

Metodología semi-automática para la anotación de corpus emocionales

Semi-automatic Methodology to Annotate Emotional Corpora

Lea Canales

Departamento de Lenguajes y Sistemas Informáticos
Universidad de Alicante
lcanales@dlsi.ua.es

Resumen: En este proyecto de tesis nos planteamos el desarrollo de una nueva metodología semi-automática que nos permita abarcar uno de los grandes retos de la tarea de detección de emociones: la creación de corpus emocionales. El desarrollo de corpus emocionales que nos permitan entrenar sistemas de reconocimiento de emociones es crucial para la creación de herramientas que nos permitan analizar y evaluar el estado emocional de los miembros de una sociedad. El análisis emocional resulta de gran importancia, ya que actualmente renombrados organismos ya realizan estos análisis a través de la interpretación de encuestas, con distintas finalidades: políticas, sociales, comerciales o educativas. Por tanto, el desarrollo de nuevos recursos y metodologías que nos permitan mejorar estos análisis automáticamente resultan de gran utilidad para este campo de investigación. Por ello, en este proyecto de tesis proponemos una metodología semi-automática con el objetivo de mejorar tanto en la calidad de los corpus como en la reducción del tiempo y coste necesario para su desarrollo.

Palabras clave: Detección de emociones, Anotación Semi-automática, Corpus Emocionales, Análisis de Sentimientos

Abstract: The objective of this PhD is to analyse and develop a new semi-automatic methodology with the aim of tackling one of the most important challenges in textual emotion detection: emotional annotation. The development of emotional corpora is crucial for improving the current emotional recognition systems, thus allowing us to analyse and evaluate people's emotions more precisely. This information and the patterns of these emotions provide a great value for politics, society, businesses or education. Therefore, the development of new resources and methodologies will allow us to improve the current emotion detection systems. Accordingly, this thesis proposes a semi-automatic methodology with the aim of improving the quality of emotional corpora and overcoming the cost and time-consuming shortcoming of manual annotation.

Keywords: Emotion detection, Semi-automatic Annotation, Emotional Corpora, Sentiment Analysis

1 Motivación

Actualmente, en Lingüística Computacional ha aumentado el interés de la detección de emociones a partir de texto escrito, debido al gran número de beneficios obtenidos a nivel empresarial, educativo, político o social.

En este campo de investigación, existen diferentes aproximaciones de sistemas de reconocimiento de emociones, pero la mayoría de ellas han sido llevadas a cabo mediante aprendizaje automático supervisado. El entrena-

miento de estos algoritmos implica la anotación de grandes cantidades de datos con categorías emocionales. Siendo este uno de los retos más importantes en este campo de investigación, ya que no es una tarea trivial debido a que la detección de emociones en texto escrito puede ser un reto incluso para los seres humanos.

La mayoría de investigación llevada a cabo hasta el momento ha demostrado los retos existentes en la anotación de corpus emo-

cionales, como por ejemplo, la obtención de acuerdo entre los anotadores o el tiempo y el coste asociado para su desarrollo. Esto es debido al hecho de que la anotación manual puede estar influenciada por un conjunto de factores, como la claridad de la guía de anotación, la dificultad de la tarea o el entrenamiento de los anotadores (Mohammad, 2016).

Como consecuencia, el número de corpus emociones disponibles en el estado de la cuestión es reducido y, además, el tamaño de los mismos es pequeño si deseamos aplicar técnicas de Aprendizaje Automático (*Machine Learning*) o Aprendizaje Profundo (*Deep Learning*).

Es por ello, que en este proyecto de tesis nos planteamos el desarrollo de una nueva metodología semi-automática para la anotación de emociones, que nos permita crear corpus emocionales en diferentes géneros, dominios e idiomas de manera eficiente, con el objetivo de crear recursos y herramientas que nos permitan analizar y evaluar el estado emocional de los miembros de una sociedad.

La metodología propuesta está compuesta de dos fases principales: 1) una pre-anotación automática de un conjunto reducido de las categorías emocionales más predominantes en cada frase; y 2) una anotación manual, en la que los anotadores seleccionarán la emoción dominante de cada frase entre las pre-seleccionadas en la primera fase.

En la siguiente sección vamos a describir los corpus emocionales más relevantes en el campo de detección de emociones, así como trabajos previos donde se ha utilizado la técnica de pre-anotación en diferentes tareas de Procesamiento de Lenguaje Natural (PLN).

2 Antecedentes y trabajos relacionados

Nuestra investigación se lleva a cabo desde el área de la Informática Afectiva o *Affective Computing* y más concretamente en su modalidad textual.

Como hemos comentado en el apartado anterior, las aproximaciones más utilizadas en esta modalidad son basadas en algoritmos de aprendizaje automático supervisado. Con el objetivo de crear estos sistemas en diferentes dominios y géneros, varios corpus emocionales anotados manualmente han sido desarrollados, como por ejemplo 185 historias

infantiles anotadas con categorías emocionales (Alm, Roth, y Sproat, 2005); entradas de blogs coleccionadas directamente de la Web, anotadas con categorías emocionales y la intensidad de cada emoción (Aman y Szpakowicz, 2007); o titulares de noticias anotados con las categorías emociones y la valencia para la tarea Affective Text del Semeval-2007 (Strapparava y Mihalcea, 2007). Todos estos corpus tienen en común un gran coste temporal y económico necesario para su desarrollo y además las dificultades para obtener acuerdo entre los anotadores, dificultando la obtención de un corpus de calidad.

El coste temporal y económico del desarrollo de un corpus es un reto compartido en muchas tareas de PLN, y la pre-anotación es una de las tareas que se ha llevado a cabo para abordarlo. La pre-anotación ha sido ampliamente estudiada en otras tareas de PLN, como POS-tagging o Reconocimiento de Entidades Nombradas, proporcionando mejoras en tiempo y calidad de los corpus desarrollados.

Marcus, Marcinkiewicz, y Santorini (1993) presentaron una de las primeras aproximaciones en las que se evaluaba el proceso de pre-anotación para la tarea de POS-tagging. Este experimento demostró que la anotación manual sin pre-anotación requirió el doble de tiempo y se obtuvo aproximadamente el doble de tasa de desacuerdo con respecto a la tarea con pre-anotación. También Fort y Sagot (2010) aplicaron la pre-anotación al POS-tagging y confirmaron la mejora en tiempo y calidad de la anotación, pero además los experimentos llevados a cabo demostraron que incluso cuando la pre-anotación no era precisa también ayudaba a mejorar la anotación manual. Más recientemente, Lingren et al. (2014) evalúan el impacto de la pre-anotación en el Reconocimiento de Entidades Nombradas del dominio médico, concluyendo que el anotador que tenía el texto pre-anotado necesitó menos tiempo para anotar que el anotador sin texto pre-anotado. Además, la pre-anotación no redujo el acuerdo o rendimiento del anotador.

Teniendo en cuenta este contexto y considerando los retos planteados en la tarea de anotación de emociones, una nueva metodología semi-automática es propuesta en este proyecto de tesis con el objetivo de mejorar tanto la calidad como el tiempo y coste asociado con el desarrollo de recursos emociona-

les, que nos permitan la anotación de corpus emocionales a gran escala y con estándares de fiabilidad.

3 Propuesta de investigación

La hipótesis de partida de este proyecto de tesis es que la reducción automática de categorías emocionales reducirá la complejidad de la tarea de anotación y por tanto la ambigüedad a la que se tienen que enfrentar los anotadores, proponiéndoles un número reducido de categorías emocionales. También los trabajos previos que demuestran las mejoras obtenidas con pre-anotación en otras tareas de PLN y el estudio de Antoine, Villaneau, y Lefeuivre (2014) donde concluyen que el acuerdo entre anotadores incrementa significativamente cuando el número de clases propuestas decrece, nos permiten considerar la viabilidad de nuestra hipótesis de partida.

Por ello, se propone una metodología semi-automática que está compuesta de dos fases principales:

- *Fase 1 - Pre-anotación automática:* proceso automático donde cada frase es pre-anotada con un número reducido de categorías emocionales.
- *Fase 2 - Anotación manual:* proceso manual donde anotadores humanos determinarán cual es la emoción dominante entre las pre-seleccionadas en la primera fase.

Nuestro objetivo es lograr una alta fiabilidad en la tarea de anotación de emociones reduciendo su complejidad automáticamente, y a su vez reducir el tiempo y el coste asociado a dicha tarea. El desarrollo de esta metodología nos permitirá la anotación de grandes cantidades de datos de manera eficiente y con garantía de estándares de fiabilidad.

4 Metodología propuesta

Como hemos mencionado en la sección anterior, la metodología que proponemos para mejorar el proceso de anotación de emociones es una metodología semi-automática compuesta de dos fases principales.

La primera fase consiste en un proceso de pre-anotación en que se pre-selecciona aquellas emociones más relacionadas con cada una de las frases. Concretamente, se pre-anotan la mitad del número de categorías emocionales con las que estemos

trabajando. Es decir, si nuestro objetivo es anotar con las seis emociones básicas de Ekman (1992), el sistema pre-anotará las frases con tres de las seis categorías. Por tanto, la metodología es adaptable a diferentes conjuntos de categorías emocionales, como por ejemplo el conjunto propuesto por Plutchik (1980) donde las emociones propuestas como básicas son: IRA, ASCO, MIEDO, FELICIDAD, TRISTEZA, SORPRESA, CONFIANZA y ANTICIPACIÓN.

Hemos desarrollado dos procesos de pre-anotación: un proceso no supervisado basado en modelos de Semántica Distribucional (*Distributional Semantic Models*) (Canales et al., 2017) y otro supervisado basado en Aprendizaje Automático.

La aproximación no supervisada esta basada en semántica distribucional y está compuesta de dos pasos principales:

1. La representación de las emociones y las frases en un espacio semántico.
2. La asociación entre frases y emociones.

En el primer paso, las frases y emociones son representados en forma de vectores distribucionales (*word embedding*). Este fase se divide en dos pasos: 1) el pre-procesamiento de las frases y las emociones, donde se construyen bolsas de palabras con vocabulario relacionado con cada una de las emociones y de las frases que deseamos anotar; y 2) la creación de un vector distribucional por cada una de las emociones y las frases, sustituyendo y sumando los vectores distribucionales de las palabras contenidas en las bolsas de palabras creadas en el paso anterior. Para la creación de las bolsas de palabras de las emociones es necesario un lexícón emocional y en nuestro caso, EmoSenticNet (Poria et al., 2013) es el recurso utilizado. Una representación gráfica de este primer paso se observa en la Figura 1.

En el segundo paso, se lleva a cabo la asociación entre frases y emociones calculando la similitud semántica entre vectores, asociado aquellas emociones cuya similitud semántica con cada frase es mayor. Con el objetivo de mejorar el conjunto de emociones pre-seleccionadas, en este proceso también se utiliza la información de polaridad y subjetividad de cada una de las frases, utilizando para ello la herramienta *Pattern* (De Smedt y Daelemans, 2012).

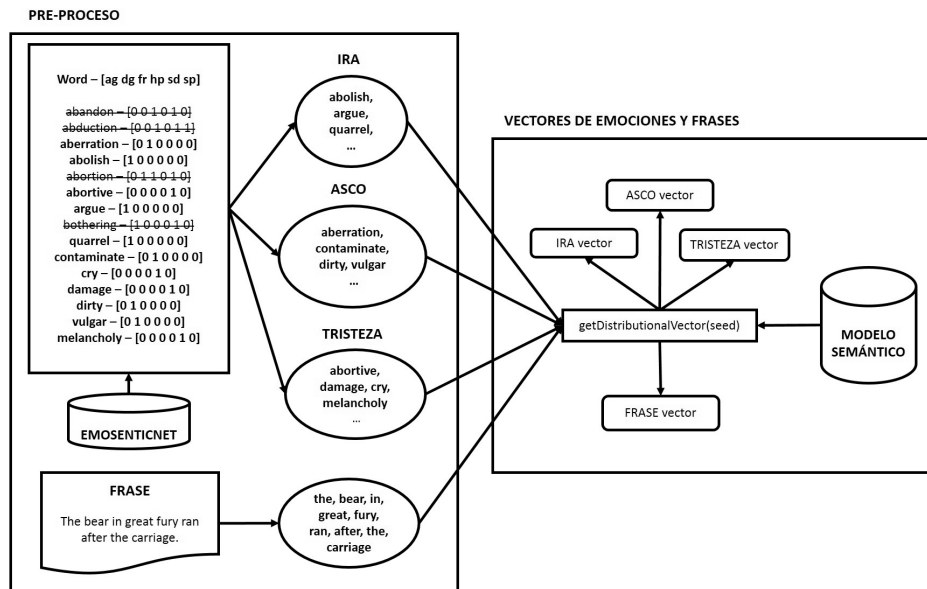


Figura 1: La representación de las emociones y las frases en el espacio semántico.

Respecto al proceso supervisado basado en Aprendizaje Automático, hemos utilizado corpus desarrollados hasta el momento en el campo de detección de emociones para crear modelos que nos permitan pre-annotar textos nuevos. En estas aproximaciones de Aprendizaje Automático, hemos utilizado diferentes conjuntos de características con el objetivo de evaluar cual de ellos nos proporcionaba mejores resultados.

Ambos procesos de pre-annotación, tanto el supervisado como el no supervisado, han sido evaluados en dos corpus emocionales: Aman corpus (Aman y Szpakowicz, 2007) y EmoTweet-28 (Liew, Turtle, y Liddy, 2016) permitiéndonos verificar la viabilidad del proceso en corpus de diferentes géneros. En el proceso de evaluación, se ha considerado que la pre-annotación era correcta si entre las emociones propuestas por el sistema se encontraba la emoción anotada en el *gold standard* de cada uno de los corpus. La evaluación muestra que la aproximación supervisada obtiene mejores resultados, pero la usabilidad de los métodos debe ser validada en la segunda fase de la metodología: la anotación manual a partir de la pre-annotación.

Por ello, el siguiente paso es el desarrollo de la anotación manual utilizando la pre-annotación. Es un proceso que ya hemos diseñado y que será llevado a cabo en las próximas semanas. En esta experimentación, nuestro objetivo es analizar las ventajas o mejo-

ras que nos pueda aportar la pre-annotación, tanto en la calidad del corpus como en la reducción del tiempo y coste necesario para su desarrollo.

5 Cuestiones de investigación

Este proyecto de tesis se enmarca dentro del campo de Informática Afectiva y Análisis de Sentimiento, lo que convierte la tarea en un reto, dado que tenemos que abarcar problemas como la ironía o la subjetividad. Considerando, por tanto, esta situación, surgen algunas cuestiones cuya puesta en común resultaría interesante:

- ¿Qué conjunto de características podemos incorporar al proceso de pre-annotación basado en Aprendizaje Automático para mejorar dicho proceso?
- ¿Qué otras técnicas no supervisadas podríamos aplicar para mejorar el proceso de pre-annotación?
- ¿Qué otras técnicas podríamos utilizar para alcanzar nuestro objetivo? ¿Podría ser *Active Learning* una buena técnica para abarcar nuestro proyecto?
- Estudios psicológicos muestran que las personas expresan sus emociones según su cultura e idioma. Por tanto, ¿es necesario crear recursos emocionales teniendo en cuenta el idioma? o ¿una simple traducción entre recursos es suficiente?

Agradecimientos

Quiero agradecer al programa de Formación de Personal Investigador (FPI) del Ministerio de Economía y Competitividad del Gobierno de España por su apoyo a través de una de sus becas pre-doctorales de investigación (BES-2013-065950) y por su apoyo a través del proyecto REDES (TIN2015-65136-C02-2-R) y RESCATA (TIN2015-65100-R). También, quiero agradecer el apoyo de la Generalitat Valenciana a través del proyecto PROMETEOII/2014/00, la Universidad de Alicante con el proyecto GRE16-01: “Plataforma inteligente para recuperación, análisis y representación de la información generada por usuarios en Internet” y a las ayudas de la Fundación BBVA para equipos de investigación científica 2016 a través del proyecto “Análisis de Sentimientos Aplicado a la Prevención del Suicidio en las Redes Sociales” (ASAP).

Bibliografía

- Alm, C. O., D. Roth, y R. Sproat. 2005. Emotions from text: Machine learning for text-based emotion prediction. En *Proceedings of the conference on HLT-EMNLP*, páginas 579–586.
- Aman, S. y S. Szpakowicz. 2007. Identifying Expressions of Emotion in Text. En *Text, Speech and Dialogue*. páginas 196–205.
- Antoine, J.-y., J. Villaneau, y A. Lefeuve. 2014. Weighted Krippendorff’s alpha is a more reliable metrics for multi-coders ordinal annotations: experimental studies on emotion, opinion and coreference annotation. En *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics*, número 1, páginas 550–559.
- Canales, L., W. Daelemans, E. Boldrini, y P. Martínez-Barco. 2017. Towards the improvement of automatic emotion pre-annotation with polarity and subjective information. En *Proceedings of the 11th biennial Recent Advances in Natural Language Processing conference (RANLP 2017)*.
- De Smedt, T. y W. Daelemans. 2012. Pattern for python. *J. Mach. Learn. Res.*, 13(1):2063–2067, Junio.
- Ekman, P. 1992. An argument for basic emotions. *Cognition and Emotion*, páginas 169–200.
- Fort, K. y B. Sagot. 2010. Influence of pre-annotation on pos-tagged corpus development. En *Proceedings of the Fourth Linguistic Annotation Workshop, LAW IV ’10*, páginas 56–63, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Liew, J. S. Y., H. R. Turtle, y E. D. Liddy. 2016. EmoTweet-28: A Fine-Grained Emotion Corpus for Sentiment Analysis. En *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*.
- Lingren, T., L. Deleger, K. Molnar, H. Zhai, J. Meinzen-Derr, M. Kaiser, L. Stoutenborough, Q. Li, y I. Solti. 2014. Evaluating the impact of pre-annotation on annotation speed and potential bias: natural language processing gold standard development for clinical named entity recognition in clinical trial announcements. *Journal of the American Medical Informatics Association : JAMIA*, 21(3):406–413.
- Marcus, M. P., M. A. Marcinkiewicz, y B. Santorini. 1993. Building a large annotated corpus of english: The penn treebank. *Comput. Linguist.*, 19(2):313–330, Junio.
- Mohammad, S. M. 2016. Sentiment Analysis: Detecting Valence, Emotions, and Other Affectual States from Text. En H. Meiselman, editor, *Emotion Measurement*. Elsevier.
- Plutchik, R. 1980. A general psychoevolutionary theory of emotion. En *Theories of Emotion*. páginas 3–33.
- Poria, S., A. Gelbukh, A. Hussain, N. Howard, D. Das, y S. Bandyopadhyay. 2013. Enhanced sentinet with affective labels for concept-based opinion mining. *IEEE Intelligent Systems*, 28(2):2–9.
- Strapparava, C. y R. Mihalcea. 2007. Semeval-2007 task 14: Affective text. En *Proceedings of the 4th International Workshop on Semantic Evaluations*, páginas 70–74.