoulos I, Manandhar S, Al-Smadi M, et al. SemEval-2016 Task 5: Aspect Based Sentiment Analysis. In: *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*. San Diego, California: Association for Computational Linguistics; 2016. p. 19-30.
 35. Devlin J, Chang MW, Lee K, Toutanova K. BERT: Pre-training of deep bidirectional transformers for language understanding. *NAACL HLT 2019 - 2019 Conf North Am Chapter Assoc Comput Linguist Hum Lang Technol - Proc Conf*. 2019;1(Mlm):4171-86.
 36. Modelos M, Armengol-Estapé J, Pàmies M, Llop-Palao J, Silveira-Ocampo J, Carrino CP, et al. MarIA: Spanish Language Models. *Proces del Leng Nat Rev*. 2022;(marzo):39-60.
 37. Henríquez Miranda C, Buelvas E. AspectSA: Unsupervised system for aspect-based sentiment analysis in Spanish. *Prospectiva* [Internet]. 2019 abr 4;17(1):87-95. Disponible en: <http://ojs.uac.edu.co/index.php/prospectiva/article/view/1961>
 38. Martínez-Seis BC, Pichardo-Lagunas O, Miranda S, Pérez-Cázarres IJ, Rodríguez-González JA. Deep learning approach for aspect-based sentiment analysis of restaurants reviews in Spanish. *Comput y Sist*. 2022;26(2):899-908.

39. Mar Rodríguez C, Montañez Castelo P, Simón Cuevas A. Análisis de sentimientos orientado al comercio electrónico de CITMATEL: proyecto de voz del cliente. *Rev Cuba Transform Digit.* 2023;4(1).
40. Toledano-López OG, Madera J, González H, Simón-Cuevas A, Demeester T, Mannens E. Fine-tuning mT5-based Transformer via CMA-ES for Sentiment Analysis. In: CEUR Workshop Proceedings [Internet]. 2022. Disponible en: <https://ceur-ws.org/Vol-3202/restmex-paper12.pdf>
41. Toledano-López OG, Álvarez-Carmona MÁ, Madera J, Simón-Cuevas A, López-Rodríguez YA, González Díez H. Polarity Prediction in Tourism Cuban Reviews Using Transformer with Estimation of Distribution Algorithms. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* [Internet]. Lecture Notes in Computer Science. Springer; 2024. p. 335-46. Disponible en: https://link.springer.com/10.1007/978-3-031-49552-6_29
42. Camacho-Collados J, Pilehvar MT. On the Role of Text Preprocessing in Neural Network Architectures: An Evaluation Study on Text Categorization and Sentiment Analysis. In: *EMNLP Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP* 2018. p. 40-6.
43. Li Y, Wang X, Xu P. Chinese text classification model based on deep learning. *Futur Internet.* 2018;10(11).
44. Yu H, Xiong J, Zhang X. Multi-view clustering by exploring complex mapping relationship between views. *Pattern Recognit Lett.* 2020;138:230-6.
45. Dong X, Wu D, Nie F, Wang R. Dependence-Guided Multi-View Clustering. In: *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2021. p. 3650-4.
46. Werner M, Laber E. Speeding up word moveras distance and its variants via properties of distances between embeddings. *Front Artif Intell Appl.* 2020;325:2204-11.
47. Hartmann J, Huppertz J, Schamp C, Heitmann M. Comparing automated text classification methods. *Int J Res Mark.* 2018;36(1):20-38.
48. Deb N, Jha V, Panjiyar AK, Gupta RK. A comparative analysis of news categorization using machine learning approaches. *Int J Sci Technol Res.* 2020;9(1):2469-72.
49. Samsudin NM, Mohd Foozy CFB, Alias N, Shamala P, Othman NF, Wan Din WIS. Youtube spam detection framework using naive bayes and logistic regression. *Indones J Electr Eng Comput Sci.* 2019;14(3):1508-17.
50. Skianis K, Nikolentzos G, Limnios S, Vazirgiannis M. Rep the Set: Neural Networks for Learning Set Representations. In: Chiappa, Silvia and Calandra R, editor. *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics. Proceedings of Machine Learning Research*; 2020. p. 1410-20.
51. Álvarez MÁ, Díaz Á, Aranda R, Rodríguez AY, Fajardo D, Guerrero R, *et al.* Overview of Rest-Mex at IberLEF 2022: Recommendation System, Sentiment Analysis and Covid Semaphore Prediction for Mexican Tourist Texts. *Proces del Leng Nat.* 2022;69.
52. Mohammad S, Bravo-Marquez F, Salameh M, Kiritchenko S. SemEval-2018 Task 1: Affect in Tweets. In: *Proceedings of the 12th International Workshop on Semantic Evaluation*. New Orleans, Louisiana: Association for Computational Linguistics; 2018. p. 1-17.
53. García-Vegaa M, Díaz-Galiano MC, García-Cumbreras M, Del Arco FMP, Montejo-Raéz A, Jiménez-Zafra SM, *et al.* Overview of TASS

2020: Introducing Emotion Detection. *CEUR Workshop Proc.* 2020;2664:163-70.

Recibido: 30/08/2024

Aprobado: 20/09/2024

Agradecimientos

Bashar Alshouha, Yoan Antonio López Rodríguez, Gabriel José Rivero Fernández, Ricardo Onel Alfonso Ayala, Harold González Guerra, Abel González Barrera, Erick Taylor Reverón, Miguel Ángel Álvarez Carmona, Raciél Yera Toledo, Luis Martínez López, Erick Mannens, Thomas Demeester, Carlos Mar Rodríguez, Anibal Hernández-González, Yoan Martínez-López, Ansel Rodríguez-González

Conflictos de intereses

Los autores declaran que no existen conflictos de intereses entre ellos, ni con la investigación presentada, ni con la institución que representa.

Contribuciones de los autores

Conceptualización: Alfredo Javier Simón Cuevas, Orlando Grabiél Toledano López, Julio Cesar Madera Quintana, Héctor Raúl González Díez, Alejandro Ariel Ramón Hernández, María Matilde García Lorenzo
Curación de datos: Alfredo Javier Simón Cuevas, Orlando Grabiél Toledano López, Alejandro Ariel Ramón Hernández, Oscar Miguel González Parets, Patricia Montañez Castelo

Análisis formal: Alfredo Javier Simón Cuevas, Julio Cesar Madera Quintana, Héctor Raúl González Díez, Jesús Serrano Guerrero, José Ángel Olivas Varela, Leticia Arco García, José Manuel Perea Ortega, Francisco P. Romero

Adquisición de fondos: Héctor Raúl González Díez, José Ángel Olivas Varela, José Manuel Perea Ortega

Investigación: Alfredo Javier Simón Cuevas, Orlando Grabiél Toledano López, Julio Cesar Madera Quintana, Héctor Raúl González Díez, Alejandro Ariel Ramón Hernández, María Matilde García Lorenzo, Jesús Serrano Guerrero, Leticia Arco García, Oscar Miguel González Parets, Patricia Montañez Castelo

Metodología: Alfredo Javier Simón Cuevas, Julio Cesar Madera Quintana, Héctor Raúl González Díez, María Matilde García Lorenzo, Jesús Serrano Guerrero, José Ángel Olivas Varela, Leticia Arco García, José Manuel Perea Ortega, Francisco P. Romero

Administración del proyecto: Alfredo Javier Simón Cuevas, Héctor Raúl González Díez

Recursos: Alfredo Javier Simón Cuevas, Julio Madera Quintana, Héctor Raúl González Díez

Software: Orlando Grabiél Toledano López, Alejandro Ariel Ramón Hernández, Oscar Miguel González Parets, Patricia Montañez Castelo

Supervisión: Alfredo Javier Simón Cuevas, Julio Cesar Madera Quintana, Héctor Raúl González Díez, María Matilde García Lorenzo

Validación: Orlando Grabiél Toledano López, Alejandro Ariel Ramón Hernández, Oscar Miguel González Parets, Patricia Montañez Castelo

Visualización: Orlando Grabiél Toledano López, Alejandro Ariel Ramón Hernández, Oscar Miguel González Parets, Patricia Montañez Castelo

Redacción original: Alfredo Javier Simón Cuevas, Orlando Grabiél Toledano López, Oscar Miguel González Parets, Patricia Montañez Castelo

Redacción-revisión y edición: Alfredo Javier Simón Cuevas, Orlando Grabiél Toledano López, Julio Cesar Madera Quintana, Héctor Raúl González Díez, María Matilde García Lorenzo, Jesús Serrano Guerrero, José Ángel Olivas Varela, Leticia Arco García, José Manuel Perea Ortega, Francisco P. Romero

Financiamientos

Los resultados científicos que se exponen aquí han sido financiados parcialmente por fondos de la Cooperación Belga al Desarrollo a través de VLIR-USO, Fondo Europeo de Desarrollo Regional (FEDER), la Junta de Extremadura (GR18135), y el Ministerio de Ciencia, Innovación y Universidades de España, a través los proyectos MERINET–TIN2016-76843-C4-2-R (AEI/FEDER, UE) y SAFER (PID2019-104735RBC42).

Cómo citar este artículo

Simón Cuevas AJ, Toledano López OG, Madera Quintana J, González Díez HR, Ramón Hernández AA, García Lorenzo MM *et al.* Contribuciones al análisis de sentimientos en opiniones de usuario aplicando técnicas y algoritmos de inteligencia artificial. An Acad Cienc Cuba [internet] 2024 [citado en día, mes y año];14(3):e1674. Disponible en: <http://www.revistaccuba.cu/index.php/revacc/article/view/1674>

El artículo se difunde en acceso abierto según los términos de una licencia Creative Commons de Atribución/Reconocimiento-NoComercial 4.0 Internacional (CC BY-NC-SA 4.0), que le atribuye la libertad de copiar, compartir, distribuir, exhibir o implementar sin permiso, salvo con las siguientes condiciones: reconocer a sus autores (atribución), indicar los cambios que haya realizado y no usar el material con fines comerciales (no comercial).[©] Los autores, 2024.

