

SINAI en TASS 2018: Inserción de Conocimiento Emocional Externo a un Clasificador Lineal de Emociones

SINAI at TASS 2018: Lineal Classification System with Emotional External Knowledge

Flor Miriam-Plaza-del-Arco¹, Eugenio Martínez Cámara²,
M. Teresa Martín Valdivia¹, L. Alfonso Ureña-López¹

¹Departamento de Informática,

Centro de Estudios Avanzados en Tecnologías de la Información y de la Comunicación
Universidad de Jaén, Campus Las Lagunillas, 23071, Jaén, Spain

²Instituto Andaluz de Investigación en Ciencia de Datos e Inteligencia Computacional (DaSCI)
Universidad de Granada

{fmplaza, maite, laurena}@ujaen.es, emcamara@decsai.ugr.es

Resumen: Este artículo describe el sistema de clasificación de emociones desarrollado por el equipo SINAI en la Tarea 4 de la competición TASS 2018. Nuestro sistema se basa en un método supervisado con SVM utilizando características emocionales. Para ello, hacemos uso de distintos lexicones emocionales y realizamos la adaptación de las distintas emociones a la polaridad de las mismas. Los resultados obtenidos nos animan a seguir trabajando en este tipo de tareas.

Palabras clave: Análisis de Sentimientos, minería de emociones, lexicones emocionales, SVM

Abstract: In this paper, we present the emotion classification system developed by the SINAI team for the Task 4 at TASS 2018 workshop. Our approach is based on a supervised learning algorithm, SVM, using emotional features, which are grounded in several emotional lexicons. The obtained results encourage us to continue working on this line.

Keywords: Sentiment Analysis, emotion mining, emotional lexicons, SVM

1 Introducción

En este trabajo se describen las aportaciones realizadas a la Tarea 4 del taller TASS (Good Or Bad News? Emotional categorization of news articles) (Martínez-Cámara et al., 2018). La tarea propone el desarrollo de sistemas que identifiquen si una noticia es segura o adecuada para un anunciante en función de la emoción que suscite el titular de la misma. Debe destacarse que los titulares proceden de noticias publicadas en periódicos de España y de diversos países de América, de manera que el corpus que se proporciona, SANSE, es una representación global de la lengua española escrita¹.

Con la aparición de la Web 2.0, la cantidad de información subjetiva en Internet ha crecido exponencialmente. Existen diferentes plataformas como son las redes sociales, portales

de noticias, blogs o microblogs en los que los usuarios publican abiertamente. Si nos centramos en el dominio de las noticias, encontramos diversos periódicos digitales como El País, El Mundo o ABC que poseen artículos de este tipo. En ellos, los usuarios pueden comentar de forma anónima o pública, dando su opinión acerca de una determinada noticia. Debido al interés de analizar este tipo de información, surge la tarea de detectar automáticamente las emociones públicas evocadas por los documentos en línea. Sin embargo, anotar los titulares de noticias con emociones es una tarea difícil incluso para los seres humanos (Katz, Singleton, y Wicentowski, 2007).

El reconocimiento de emociones en textos tiene múltiples beneficios en diferentes sectores. Algunos ejemplos se listan a continuación:

1. Ayuda a las empresas a configurar sus

¹Los detalles de la tarea y el corpus pueden leerse en (Martínez-Cámara et al., 2018).

estrategias de marketing en función de las emociones de los consumidores (Bougie, Pieters, y Zeelenberg, 2003).

2. Permite mejorar los sistemas de recomendación basados en filtros colaborativos (Badaro et al., 2013).
3. Los políticos pueden adaptar su discurso en base a la reacción de la población (Pang, Lee, y others, 2008).
4. Puede ayudar a predecir el mercado de valores.
5. Puede identificar si el titular de una noticia es seguro o inseguro para un anunciante en función de las emociones expresadas en el mismo.

En este trabajo, proponemos un sistema de clasificación de emociones en español para el reconocimiento de emociones en titulares de noticias. Para ello, usamos diferentes lexicones afectivos para extraer una serie de características y, además, integramos dicho conocimiento en un clasificador supervisado. Los resultados obtenidos nos animan a seguir trabajando en esta tarea.

2 Minería de emociones

La minería de la opinión y la minería de la emoción forman parte del área de Análisis de Sentimientos (AS), pero tienen diferentes objetivos. La minería de la opinión trata el estudio de las opiniones expresadas en los textos y su tarea básica es la detección de la polaridad (Liu, 2015), mientras que la minería de la emoción se relaciona con el estudio de las emociones y su tarea básica es el reconocimiento de emociones (Cambria, 2016). La detección de la polaridad suele ser una tarea de clasificación binaria en la que se obtiene como resultado si una opinión es positiva o negativa, mientras que el reconocimiento de emociones trata de identificar diferentes categorías emocionales en un texto como el miedo, la alegría, la tristeza o la repulsión.

En la última década, la mayor parte del trabajo se ha centrado en la tarea de clasificación de la polaridad. Sin embargo, una de las áreas más complejas que aún no se ha estudiado en profundidad es la minería de la emoción. Algunos trabajos están comenzando a explorar el potencial de los sistemas de detección y clasificación de emociones (Mohammad, 2017). No obstante, la mayor parte de

los trabajos están orientados a tratar documentos en inglés, encontrando muy pocos trabajos en español.

La mayor parte de los trabajos de reconocimiento de emociones, se centran en el uso e integración de diferentes recursos como lexicones y corpus. En concreto, los lexicones afectivos son muy valiosos ya que proporcionan información acerca del tipo de emoción e intensidad que expresa cada palabra del texto. Además, pueden ser utilizados en varios enfoques del AS, como características para la clasificación en métodos de aprendizaje automático (Liu y Zhang, 2012), o para generar una puntuación de afecto para cada documento, basada en el valor de intensidad de cada palabra (Socher et al., 2013).

Sin embargo, la disponibilidad de recursos emocionales es escasa y la mayoría de ellos los encontramos en inglés. Por ejemplo, Strapparava y Valitutti (2004) desarrollaron el recurso WordNet-Affect formado por un conjunto de términos afectivos en inglés y basado en los synsets de WordNet. Linguistic Inquiry and Word Count (LIWC) es otro lexicon emocional que divide las palabras en diferentes categorías, incluidos los estados emocionales. Mohammad y Turney (2010) desarrollan un recurso denominado NRC Affect Intensity Lexicon el cual proporciona valores de intensidad de afecto para cuatro emociones básicas (*enfado, miedo, tristeza, alegría*). Por último Mohammad y Turney (2010) crean otro recurso denominado NRC Word-Emotion Association Lexicon (Emolex).

Con respecto a la disponibilidad de recursos emocionales en español, encontramos que el número es muy limitado. Hasta donde sabemos, el único lexicon emocional que podemos encontrar es Spanish Emotion Lexicon (SEL) (Sidorov et al., 2012) aunque los resultados obtenidos en diferentes experimentos no son muy prometedores.

En este trabajo, usamos los lexicones emocionales como características para la clasificación de noticias haciendo uso de un algoritmo de aprendizaje automático.

3 Descripción del sistema

El equipo SINAI ha participado con un sistema, que se describirá a continuación, en la Subtarea 1 de la Tarea 4 de la edición de 2018 del taller TASS.

En esta sección, se presentan los elementos que forman parte del sistema desarrollado

para la competición. En primer lugar, se describen los lexicones utilizados para la obtención de las características emocionales dadas en una noticia. En segundo lugar, se explica el proceso seguido para la extracción de las características haciendo uso de los lexicones y, por último, se explica la etapa de clasificación llevada a cabo con un algoritmo supervisado.

3.1 Lexicones de emoción y de polaridad

Para el reconocimiento de la emoción a nivel de oración, es esencial el análisis de las palabras del titular de la noticia para determinar su emoción. Por esta razón, en este sistema utilizamos dos lexicones de emoción y uno de polaridad con el fin de determinar la carga emocional que expresan las palabras del titular de una noticia. Los lexicones empleados son:

- **NRC Word-Emotion Association Lexicon (Emolex)** (Mohammad y Turney, 2010). Está constituido por 14.182 términos en inglés asociados a una o más emociones: ira, miedo, anticipación, confianza, sorpresa, tristeza, alegría. Además, este recurso está disponible en más de cien idiomas, incluido el español. Todas estas versiones se han generado mediante la traducción de los términos en inglés haciendo uso del traductor de Google.
- **Spanish Emotion Lexicon (SEL)** (Sidorov et al., 2012). Incluye 2.036 palabras en español que se asocian con la medida del Factor de Probabilidad de Uso Afectivo con respecto a al menos una emoción básica: alegría, ira, miedo, tristeza, sorpresa y repulsión. Fue anotado manualmente por 19 anotadores (escala: nula, baja, media, alta) y se implementaron ciertos umbrales de acuerdo.
- **Lexicón Mejorado de Opiniones en Español (iSOL²)** (Molina-González et al., 2013). Es una lista de palabras de opinión en español independiente del dominio. Los autores tradujeron automáticamente con el traductor Reverso³ el lexicón de Bing Liu (Hu y Liu, 2004), y posteriormente corrigieron manualmente

la traducción. iSOL contiene 2.509 palabras positivas y 5.626 negativas.

3.2 Extracción de características

Para la extracción de características emocionales de una noticia se ha tenido en cuenta la siguiente hipótesis: si una noticia genera una emoción positiva o neutra, es segura para incorporar anuncios, pero si genera una emoción negativa, la noticia no es segura para añadir anuncios. En base a esta hipótesis, se han utilizado los lexicones mencionados anteriormente mapeando cada emoción de las distintas palabras en polaridad positiva o negativa. Este mapeo se ha realizado en el caso de los lexicones que contienen palabras clasificadas en emociones: Emolex y SEL. Sin embargo, para el caso del lexicón iSOL no ha sido necesario ya que se trata de un lexicón de polaridad. Las Tablas 1 y 2 muestran la correspondencia realizada entre las emociones de los distintos lexicones y la polaridad (positiva o negativa).

Emoción	Polaridad
anger	negativa
fear	negativa
sadness	negativa
joy	positiva
disgust	negativa
surprise	positiva

Tabla 1: Correspondencia entre las emociones de SEL y la polaridad positiva/negativa

Emoción	Polaridad
enfado	negativa
miedo	negativa
tristeza	negativa
alegría	positiva
asco	negativa
sorpresa	positiva
anticipación	negativa
confianza	positiva

Tabla 2: Correspondencia entre las emociones de Emolex y la polaridad positiva/negativa

El método seguido para la extracción de características emocionales es el siguiente:

1. Se comprueba la presencia de cada uno

²<http://sinai.ujaen.es/isol/>

³<http://www.reverso.net/>

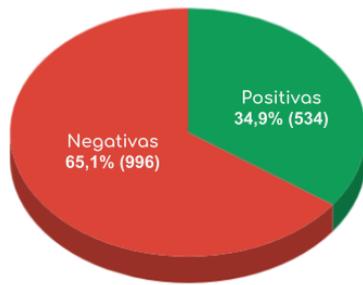


Figura 1: Cobertura del número de palabras de la noticia detectadas en los lexicones

de los términos que componen el titular de la noticia en los distintos lexicones.

2. Si el término está presente en algún lexicon, se mapea la emoción de dicho término y se obtiene la polaridad positiva o negativa.
3. Para cada titular de noticia, obtenemos un valor de polaridad positiva y otro de polaridad negativa. Este valor se obtiene realizando el sumatorio de la polaridad de cada una de las palabras del titular presente en los distintos lexicones (Emo-lex, SEL e iSOL).
4. Como resultado, obtenemos un vector de dos valores (positivo, negativo) para cada lexicon, es decir, extraemos un total de seis características para cada noticia.

En la Tabla 3 podemos ver un ejemplo de algunas de las noticias junto a las emociones reconocidas por el sistema.

3.3 Aprendizaje supervisado

En primer lugar, preprocesamos el corpus de noticias proporcionado. El preprocesamiento que llevamos a cabo se explica a continuación:

- Tokenización con TweetTokenizer⁴ de NLTK.
- Eliminación de palabras vacías (*stop-words*) con NLTK.⁵
- Conversión de las palabras a minúscula.

Seguidamente, se realiza la unión de los vectores obtenidos con los distintos lexicones de emociones y de polaridad, y la representación TF-IDF de la noticia. Para llevar a cabo la clasificación, optamos por el algoritmo

⁴<http://www.nltk.org/api/nltk.tokenize.html>

⁵<https://pythonspot.com/nltk-stop-words/>

de aprendizaje supervisado SVM, y concretamente su implementación C-SVC. Se usó un núcleo lineal y el valor de C fue 1.0.

Por último, para ejecutar nuestros experimentos, se utilizaron los conjuntos de datos proporcionados por los organizadores de la tarea de la siguiente manera. Durante el período de pre-evaluación, entrenamos el sistema con el conjunto de entrenamiento (*train*) y evaluamos con el conjunto de desarrollo (*dev.*). Sin embargo, durante el período de evaluación, entrenamos el sistema con el conjunto *train* y *dev.*, y evaluamos con el conjunto de evaluación (*test*).

4 Resultados obtenidos

Los resultados proporcionados por nuestro sistema sobre el corpus SANSE para la Tarea 4 se muestran en la Tabla 3 y 4. Las medidas oficiales de la competición son Macro-Precisión, Macro-Recall, Macro-F1 y Accuracy.

En la Tabla 5 podemos observar los resultados obtenidos por clase (SAFE, UNSAFE). Como podemos apreciar, el modelo muestra mejores resultados para la clase UNSAFE. Sin embargo, para la clase SAFE obtenemos un recall bajo de 62,2%. Esto se debe a que en el mapeado realizado (emoción-polaridad), el número de emociones negativas es mayor que el número de emociones positivas como se puede observar en la Tabla 1 y 2 y, por consiguiente, hay más palabras negativas que positivas en los lexicones emocionales. En la Figura 1 vemos claramente que el número de palabras negativas encontradas en la noticia (65,1%) es mayor que el número de palabras positivas (34,9%). Por tanto, se recuperan pocas palabras con polaridad positiva. Por otra parte, como podemos ver en la Tabla 4 los resultados obtenidos a nivel general por nuestro sistema son de 73,33% en el caso de la Macro-Precisión, 72,22% de Macro-Recall, 72,8% de Macro-F1 y 74,2% de Accuracy.

5 Conclusiones

En este trabajo presentamos un método de clasificación de emociones desarrollado con aprendizaje supervisado que hace uso de características emocionales extraídas a partir de diferentes lexicones y que se utilizan como entrada para un clasificador supervisado. Los resultados obtenidos demuestran la utilidad de utilizar características emocionales para

Titular noticia	Emociones positivas	Emociones negativas
Obesidad _[tristeza, asco] , un problema _[tristeza, miedo] de gravedad _[tristeza, miedo] mayor _[tristeza, confianza] en México	1	3
Secretario de estado comprometido _[alegría, confianza, anticipación] con la paz _[alegría, confianza, anticipación] en Bolivia	2	1
Detenido un yihadista que quería matar _[enfado, miedo, tristeza, anticipación] a Españoles	0	4

Tabla 3: Emociones detectadas en una noticia

Macro-P	Macro-R	Macro-F1	Acc.
0.733	0.722	0.728	0.742

Tabla 4: Resultados obtenidos sobre el corpus SANSE

Categoría	Precisión	Recall	F1
SAFE	0.702	0.622	0.660
UNSAFE	0.764	0.823	0.792

Tabla 5: Evaluación categórica

realizar la clasificación ya que son aceptables. Además, vemos que es necesario el desarrollo de lexicones emocionales en español ya que el único que encontramos actualmente es SEL. Nuestro próximo estudio se enfocará en explorar más lexicones afectivos en otros idiomas con el objetivo de adaptarlos al español ya que como se demostró en la tarea compartida del WASSA-2017 (Mohammad y Bravo-Marquez, 2017), el uso de lexicones es beneficioso para la clasificación de emociones.

Agradecimientos

Este trabajo ha sido parcialmente subvencionado por el Fondo Europeo de Desarrollo Regional (FEDER), el proyecto REDES (TIN2015-65136-C2-1-R) y el proyecto SMART-DASCI (TIN2017-89517-P) del Gobierno de España. Eugenio Martínez Cámara fue financiado por el programa Juan de la Cierva Formación (FJCI-2016-28353) del Gobierno de España.

Bibliografía

Badaro, G., H. Hajj, W. El-Hajj, y L. Nachman. 2013. A hybrid approach with collaborative filtering for recommender systems. En *Wireless Communications and*

Mobile Computing Conference (IWCMC), 2013 9th International, páginas 349–354. IEEE.

Bougie, R., R. Pieters, y M. Zeelenberg. 2003. Angry customers don't come back, they get back: The experience and behavioral implications of anger and dissatisfaction in services. *Journal of the Academy of Marketing Science*, 31(4):377–393.

Cambria, E. 2016. Affective computing and sentiment analysis. *IEEE Intelligent Systems*, 31(2):102–107.

Hu, M. y B. Liu. 2004. Mining and summarizing customer reviews. En *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, páginas 168–177. ACM.

Katz, P., M. Singleton, y R. Wicentowski. 2007. Swat-mp: the semeval-2007 systems for task 5 and task 14. En *Proceedings of the 4th international workshop on semantic evaluations*, páginas 308–313. Association for Computational Linguistics.

Liu, B. 2015. *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge University Press.

Liu, B. y L. Zhang. 2012. A survey of opinion mining and sentiment analysis. En *Mining text data*. Springer, páginas 415–463.

Martínez-Cámara, E., Y. Almeida-Cruz, M. C. Díaz-Galiano, S. Estévez-Velarde, M. A. García-Cumbreras, M. García-Vega, Y. Gutiérrez, A. Montejo Ráez, A. Montoyo, R. Muñoz, A. Piad-Morffis, y J. Villena-Román. 2018. Overview of TASS 2018: Opinions, health and emotions. En E. Martínez-Cámara Y. Almeida-Cruz M. C. Díaz-Galiano

- S. Estévez-Velarde M. A. García-Cumbreras M. García-Vega Y. Gutiérrez A. Montejo Ráez A. Montoyo R. Muñoz A. Piad-Morffis, y J. Villena-Román, editores, *Proceedings of TASS 2018: Workshop on Semantic Analysis at SEPLN (TASS 2018)*, volumen 2172 de *CEUR Workshop Proceedings*, Sevilla, Spain, September. CEUR-WS.
- Mohammad, S. M. 2017. Word affect intensities. *arXiv preprint arXiv:1704.08798*.
- Mohammad, S. M. y F. Bravo-Marquez. 2017. Wassa-2017 shared task on emotion intensity. *arXiv preprint arXiv:1708.03700*.
- Mohammad, S. M. y P. D. Turney. 2010. Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. En *Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text*, páginas 26–34. Association for Computational Linguistics.
- 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.
- Pang, B., L. Lee, y others. 2008. Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, 2(1–2):1–135.
- Sidorov, G., S. Miranda-Jiménez, F. Viveros-Jiménez, A. Gelbukh, N. Castro-Sánchez, F. Velásquez, I. Díaz-Rangel, S. Suárez-Guerra, A. Trevino, y J. Gordon. 2012. Empirical study of machine learning based approach for opinion mining in tweets. En *Mexican international conference on Artificial intelligence*, páginas 1–14. Springer.
- Socher, R., A. Perelygin, J. Wu, J. Chuang, C. D. Manning, A. Ng, y C. Potts. 2013. Recursive deep models for semantic compositionality over a sentiment treebank. En *Proceedings of the 2013 conference on empirical methods in natural language processing*, páginas 1631–1642.
- Strapparava, C. y A. Valitutti. 2004. Wordnet affect: an affective extension of wordnet. En *Lrec*, volumen 4, páginas 1083–1086. Citeseer.