

DETECCIÓN AUTOMÁTICA DE EMOCIONES EN TEXTOS DE REDES SOCIALES MEDIANTE ALGORITMOS INTELIGENTES



Instituto de Investigación y Transferencia en Tecnología (ITT)

Expositor: Lic. ESNAOLA, Leonardo

Grupo de Trabajo: J. P. Tessore, B. Cicerchia, H. Ramón, C. Russo, M. Sarobe

<http://www.itt.unnoba.edu.ar>

contacto@itt.unnoba.edu.ar

INTRODUCCIÓN

El procesamiento del lenguaje natural (PLN) o minería de textos es un área de la inteligencia artificial, la misma se encarga de desarrollar algoritmos que permitan extraer información relevante a partir de diversos contenidos en forma de texto. Con el auge de las redes sociales, la generación de este tipo de contenido ha crecido en forma exponencial, creando la materia prima necesaria para aplicar algoritmos de minería de textos que posibiliten extraer patrones significativos. Dentro de las tareas de la minería de textos podemos mencionar: clasificación de subjetividad; determinación de connotación positiva o negativa; detección de emociones; resolución de ambigüedades; extracción de palabras de opinión y/o aspectos, etc.

Este trabajo pretende construir, en primer lugar, una base de datos lo suficientemente representativa, a partir del contenido extraído de las redes sociales, para luego realizar sobre el mismo distintas tareas de minería de textos que permitan analizarlos con el objetivo de construir un procesador automático de emociones. Esta herramienta debe ser capaz de realizar una detección automática de las emociones presentes en un texto, basándose para ello en el análisis previo de las emociones manifestadas por las personas en redes sociales como Facebook, cuando depositan un comentario sobre alguna publicación determinada.

Las aplicaciones potenciales de una herramienta de este tipo pueden ser muchas, como detectar la emoción reflejada a partir de la opinión de grandes grupos de personas con respecto a determinados productos, servicios¹, o también por ejemplo, políticas públicas; puede utilizarse para identificar necesidades insatisfechas o quejas de los ciudadanos; también podría aplicarse en seguridad, a la detección automática de factores de riesgo en redes sociales, estos pueden ser por ejemplo amenazas, acosos o casos relacionados con el *cyberbulling*². Máxime, la mayoría de los trabajos en esta área han sido realizados para textos en idioma inglés, consecuentemente existen muchas oportunidades de investigación y desarrollo para el idioma español.

¹ J. Bollen, H. Mao, X. Zeng, Twitter mood predicts the stock market, Journal of Computational Science 2 (2012) 1-8.

² N. Sintaha, S. Bin Satter, N. Zawad, C. Swarnaker & A. Hassan (2016). Cyberbullying Detection Using Sentiment Analysis In Social Media. Brac University.

DESARROLLO

Para cumplir con el objetivo propuesto y poder desarrollar esta herramienta, se elaboró un proceso cuyas etapas se introducen en la figura siguiente (ver Figura 1).

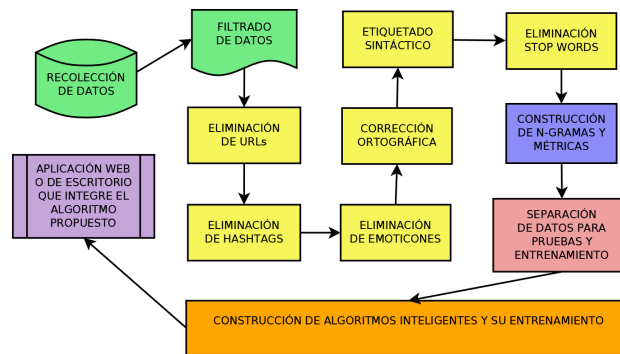


Figura 1. Etapas del proceso de construcción de la herramienta

A continuación, se introduce sucintamente, el trabajo realizado en las distintas etapas identificadas.

Recolección de datos. Esta etapa involucra la recolección de datos de una red social. Se ha seleccionado para ello Facebook, una plataforma que a enero del 2018, según un estudio³, cuenta con más de 2.167 millones de usuarios activos de todas partes del mundo.

En primer lugar se eligieron trece portales de noticias populares del país, con presencia en la red social, correspondientes a: Clarín, La Nación, Página 12, El Cronista, Ámbito Financiero, Todo Noticias, Crónica, CNN en Español, C5N, Agencia Télam, Deportivo Olé, Teleshov e Infobae, portales en los que se publican artículos de diversa índole, ya que algunos de estos medios se especializan en economía, otros en espectáculos, otros en interés general, etc.

Mediante la herramienta *Facebook API Graph*⁴ se recopilaron los comentarios y reacciones de las personas en diferentes publicaciones de los portales seleccionados. Se han reunido alrededor de 800.000 (ochocientos mil) comentarios etiquetados con su correspondiente emoción, almacenados ahora en una base de datos relacional.

Filtrado de datos. La etapa de recolección de datos descrita anteriormente captura todos los comentarios y todas las reacciones de las personas en las diferentes publicaciones, pero no todas las personas que depositan un comentario manifiestan una reacción; consecuentemente es necesario realizar una asociación entre el comentario y la reacción. Esta etapa de filtrado se realiza mediante una consulta a la Base de Datos que tiene como salida solo aquellos comentarios que tienen asociada una reacción.

Eliminación de URLs. Muchos de los comentarios incluyen *URLs*; estos elementos no tienen un verdadero valor analizable o vinculable a una emoción; consecuentemente se ha decidido quitarlos de los

³ Digital in 2018: World's internet users pass the 4 billion mark - We Are Social. (2018, January 30). Retrieved from <https://wearesocial.com/blog/2018/01/global-digital-report-2018>.

⁴ Facebook, "Facebook api." [Online]. Available: <http://developers.facebook.com>.

comentarios para facilitar el procesamiento del texto mediante expresiones regulares que permiten detectar si el fragmento de texto analizado es en efecto una *URL*.

Eliminación de Hashtags. Otra de las particularidades de los textos provenientes de redes sociales es la inclusión de *Hashtags*, cadenas de caracteres formadas por una o varias palabras concatenadas y precedidas por el carácter numeral (#). Los Hashtags permiten agrupar distintos comentarios vinculándolos a un tópico o tema determinado. En una primera instancia se había decidido “repararlos”, haciendo una corrección ortográfica para intentar recuperar el texto a partir del *hashtag*, pero todos los correctores ortográficos probados presentaban problemas con los caracteres “especiales” de nuestro idioma, como la “ñ” o la diéresis, las palabras acentuadas o aquellas escritas en lunfardo, por ejemplo “guita”, todas ellas válidas. Así, el proceso se hacía extremadamente complejo para considerar todas las posibilidades. Máxime, los comentarios que utilizaban *hashtags* representaban menos del 0.8% del total, motivo por el cual la decisión fue también eliminarlos.

Eliminación de emoticones. Los emoticones son elementos que también están muy presentes en los textos provenientes de redes sociales. Éstos reflejan en sí mismos una emoción, estado de ánimo, o hacen referencia a una situación u objeto. Dado que el interés de este trabajo se remite al análisis de textos, también se ha decidido eliminarlos de los comentarios con el propósito de facilitar el proceso. Para ello se ha utilizado una base de datos de los emoticones en texto plano, como por ejemplo: “:-)”, representando una cara que sonríe y, tras iterar sobre cada uno de los comentarios, se remueven sus ocurrencias. Otros emoticones más complejos, como por ejemplo “☺”, se representan mediante caracteres *Unicode*, los cuales se eliminaron a partir su reconocimiento a través de expresiones regulares.

Corrección ortográfica. Este proceso no es trivial cuando se analiza contenido proveniente de las redes sociales, principalmente por los motivos que se expresan a continuación. La mayoría de las herramientas de corrección automática ofrecen un conjunto de sugerencias ante la presencia de una palabra incorrecta. Dichas sugerencias, sin embargo no cuentan con algún nivel de confianza o ponderación que permita automáticamente decidir cuál es la palabra más adecuada para efectuar el reemplazo. Por otro lado, muchas de las palabras que conforman los comentarios, en general, no son reconocidas por las herramientas de corrección automática pues provienen del lenguaje informal. En Argentina se utilizan muchas palabras del lunfardo; hay palabras que, no estando correctamente escritas, son de uso popular y deberían ser tenidas en cuenta en cualquier estudio que involucre analizar contenido proveniente de redes sociales. Surge, en consecuencia, la necesidad de crear un diccionario enriquecido que incorpore variantes utilizadas en estos medios y que no están presentes en el lenguaje formal. Actualmente se está trabajando en esta etapa.

Etiquetado sintáctico. Concluida la etapa de corrección ortográfica es necesario realizar el etiquetado sintáctico, que permite reconocer la estructura de los comentarios, identificando sustantivos, verbos, adverbios, adjetivos, etc. Esta etapa está fuertemente vinculada al proceso anterior, dado que si se propone un diccionario de texto enriquecido que incorpore el lunfardo y palabras de uso popular frecuentemente utilizadas, será necesario también entrenar al etiquetador sintáctico con este nuevo corpus. Actualmente se utiliza la herramienta *NLTK*⁵, de software libre, entrenada previamente con un corpus de textos en idioma español extraído de Wikipedia.

⁵ Nltk.org. (2017). Natural Language Toolkit - NLTK 3.2.5 documentation. [online] Available at: <http://www.nltk.org/> [Accessed 28 Nov. 2017].

Eliminación de Stop Words. Otra etapa importante en el PLN consiste en la eliminación de aquellas palabras que suelen denominarse “vacías”, dado que las mismas no poseen un significado concreto: artículos, pronombres, preposiciones, etc.

No hay una lista definitiva de palabras vacías que todas las herramientas de procesamiento de lenguaje natural incorporen.

Construcción de N-gramas y Métricas. Los n-gramas nos permiten detectar palabras o secuencias de ellas, para por ejemplo, asociarlas a una determinada clase, en este caso una emoción. Para ciertos valores de los n-gramas tienen nombres especiales. Por ejemplo: los n-gramas de una palabra también se llaman unigramas; los n-gramas de dos palabras se llaman bigramas o digramas; los n-gramas de tres palabras también se llaman trigramas.

En algunos de los trabajos relevados⁶, se definen un conjunto de métricas, que permiten determinar la importancia de un término en un documento y la relevancia de ese término para expresar una emoción. A continuación se enumeran algunas de las más métricas recopiladas: “*Document frequency*”, “*Information gain*”, “*Point of Mutual Information*”, “*Odds Ratio*”, “*Chi square*”, “*Weighted Log Likelihood Ratio*”, “*Weighed Frequency and Odds*”, “*Latent semantic index*”, entre otras.

Separación de datos de prueba y entrenamiento. Para el entrenamiento de los algoritmos inteligentes que se propongan es necesario dividir los datos muestrales recolectados, en este caso los comentarios con su correspondiente reacción, en dos conjuntos: uno de los conjuntos estará reservado para el entrenamiento del algoritmo (generalmente conformado por el 80% del total de datos disponible); y un conjunto de datos reservado para pruebas (el complemento del total elegido para entrenamiento), conformado por datos nunca vistos por el algoritmo previamente y a través de los cuales será posible evaluar la exactitud alcanzada en la predicción. Ambos conjuntos deben conformarse seleccionando datos al azar pero se considera apropiado estratificados, es decir, mantener la proporción de muestras de cada clase (donde cada clase se corresponderá con una reacción).

Construcción y entrenamiento de algoritmos inteligentes. Esta etapa comprende la construcción y entrenamiento de algoritmos inteligentes, usando técnicas como redes neuronales, árboles de decisión, máquinas de soporte vectorial, etc. En primer lugar será necesario numerizar los datos muestrales (los comentarios), dado que estas técnicas exigen representar la entrada de una manera numérica. Posteriormente se realiza el entrenamiento del algoritmo para que aprenda la clasificación, es decir, la reacción vinculada a cada fragmento de texto. Finalmente se compara la respuesta sobre los datos de prueba del algoritmo con la respuesta esperada para medir exactitud y confiabilidad en la predicción de cada clase.

Construcción de una herramienta que integre el algoritmo propuesto. Una vez que se hayan seleccionado los algoritmos más apropiados, será necesario crear una interfaz amigable que encapsule su utilización a través de, por ejemplo, una aplicación Web o de escritorio, o un Servicio Web, o aquella interfaz que mejor se adapte al usuario.

⁶ Z.-H. Deng, K.-H. Luo, H.-L. Yu, A study of supervised term weighting scheme for sentiment analysis, *Expert Systems with Applications* 41 (2014) 3506–3513.

⁷ Pointwise mutual information on text. [online] Stackoverflow.com. Available at: <https://stackoverflow.com/questions/13488817/pointwise-mutual-information-on-text> [Accessed 28 Nov. 2017].