

Un enfoque semántico en la selección de características basadas en léxico para la detección de emociones

A semantic approach in the lexicon-based feature selection for emotion detection

Harold González-Guerra¹, Alfredo Simón-Cuevas¹, José M. Perea-Ortega²,
José A. Olivas³

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

²Universidad de Extremadura, Badajoz, España

³Universidad de Castilla-La Mancha, Ciudad Real, España

hgonzalez@ceis.cujae.edu.cu, asimon@ceis.cujae.edu.cu, jmperea@unex.es,
joseangel.olivas@uclm.es

Resumen: La detección de emociones es una tarea del análisis de sentimientos que trata la extracción y el análisis de las emociones en textos. Reconocer emociones implícitas es uno de los principales desafíos en enfoques basados en palabras claves o lexicones. Este trabajo presenta un enfoque híbrido de detección de emociones, que combina la selección de características relevantes de emoción basada en un lexicon, con un enfoque clásico de aprendizaje para determinar la emoción. El proceso de selección de características propuesto se centra en capturar el significado emocional del texto mediante el cálculo de la relación semántica entre su contenido y el vocabulario del lexicon, con el objetivo de incrementar el reconocimiento de emociones implícitas. La solución propuesta fue evaluada en la clasificación de emociones en tweets en español incluidos en el corpus AIT, con diferentes alternativas para computar la relación semántica y varios algoritmos de clasificación, obteniéndose resultados muy prometedores.

Palabras clave: Detección de emociones, selección de características, medidas semánticas.

Abstract: Emotion detection is a task of sentiment analysis that deals with the extraction and analysis of emotions in texts. Recognizing implicit emotions is one of the main challenges in keyword or lexicon-based approaches. This paper presents a hybrid emotion detection approach, which combines lexicon-based emotion-relevant feature selection with a classical learning-based approach to determine the emotion. The proposed feature selection process focuses on capturing the emotional meaning of the text by computing the semantic relationship between its content and the lexicon vocabulary, with the goal of increasing implicit emotion recognition. The proposed solution was evaluated on the classification of emotions in Spanish tweets included in the AIT corpus, with different alternatives to compute the semantic relation and several classification algorithms, obtaining very promising results.

Keywords: Emotion detection, feature selection, semantic measures.

1 *Introducción*

Las emociones son rasgos básicos que nos caracterizan como humanos y que influyen en las acciones, los pensamientos y, por supuesto, en nuestra forma de comunicarnos. A pesar de no considerarse entidades propiamente

lingüísticas, las emociones se expresan a través del lenguaje por lo que, desde hace varios años, han sido estudiadas por investigadores de diferentes disciplinas como la psicología, la sociología, la medicina, o la informática (Ekman, 1992).

En los últimos años, la comunidad científica relacionada con el PLN ha mostrado especial interés en la detección de emociones en conversaciones textuales, ya que su investigación puede encontrar varias aplicaciones en el mundo digital actual (Strapparava, 2016; Mohammad et al., 2018; Chatterjee et al., 2019). Por ejemplo, en el ámbito de la atención al cliente, redes sociales como Twitter están ganando protagonismo y los clientes esperan respuestas rápidas. En caso de un gran flujo de tweets el tiempo de respuesta aumenta, por lo que si los tweets se pudieran priorizar según su contenido emocional, la satisfacción del cliente seguramente aumentaría. Por otro lado, en esta era de la mensajería instantánea, dado que los usuarios están constantemente enviando mensajes, podría ocurrir que enviaran mensajes de enfado inapropiados a otros usuarios. En estos casos, si se utilizara una aplicación de detección de emociones, se podrían tomar medidas como, por ejemplo, mostrar una advertencia al usuario antes de enviar el mensaje.

La tarea de la detección de emociones en textos presenta un importante desafío al carecer de la ayuda que, en cualquier comunicación visual, proporcionan las expresiones faciales y las modulaciones de voz. Además, el reto de detectar las emociones en un texto se ve agravado por la dificultad de comprender algunos aspectos relacionados con la comunicación, como pueden ser el contexto, el sarcasmo, la ambigüedad del propio lenguaje natural, o la creciente jerga que está provocando el uso masivo de aplicaciones de mensajería instantánea (Shivhare y Khethawat, 2012; Khan et al., 2016). En la literatura existen varios enfoques para abordar esta tarea. Uno de los más utilizados es el basado en reglas (Strapparava y Mihalcea, 2008; Sykora et al., 2013), que trata de explotar el uso de palabras clave y su coocurrencia con otras palabras que tienen asociado un determinado valor emocional o afectivo. Ese valor asociado a determinadas palabras del lenguaje suele establecerse a partir de diferentes recursos léxicos existentes, algunos muy conocidos como WordNet-Affect o SentiWordNet. Por esa razón, a los métodos que siguen este enfoque también se les conoce como métodos basados en palabras clave o en lexicón.

Más recientemente, se reportan acercamientos donde se combinan el uso de lexicón con modelos de aprendizaje automático, los

cuales son reconocidos dentro de los enfoques híbridos (Alswaidan y Menai, 2020). Disponer de un lexicón o palabras clave con determinado valor emocional o afectivo permite determinar con alta eficacia el estado emocional de los textos, sin embargo, se reconocen ciertas limitaciones en los enfoques basados en lexicón que están relacionadas con la cobertura del contenido textual (Hemmatian y Sohrabi, 2019; Chakriswaran et al., 2019; Acheampong, Wenyu, y Nunoo-Mensah, 2020). El procesamiento del vocabulario en un lexicón desde una perspectiva semántica permite incrementar su cobertura para identificar características emocionales de un texto, sin necesidad de aumentar el tamaño de dicho vocabulario. Sin embargo, esta orientación semántica del uso de los lexicones, por ejemplo, mediante cómputo de relaciones semánticas subyacente entre las palabras de emoción y el contenido textual a procesar, ha sido muy poco explotado. Precisamente, estudios realizados sobre soluciones reportadas en la literatura señalan como limitaciones el no uso de características semánticas en la detección de emociones en textos (Alswaidan y Menai, 2020).

En este trabajo se propone un método para la detección de emociones en textos cortos escritos en español (tweets), que se basa en la semántica de las palabras clave del texto. Nuestra hipótesis se centra en que si logramos determinar un buen grado de afinidad semántica del texto del tweet con cada emoción a detectar (relación semántica tweet-emoción), conseguiremos una selección de características más enfocada a las emociones, lo que supondría un mejor proceso de aprendizaje en el algoritmo encargado de la clasificación. Por tanto, la principal novedad que aporta este trabajo está relacionada con el enfoque propuesto para capturar el grado de afinidad semántica entre el contenido del tweet y cada uno de los vocabularios que caracterizan a cada emoción, lo que permite una selección de características de la opinión donde el análisis de relevancia esté más orientado al dominio del problema (emociones), en lugar de a aspectos estadísticos dentro del contenido, como comúnmente se suele adoptar. Como principal contribución destaca el enfoque semántico propuesto con el que se trata el vocabulario del lexicón y su relación con el contenido del texto. En general, las soluciones basadas en conocimiento (lexicones)

hacen uso de este tipo de recursos como si fueran una lista de palabras sin significado. Nuestra contribución aporta una solución al desafío que representa extender la cobertura de ese conocimiento en el procesamiento del contenido textual a clasificar y lograr identificar características del texto más relevantes para predecir las clases de emociones.

El resto del artículo se estructura de la siguiente manera: los principales trabajos relacionados se exponen en la Sección 2; la estructura general y los principios del método propuesto se describen en la Sección 3; los datos experimentales, la evaluación, y el análisis de los resultados obtenidos se muestran en la Sección 4; en la Sección 5 se resume el trabajo y se propone una dirección futura.

2 Trabajos relacionados

En la literatura se pueden encontrar diferentes enfoques para abordar la tarea de la detección de emociones, siendo incluso clasificados en varias categorías por distintos investigadores (Cambria, 2016; Gupta et al., 2017; Sailunaz et al., 2018; Acheampong, Wenyu, y Nunoo-Mensah, 2020). La clasificación más genérica es la propuesta por Cambria (2016), que establece dos categorías: los basados en reglas y los basados en aprendizaje automático. Al primero pertenecen aquellos métodos que hacen uso de recursos léxicos tales como lexicones, bolsas de palabras e incluso ontologías. El segundo engloba aquellos métodos que aplican algoritmos de aprendizaje basados en características lingüísticas.

En los **enfoques basados en reglas** algunos métodos explotan el uso de palabras clave en los textos y su coocurrencia con otras palabras clave con valor emocional/afectivo explícito (Strapparava y Mihalcea, 2008). Para ello, se utilizan varios recursos léxicos como WordNet-Affect (Strapparava y Valitutti, 2004) y SentiWordNet (Esuli y Sebastiani, 2006) para el idioma inglés. En cuanto a la disponibilidad de este tipo de recursos en otros idiomas, el número es muy limitado. Para el español, destacan el Spanish Emotion Lexicon (SEL) (Sidorov et al., 2012) y el Improved Spanish Opinion Lexicon (iSOL) (Molina-González et al., 2013). Otros métodos basados en este enfoque también tratan de explotar la sintaxis de las palabras clave mediante el uso de etiquetadores POS y, aunque suelen obtener buena precisión, sufren de una baja cobertura (*recall*) porque

muchos textos no contienen palabras afectivas a pesar de transmitir emociones (Gupta et al., 2017).

La mayoría de los trabajos presentados en el conocido foro de evaluación SemEval durante los últimos años (Mohammad et al., 2018; Chatterjee et al., 2019), utilizan léxicos de afecto y concluyen que son una fuente de información muy valiosa porque proporcionan información previa sobre el tipo de emoción asociada a cada palabra del texto. Además, en WASSA, otro foro de evaluación relacionado con estas tareas, también se demostró que el uso de características procedentes de léxicos de afecto es útil para tareas de minería de emociones (Mohammad y Bravo-Marquez, 2017). En este sentido, Bandhakavi et al. (2017) estudian el problema de la selección de características de emoción utilizando recursos léxicos específicos del dominio y léxicos de emoción de propósito general. Aún así, existen varias revisiones recientes que ponen de manifiesto las limitaciones asociadas al uso de la semántica que se encuentran en estos procesos para detectar o clasificar emociones en los textos (Acheampong, Wenyu, y Nunoo-Mensah, 2020; Alswaidan y Menai, 2020).

En los **enfoques basados en aprendizaje automático** (*Machine Learning*, ML) la mayoría de los métodos se basan en la extracción de características, tales como la presencia de n-gramas frecuentes, la negación, la puntuación, los emoticonos, los hashtags, etc., para así formar una representación de características del texto que luego se utiliza como entrada por los clasificadores para predecir la salida (Canales y Martínez-Barco, 2014; Liew y Turtle, 2016). Estos métodos suelen requerir un arduo proceso de selección de características y no logran una alta cobertura (*recall*) debido a las diversas formas de representar las emociones.

Dentro de los enfoques ML se han reportado métodos basados en redes neuronales profundas, que han tenido un éxito considerable en diversas tareas aplicadas a texto, habla e imagen. Variaciones de las redes neuronales recurrentes (RNN), como la LSTM y la BiLSTM, han sido eficaces para modelar información secuencial. Por su parte, la introducción de las redes neuronales convolucionales (CNN) en el dominio del texto ha demostrado su capacidad para clasificar características de las emociones (Mundra et

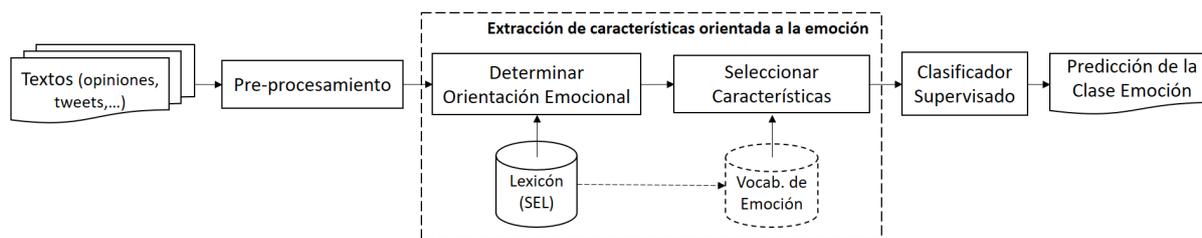


Figura 1: Flujo de trabajo de la solución propuesta.

al., 2017). A pesar de los buenos resultados, los enfoques basados en aprendizaje profundo presentan algunas desventajas respecto a modelos tradicionales de aprendizaje, lo que en algunos contextos pueden llevar a descartar su elección. Algunas de ellas son: demanda de mucha mayor cantidad de datos bien etiquetados para entrenamiento, mayores exigencias en la capacidad de cómputo, y por último, su naturaleza de funcionamiento tipo “caja negra”, que no permite una adecuada comprensión del proceso de aprendizaje (Kowsari et al., 2019).

El gran éxito de los métodos tradicionales basados en aprendizaje automático, junto con el hecho de que recursos léxicos como SEL aportan información muy valiosa sobre el tipo de emoción asociada a cada palabra, nos ha motivado a probar una combinación de ambos enfoques para detectar emociones. Partiendo del uso de SEL, proponemos calcular la relación semántica texto-emoción para así lograr una selección de características más enfocada a las emociones, lo que se supone redundará en un mejor aprendizaje del algoritmo de clasificación.

3 Solución propuesta

Esta sección describe la solución propuesta en este trabajo. Las principales etapas del proceso de desarrollo se describen en la Figura 1. En primer lugar, los textos de opinión son preprocesados y normalizados. A continuación, se lleva a cabo el proceso de determinar la orientación emocional de los textos, que consiste en establecer un *grado de afinidad* (relación semántica) de cada texto con cada vocabulario asociado a cada emoción del lexicón SEL. Finalmente, se realiza el proceso de selección de características, tomando como base el *grado de afinidad* obtenido en la fase anterior, y que permite generar los vectores característicos que serán utilizados por el algoritmo de aprendizaje supervisado.

3.1 Preprocesamiento

La naturaleza no estructurada de los textos, más aún en el caso de los tweets y textos de opinión, requirió la ejecución de tareas de preprocesamiento o normalización. Estas tareas consistieron en tokenizar los tweets usando NLTK TweetTokenizer¹, convertir todo el texto a minúscula y eliminar stop words, signos de puntuación y caracteres raros. No fue necesario ningún proceso de traducción previa porque dicha herramienta soporta trabajar con textos en castellano.

3.2 Extracción de características orientadas a la emoción

3.2.1 Determinar la orientación emocional

En esta tarea se determina la afinidad del contenido de los tweets a cada una de las emociones incluidas en el lexicón de emociones, desde el punto de vista semántico. Específicamente, el lexicón SEL (caso de estudio) contiene un total de 2.036 palabras en español que se organizan en 6 emociones diferentes: enojo (*anger*), miedo (*fear*), tristeza (*sadness*), alegría (*joy*), sorpresa (*surprise*), y repulsión (*disgust*). Por tanto, cada emoción está representada a través de un vocabulario de términos, donde cada término está etiquetado con un valor PFA (*Probability Factor of Affective*) (Sidorov et al., 2012). Ejemplos de palabras que se incluyen en cada categoría son: *amistad, bienestar, carcajada, celebrar...* (alegría), *enfadado, enfurecer, enrabiar, ira...* (enojo), *espeluznante, fobia, temor, terror...* (miedo), *asqueroso, detestable, inmundo, repugnante...* (repulsión), *asombroso, increíble, maravilloso, perplejo...* (sorpresa), *infeliz, luto, pena, pérdida...* (tristeza). La Figura 2 muestra el número de palabras por emoción que contiene SEL. Al determinar

¹<http://www.nltk.org/api/nltk.tokenize.html>

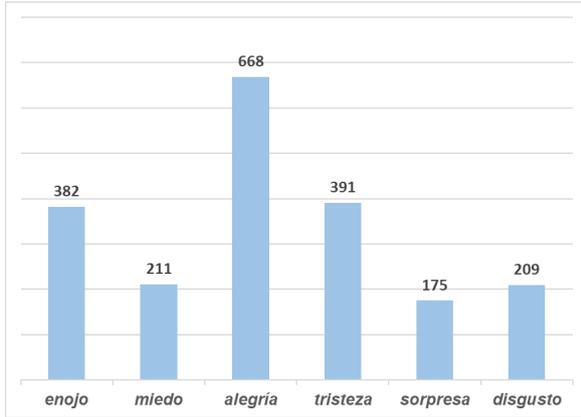


Figura 2: Estadísticas del lexicon SEL (# palabras por emoción).

esta afinidad (orientación) es posible realizar un proceso de selección de características de los tweets más enfocado en las emociones, lo cual no es posible en los modelos de relevancia basados en frecuencia, tales como TF o TF-IDF.

La determinación de la afinidad emocional se basa en el análisis de la relación semántica existente entre el contenido de los textos de opinión y cada emoción del lexicon, tomando como referencia su vocabulario. En este trabajo se proponen evaluar dos variantes para el cómputo de esa relación semántica opinión-emoción: (1) basada en el corpus de opiniones, y (2) basada en un recurso semántico externo. Entre las medidas basadas en corpus que han sido aplicadas en el ámbito del análisis de sentimientos se encuentran *Pointwise Mutual Information* (PMI) y *chi-square*, entre otras, siendo la primera la adoptada en esta propuesta. En el caso de la variante que se apoya en recursos semánticos, se utilizarán métricas basadas en WordNet (Pedersen, Patwardhan, y Michelizzi, 2004).

La relación semántica entre los tweets y las emociones del lexicon se representa mediante una *matriz de afinidad*, la cual es construida para cada tweet u opinión que se vaya a procesar. En esta matriz, las filas identifican cada una de las emociones del lexicon SEL, y las columnas las palabras de la opinión que constituyen características candidatas. La orientación emocional $O(w_i, E_j)$ entre cada palabra w_i de la opinión y la cantidad de términos del vocabulario de la emoción SEL presentes en la opinión E_j , es calculada según la Ecuación 1, y se almacena en la intersección término-emoción de la *matriz de*

afinidad. Esta matriz constituye la base para la siguiente fase (selección de características basada en la emoción) donde son seleccionadas las características más relevantes (las que más relación tengan con las emociones).

$$O(w_i, E_j) = \frac{\sum_{v_{i,j} \in E_j} rel_sem(w_i, v_{i,j})}{|E_j|} \quad (1)$$

Según se muestra en la Figura 2, en SEL existe un desbalance significativo en el tamaño de los vocabularios de cada emoción, lo que se traduce también en una mayor dispersión de valores PFA en las palabras que los componen, interpretándose este también como un valor de relevancia para la emoción. Todo ello indica que en un acercamiento basado en la relación semántica entre opinión-emoción, como el que se propone, se podría estar determinando esa relación semántica a partir de rangos de valores muy diferentes de PFA y, derivado de ello, originarse afectaciones en el cálculo de $O(w_i, E_j)$. Las emociones que más palabras tengan en su vocabulario tendrán más influencia en el cálculo de la orientación de la opinión, dado que hay más probabilidad de que los valores resultantes de este cálculo sean superiores. Una manera de reducir el efecto de este problema es mediante un mecanismo de poda o filtro del vocabulario de las emociones a partir de un umbral de relevancia mínima PFA_{min} .

3.2.2 Pointwise Mutual Information (PMI)

La medida PMI se deriva de la teoría de la información y proporciona una vía formal de modelar la información mutua entre las características (ej. palabras de la opinión) y las clases (ej. emociones a clasificar) (Aggarwal y Zhai, 2013). En el ámbito de la clasificación de textos, este tipo de medidas constituye una alternativa para evaluar la relevancia de características potenciales del texto con respecto a clases específicas. Esta información mutua puntual entre la palabra w_i y la clase de emoción E_j se define sobre la base de la coocurrencia entre las palabras $v_{i,j}$ del vocabulario de la emoción E_j y la palabra w_i de la opinión, y se calcularía según la Ecuación 2, siendo en este caso $rel_sem(w_i, v_{i,j}) = PMI(w_i, v_{i,j})$. Este planteamiento se basa en la suposición de que las palabras afectivas (incluidas en el lexicon) que coocurren con frecuencia con las palabras de la opinión tienden a estar semánticamente

relacionadas (Agrawal y An, 2012).

$$PMI(w_i, v_{i,j}) = \frac{coocurrencia(w_i, v_{i,j})}{ocurr(w_i) * ocurr(v_{i,j})} \quad (2)$$

3.2.3 Métricas basadas en WordNet (WN)

Las medidas de relación semántica basadas en WordNet constituyen otra de las alternativas viables para demostrar la hipótesis que ha motivado este trabajo. A diferencia del PMI, donde el cómputo se sustenta en el procesamiento solo del contenido de los textos, este tipo de medidas explotan la estructura topológica que forman los *synsets* y sus relaciones en WordNet, como recurso semántico externo. La relación semántica es un concepto que abarca la relación entre dos palabras, tanto por similitud entre significados, como por vínculos contextuales (ej. funcional, asociación, parte de, etc.) (Budanitsky y Hirst, 2006). Diversas métricas han sido definidas para medir ambos tipos de relaciones, cuyas implementaciones se ofrecen en el paquete de software libre WordNet::Similarity² (Pedersen, Patwardhan, y Michelizzi, 2004). La disponibilidad de estas métricas constituye una ventaja, dado que se podrían evaluar diferentes alternativas individuales y también combinarlas. No obstante, se reconoce que el uso de estas métricas tiene como limitante que las palabras que se comparan estén incluidas en algún *synset* en WordNet. En esta propuesta se ha adoptado para su evaluación la métrica JCN (Jiang y Conrath, 1997), siendo entonces $rel_sem(w_i, v_{i,j}) = JCN(w_i, v_{i,j})$. Resultados reportados en Budanitsky y Hirst (2006) muestran que, en un análisis experimental comparativo, la medida JCN es una de las que mejor se comporta.

3.2.4 Poda del vocabulario de las emociones

El valor PFA que posee cada palabra en el lexicón SEL sugiere cierta imprecisión y vaguedad subyacente en la relación palabra-clase de emoción, cuyo grado o pertenencia a la clase se expresa a través de dicho valor de PFA. Esta imprecisión se puede propagar hacia el resultado final afectando su calidad. En este sentido, se propone un mecanismo de poda o filtrado del vocabulario del lexicón que permite descartar el subconjunto de vocabulario

que mayor imprecisión posee, aspecto no considerado en otras soluciones reportadas.

El mecanismo propuesto de poda o filtrado del vocabulario de las emociones a partir del PFA_{min} permite reducir la dispersión en los valores de PFA de las palabras a considerar en el cómputo de la afinidad emocional de las opiniones, y llevar a cabo una selección de características guiada por un vocabulario de mayor relevancia para las emociones. De esta forma, también se reduce la carga computacional de este proceso, dado que se disminuiría la cantidad de interacciones en el cómputo de las relaciones semánticas, aspecto muy relevante cuando se utilizan medidas basadas en WordNet. En este sentido, se definieron y evaluaron tres criterios para obtener el umbral para la poda del vocabulario de las emociones en el lexicón:

1. Tomar como PFA_{min} el valor mínimo de la media de PFA calculada para cada una de las emociones del SEL (**enfoque optimista**).
2. Tomar como PFA_{min} el valor máximo de la media de PFA calculada para cada una de las emociones del SEL (más restrictivo que el anterior, **enfoque pesimista**).
3. Tomar como PFA_{min} el valor resultante de aplicar un operador de agregación compensatorio (**enfoque fuzzy** de la selección del umbral de poda), que permita obtener un único valor representativo de los valores medios de PFA obtenidos de los vocabularios de cada emoción (e_i), como es el caso del operador propuesto por Zimmermann y Zysno (1980) (Ecuación 3), donde γ es el grado de compensación proporcionado y se podría calcular según Ecuación 4 (Yager y Rybalov, 1998), en la que $T(e_1, e_2, \dots, e_n)$ es una función t-norma y se podría calcular según la Ecuación 5.

$$PFA_{min}(e_1, e_2, \dots, e_n) = \left(\prod_{i=1}^n e_i \right)^{1-\gamma} * \left(1 - \prod_{i=1}^n (1 - e_i) \right)^\gamma \quad (3)$$

$$\gamma = \frac{T(e_1, e_2, \dots, e_n)}{T(e_1, \dots, e_n) + T(1 - e_1, \dots, 1 - e_n)} \quad (4)$$

²<http://wn-similarity.sourceforge.net>

$$T(e_1, e_2, \dots, e_n) = \prod_{i=1}^n e_i \quad (5)$$

Luego de obtenido el umbral de poda (PFA_{min}), se filtra el vocabulario de cada emoción seleccionando solo aquellas palabras que tengan un $PFA_i \geq PFA_{min}$, y se lleva a cabo el análisis de la relación semántica entre cada una de las opiniones y las emociones del léxico (podadas).

3.3 Selección de características

En esta fase se seleccionan las características de las opiniones, tomando como base la *matriz de afinidad* construida en la fase anterior y con el objetivo de generar los vectores característicos que serán utilizados por los algoritmos de clasificación supervisados. Inicialmente, a partir la *matriz de afinidad* de cada opinión (o), se obtiene un *grado de afinidad semántica* (SAD, *Semantic Affinity Degree*) con respecto a cada una de las emociones (E_j), según la Ecuación 6. Luego, se selecciona la clase de emoción con la cual la opinión tiene mayor afinidad, siendo esta la que arroje un valor más alto de $SAD(o, E_j)$. La emoción sobre la que se exprese mayor afinidad será la que determine qué palabras afectivas (de las candidatas representadas) caracterizan la opinión (tweet).

$$SAD(o, E_j) = \frac{\sum_{i=1/w_i \in o}^n O(w_i, E_j)}{|o|} \quad (6)$$

A partir de identificar la emoción E_j con mayor $SAD(o, E_j)$, se filtra la *matriz de afinidad* de la opinión que se está procesando, eliminando las filas correspondientes a las emociones restantes. El vector característico de la opinión se construirá con las palabras de la opinión que posean $O(w_i, E_j) > 0$. En este enfoque se logra un proceso de selección de características donde la evaluación de su relevancia tiene más en cuenta la semántica alrededor de las emociones, a diferencia de otras propuestas donde el peso fundamental de la relevancia está en enfoques basados en la frecuencia (Plaza-del Arco et al., 2020).

Esta propuesta de selección propicia la reducción de características redundantes en la construcción de los vectores de las opiniones, dado que las características se determinan por una emoción en particular. También permite reducir las características no informativas

(o poco informativas) y que no tengan un alto poder discriminatorio, debido a la irrelevancia o redundancia con respecto a la clase (una misma característica es relevante en diferentes grados para varias clases). Todo ello propiciaría la mejora de los resultados del reconocimiento de emociones basado en un léxico.

Luego de identificadas las características de cada opinión, se procede a la última tarea para la construcción del vector de las opiniones referente al pesado de las características. El valor de peso de cada una de las características debe expresar un grado de relevancia de la misma, y como parte de este trabajo se estudiaron algunas alternativas de peso, tales como $O(w_i, E_j)$, como relevancia directa de la característica w_i en función de la emoción que determinó su selección; y la frecuencia de ocurrencia de esa palabra características dentro del corpus de opiniones, dado que es la alternativa más común en soluciones de clasificación supervisada de textos. Sin embargo, resultados parciales experimentales arrojaron que un modelo binario de representación del vector obtuvo mejores resultados que esas dos alternativas, confirmándose lo reportado en Agarwal y Mittal (2016), donde se plantea que, en el ámbito del análisis de sentimientos, este tipo de modelos ofrece mejores resultados que el basado en la frecuencia. En este sentido, se adoptó el modelo de representación binario para construir el vector característico de las opiniones, teniendo en cuenta que el peso de una característica w_i tiene valor 1 si el $O(w_i, E_j) \neq 0$, y valor 0 en caso contrario.

4 Resultados experimentales

La solución propuesta fue evaluada en la clasificación de cuatro emociones (*anger*, *fear*, *joy*, *sadness*) en tweets escritos en español, utilizando el corpus AIT empleado en *SemEval-2018 Task 1: Affect in Tweets* (subtarea *EI-oc*). A partir de ese corpus, se tomaron aleatoriamente 800 tweets por cada una de las emociones (3.200 tweets en total) para conformar el corpus de prueba. La Tabla 1 muestra la estadística del corpus de tweets utilizado durante la experimentación.

Los experimentos fueron realizados utilizando los algoritmos de clasificación *Support Vector Machine* (SVM), *Logistic Regression* (LR), *Multilayer Perceptron* (MLP) y *Naive Bayes* (NB), implementados en la librería

Emoción	#tweets	#palabras	media pal. por tweet
<i>enojo</i>	800	11.183	13,87
<i>miedo</i>	800	11.156	13,94
<i>alegría</i>	800	10.091	12,61
<i>tristeza</i>	800	10.967	13,70
Total	3.200	43.317	-

Tabla 1: Estadística del corpus de tweets utilizado durante la experimentación.

scikit-learn³ de Python. El 80 % (2.560) de los tweets incluidos en el corpus de prueba se utilizaron para el entrenamiento de los algoritmos de clasificación y el 20 % (640) para test. Los experimentos realizados pretendieron, por un lado, evaluar el comportamiento de las medidas de relación semántica seleccionadas y, por otro, evaluar el comportamiento de la poda del lexicón por relevancia del vocabulario (en sus varias alternativas).

En la evaluación de los resultados se utilizaron las métricas de *Precision* (P), *Recall* (R), y *F-score* (F1), y se definieron los resultados de la selección de características basada en la frecuencia de términos (TF) como caso base. En las Tablas 2-5, se muestran los resultados obtenidos con cada uno de los clasificadores, SVM, LR, MLP, y NB, respectivamente.

Los resultados obtenidos muestran que el clasificador que mejores resultados obtiene es SVM, al igual que los experimentos reportados en Plaza-del Arco et al. (2020). La Figura 3 muestra una comparativa de los mejores resultados F1 obtenidos por cada clasificador, independientemente de la medida utilizada y el criterio empleado para la obtención del PFA_{min} . En todas las pruebas realizadas, respecto a la alternativa más comúnmente usada, se obtienen mejores resultados cuando el proceso de selección de características está guiado solo por el análisis de relevancia a partir de la afinidad semántica de la opinión respecto a cada una de las emociones, utilizando como referencia el vocabulario que las representa en el lexicón SEL, lo que pone en evidencia la contribución del enfoque propuesto.

Entre las métricas empleadas para el cómputo de la relación semántica destaca PMI con respecto a la métrica JCN basada en WordNet, independientemente del clasificador y del tratamiento del lexicón (Figura 4). Una de las posibles causas del compor-

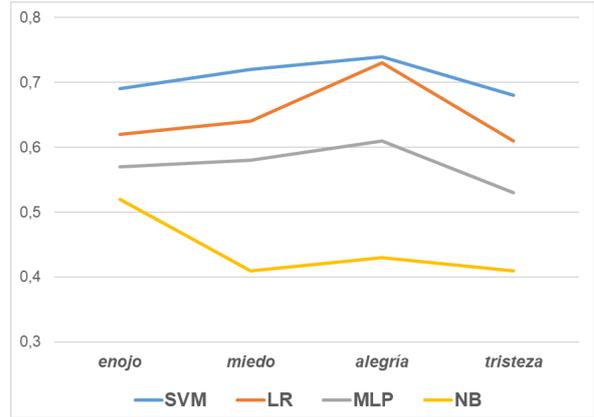


Figura 3: Comparativa de los mejores resultados F1 entre los clasificadores utilizados.

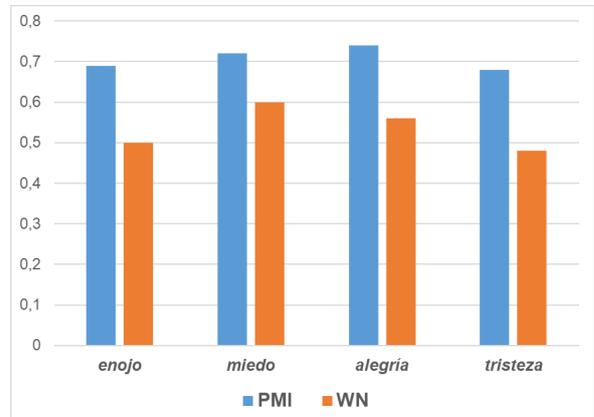


Figura 4: Comparativa de los mejores resultados F1 entre las medidas PMI y WN.

tamiento de la métrica basada en WordNet es la ausencia de palabras del lexicón o los tweets en ese recurso externo, lo cual es más probable en su versión en español. No obstante, considerando que los resultados en general son bastante cercanos, podría resultar una alternativa prometedora combinar estas y otras métricas en la evaluación de la afinidad emocional de las opiniones.

Cabe destacar también en estos resultados la contribución que representa trabajar con un vocabulario de emociones que tenga un mayor equilibrio en cuanto a grados de relevancia de los términos que integre, así como su cantidad de términos. La exactitud (medida *Acc*) alcanzada por todos los algoritmos de clasificación fue superior cuando los vocabularios del lexicón fueron podados a partir del umbral de relevancia mínimo (PFA_{min}), independientemente de los criterios de selección de este umbral y de la medida utilizada (Figura 5). El aporte al mejoramiento de cada

³<http://scikit-learn.org>

Lexicón	Medidas	<i>enojo</i>			<i>miedo</i>			<i>alegría</i>			<i>tristeza</i>			Acc	
		P	R	F1	P	R	F1	P	R	F1	P	R	F1		
Podado	Optimista	PMI	.62	.76	.68	.74	.70	.72	.80	.69	.74	.65	.62	.63	.69
		WN	.36	.66	.42	.60	.40	.48	.63	.51	.56	.51	.38	.43	.52
	Pesimista	PMI	.60	.73	.66	.67	.61	.64	.72	.70	.70	.65	.60	.62	.66
		WN	.40	.68	.50	.61	.50	.60	.64	.47	.54	.50	.39	.44	.52
	Fuzzy	PMI	.62	.77	.69	.74	.58	.65	.65	.66	.66	.69	.66	.68	.67
		WN	.38	.65	.48	.52	.37	.43	.65	.46	.53	.52	.45	.48	.51
Sin podar	PMI	.51	.62	.55	.53	.50	.51	.57	.51	.53	.48	.49	.48	.52	
	WN	.45	.38	.41	.48	.42	.44	.55	.44	.49	.36	.61	.45	.46	
TF (caso base)		.33	.35	.33	.35	.30	.32	.35	.36	.35	.34	.33	.33	.35	

Tabla 2: Resultados obtenidos con el clasificador SVM.

Lexicón	Medidas	<i>enojo</i>			<i>miedo</i>			<i>alegría</i>			<i>tristeza</i>			Acc	
		P	R	F1	P	R	F1	P	R	F1	P	R	F1		
Podado	Optimista	PMI	.53	.73	.61	.64	.65	.64	.83	.66	.73	.63	.51	.57	.64
		WN	.38	.69	.49	.57	.37	.45	.54	.53	.54	.45	.26	.33	.46
	Pesimista	PMI	.50	.71	.59	.65	.60	.62	.71	.68	.69	.71	.52	.60	.63
		WN	.37	.65	.47	.60	.45	.51	.51	.45	.48	.44	.28	.34	.48
	Fuzzy	PMI	.53	.75	.62	.61	.57	.59	.69	.61	.64	.69	.54	.61	.62
		WN	.38	.73	.50	.54	.33	.40	.47	.43	.45	.48	.28	.35	.44
Sin podar	PMI	.49	.65	.55	.53	.52	.52	.54	.57	.55	.68	.43	.53	.55	
	WN	.34	.61	.44	.51	.30	.37	.51	.46	.48	.45	.33	.38	.44	
TF (caso base)		.33	.32	.32	.32	.31	.31	.37	.34	.37	.34	.30	.32	.34	

Tabla 3: Resultados obtenidos con el clasificador LR.

Lexicón	Medidas	<i>enojo</i>			<i>miedo</i>			<i>alegría</i>			<i>tristeza</i>			Acc	
		P	R	F1	P	R	F1	P	R	F1	P	R	F1		
Podado	Optimista	PMI	.55	.51	.53	.56	.60	.58	.62	.53	.57	.45	.52	.48	.54
		WN	.38	.62	.47	.53	.39	.45	.55	.48	.51	.44	.36	.39	.46
	Pesimista	PMI	.55	.59	.57	.53	.54	.53	.59	.63	.61	.58	.49	.53	.56
		WN	.38	.64	.48	.55	.43	.48	.57	.44	.48	.41	.34	.37	.47
	Fuzzy	PMI	.56	.56	.56	.50	.42	.46	.57	.63	.60	.47	.50	.49	.53
		WN	.38	.62	.47	.57	.40	.47	.50	.47	.48	.51	.40	.44	.49
Sin podar	PMI	.49	.48	.48	.47	.34	.40	.51	.45	.47	.41	.42	.41	.47	
	WN	.37	.59	.46	.45	.37	.40	.57	.41	.47	.47	.38	.42	.45	
TF (caso base)		.26	.25	.28	.28	.29	.28	.35	.31	.33	.33	.32	.32	.31	

Tabla 4: Resultados obtenidos con el clasificador MLP.

Lexicón	Medidas	<i>enojo</i>			<i>miedo</i>			<i>alegría</i>			<i>tristeza</i>			Acc	
		P	R	F1	P	R	F1	P	R	F1	P	R	F1		
Podado	Optimista	PMI	.43	.63	.52	.51	.29	.37	.48	.34	.40	.37	.47	.41	.45
		WN	.34	.81	.48	.40	.24	.32	.45	.23	.34	.34	.30	.32	.38
	Pesimista	PMI	.42	.67	.51	.41	.25	.31	.40	.26	.31	.38	.43	.41	.40
		WN	.35	.81	.49	.30	.31	.30	.55	.20	.26	.36	.25	.29	.39
	Fuzzy	PMI	.42	.65	.51	.42	.23	.30	.46	.41	.43	.41	.41	.41	.43
		WN	.32	.80	.46	.60	.32	.41	.41	.12	.18	.43	.31	.36	.38
Sin podar	PMI	.32	.44	.32	.31	.21	.21	.23	.21	.21	.32	.34	.32	.31	
	WN	.32	.75	.44	.49	.29	.36	.37	.30	.33	.33	.25	.28	.37	
TF (caso base)		.29	.28	.28	.28	.29	.28	.34	.30	.32	.32	.28	.29	.31	

Tabla 5: Resultados obtenidos con el clasificador NB.

una de las métricas de evaluación de los criterios para la obtención del PFA_{min} evaluados muestra cierta dispersión, con influencia del clasificador y la métrica de relación semántica que se emplee. En general, los mejores resultados se concentran en los criterios *optimista*

y *fuzzy*, siendo el primero con el que se obtienen los mejores resultados combinando SVM y PMI.

Por último, cabe reseñar que, aunque los experimentos no se hayan realizado con el 100 % del corpus (63,4 % del total de tweets

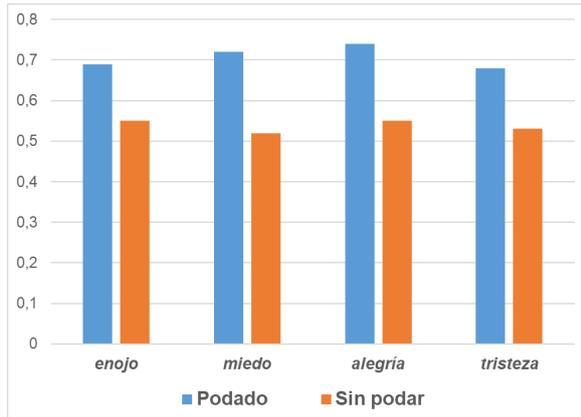


Figura 5: Comparativa de los mejores resultados F1 entre aplicar poda y sin podar.

de entrenamiento), los resultados obtenidos se consideran muy prometedores dado que, si establecemos una comparación con el mejor resultado de exactitud (Acc) obtenido en Plaza-del Arco et al. (2020) con el lexicon SEL y el clasificador SVM, que fue de 0,76, nuestra solución alcanza 0,69 con ese mismo clasificador pero con un 46 % menos de tweets de entrenamiento.

5 Conclusiones y trabajo futuro

En este trabajo se aborda la tarea de la detección de emociones en textos mediante un enfoque híbrido, que combina un proceso de selección de características basado en lexicon con un enfoque clásico de aprendizaje automático. La novedad de la propuesta radica en el enfoque semántico propuesto para capturar el grado de afinidad entre el contenido de opinión y el vocabulario que caracteriza cada emoción, de manera que se consigue una selección de características en las opiniones que es más adecuada para la tarea de clasificación objetivo.

Para evaluar la solución propuesta se llevaron a cabo diversos experimentos utilizando el corpus AIT de emociones en tweets en español. En los experimentos se evaluaron diferentes alternativas para calcular el grado de afinidad semántica texto-emoción y se probaron varios algoritmos de clasificación. Como principal conclusión, se considera que el enfoque propuesto es bastante prometedor a la hora de realizar una selección de características más adecuada, es decir, más enfocada en las emociones. Los buenos resultados obtenidos así lo avalan, habida cuenta que en la experimentación se utilizaron un 37 % menos

de tweets del total disponible en el corpus.

Como parte del trabajo futuro, además de aplicar el método propuesto a diferentes corpus de tweets, evaluar otras medidas de relación semántica y otros operadores de agregación, pretendemos ampliar este enfoque para entrenar modelos de clasificación que tengan en cuenta otra información lingüística del contexto de la opinión como, por ejemplo, la negación.

Agradecimientos

Este trabajo ha sido parcialmente financiado por el Fondo Europeo de Desarrollo Regional (FEDER), la Junta de Extremadura (GR18135), y el Ministerio de Ciencia, Innovación y Universidades de España, a través del proyecto SAFER (PID2019-104735RB-C42).

Bibliografía

- Acheampong, F. A., C. Wenyu, y H. Nunoo-Mensah. 2020. Text-based emotion detection: Advances, challenges, and opportunities. *Engineering Reports*, 2(7):1–24.
- Agarwal, B. y N. Mittal. 2016. Prominent Feature Extraction for Sentiment Analysis. *Prominent Feature Extraction for Sentiment Analysis*, i:21–45.
- Aggarwal, C. C. y C. X. Zhai. 2013. *Mining text data*, volumen 9781461432234. Springer.
- Agrawal, A. y A. An. 2012. Unsupervised emotion detection from text using semantic and syntactic relations. En *Proceedings - 2012 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2012*, páginas 346–353.
- Alswaidan, N. y M. E. B. Menai. 2020. A survey of state-of-the-art approaches for emotion recognition in text. *Knowledge and Information Systems*, 62(8):2937–2987.
- Bandhakavi, A., N. Wiratunga, D. Padmanabhan, y S. Massie. 2017. Lexicon based feature extraction for emotion text classification. *Pattern Recognition Letters*, 93:133–142.
- Budanitsky, A. y G. Hirst. 2006. Evaluating WordNet-based Measures of Lexical Semantic Relatedness. *Computational Linguistics*, 32(1):13–47.

- Cambria, E. 2016. Affective Computing and Sentiment Analysis. *IEEE Intelligent Systems*, 31(2):102–107.
- Canales, L. y P. Martínez-Barco. 2014. Emotion detection from text: A survey. En *Proceedings of the Workshop on Natural Language Processing in the 5th Information Systems Research Working Days (JISIC)*, páginas 37–43. Association for Computational Linguistics.
- Chakriswaran, P., D. R. Vincent, K. Srinivasan, V. Sharma, C.-Y. Chang, y D. G. Reina. 2019. Emotion ai-driven sentiment analysis: A survey, future research directions, and open issues. *Applied Sciences*, 9(24).
- Chatterjee, A., K. N. Narahari, M. Joshi, y P. Agrawal. 2019. SemEval-2019 Task 3: EmoContext Contextual Emotion Detection in Text. En *Proceedings of the 13th International Workshop on Semantic Evaluation (SemEval-2019)*, páginas 39–48. Association for Computational Linguistics.
- Ekman, P. 1992. An Argument for Basic Emotions. *Cognition and Emotion*, 6(3-4):169–200.
- Esuli, A. y F. Sebastiani. 2006. SentiWordNet: A Publicly Available Lexical Resource for Opinion Mining. *Proceedings of LREC 2006*, páginas 417–422.
- Gupta, U., A. Chatterjee, R. Srikanth, y P. Agrawal. 2017. A sentiment-and-semantics-based approach for emotion detection in textual conversations. *arXiv*.
- Hemmatian, F. y M. K. Sohrabi. 2019. A survey on classification techniques for opinion mining and sentiment analysis. *Artificial Intelligence Review*, 52(3):1495–1545.
- Jiang, J. J. y D. W. Conrath. 1997. Semantic similarity based on corpus statistics and lexical taxonomy. En *Proceedings of the 10th Research on Computational Linguistics International Conference*, páginas 19–33. The Association for Computational Linguistics and Chinese Language Processing (ACLCLP), Agosto.
- Khan, M. T., M. Durrani, A. Ali, I. Inayat, S. Khalid, y K. H. Khan. 2016. Sentiment analysis and the complex natural language. *Complex Adaptive Systems Modeling*, 4(1):2.
- Kowsari, K., K. Jafari Meimandi, M. Heidarysafa, S. Mendu, L. Barnes, y D. Brown. 2019. Text classification algorithms: A survey. *Information*, 10(4).
- Liew, J. S. Y. y H. R. Turtle. 2016. Exploring fine-grained emotion detection in tweets. En *Proceedings of the NAACL Student Research Workshop*, páginas 73–80. Association for Computational Linguistics.
- Mohammad, S. M. y F. Bravo-Marquez. 2017. Wassa-2017 shared task on emotion intensity.
- Mohammad, S. M., F. Bravo-Marquez, M. Salameh, y S. Kiritchenko. 2018. SemEval-2018 Task 1: Affect in Tweets. En *Proceedings of International Workshop on Semantic Evaluation (SemEval-2018)*, New Orleans, LA, USA.
- Molina-González, M. D., E. Martínez-Cámara, M. T. Martín-Valdivia, y J. M. Perea-Ortega. 2013. Semantic orientation for polarity classification in Spanish reviews. *Expert Systems with Applications*, 40(18):7250–7257.
- Mundra, S., A. Sen, M. Sinha, S. Mannarswamy, S. Dandapat, y S. Roy. 2017. Fine-grained emotion detection in contact center chat utterances. En *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volumen 10235 LNAI, páginas 337–349.
- Pedersen, T., S. Patwardhan, y J. Michelizzi. 2004. WordNet::Similarity - Measuring the relatedness of concepts. En *Proceedings of the National Conference on Artificial Intelligence*, páginas 1024–1025.
- Plaza-del Arco, F. M., M. T. Martín-Valdivia, L. A. Ureña-López, y R. Mitkov. 2020. Improved emotion recognition in Spanish social media through incorporation of lexical knowledge. *Future Generation Computer Systems*, 110:1000–1008.
- Sailunaz, K., M. Dhaliwal, J. Rokne, y R. Alhajj. 2018. Emotion detection from text and speech: a survey. *Social Network Analysis and Mining*, 8(1).

- Shivhare, S. N. y S. Khethawat. 2012. Emotion detection from text.
- Sidorov, G., S. Miranda-Jiménez, F. Viveros-Jiménez, A. Gelbukh, N. Castro-Sánchez, F. Velásquez, I. Díaz-Rangel, S. Suárez-Guerra, A. Treviño, y J. Gordon. 2012. Empirical study of opinion mining in Spanish tweets. *LNAI 7629*, páginas 1–14.
- Strapparava, C. y A. Valitutti. 2004. WordNet-Affect: an affective extension of WordNet. *Proceedings of the 4th International Conference on Language Resources and Evaluation*, páginas 1083–1086.
- Strapparava, C. 2016. Emotions and NLP: Future Directions. En *Proceedings of NAACL-HLT 2016*, página 180. Association for Computational Linguistics.
- Strapparava, C. y R. Mihalcea. 2008. Learning to identify emotions in text. *Proceedings of the 2008 ACM symposium on Applied computing - SAC '08*, páginas 1556–1560.
- Sykora, M. D., T. W. Jackson, A. O'Brien, y S. Elayan. 2013. Emotive ontology: Extracting fine-grained emotions from terse, informal messages. En *Proceedings of the IADIS International Conference Intelligent Systems and Agents 2013, ISA 2013, Proceedings of the IADIS European Conference on Data Mining 2013, ECDM 2013*.
- Yager, R. R. y A. Rybalov. 1998. Full reinforcement operators in aggregation techniques. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 28(6):757–769.
- Zimmermann, H. J. y P. Zysno. 1980. Latent connectives in human decision making. *Fuzzy Sets and Systems*, 4(1):37–51.