

Minería de Opiniones basado en la adaptación al español de ANEW sobre opiniones acerca de hoteles

Opinion Mining based on the spanish adaptation of ANEW on hotel customer comments

Carlos Henriquez Miranda
Universidad Autónoma del caribe
Cl. 90, Barranquilla, Colombia
chenriquez@uac.edu.co

Jaime Guzmán
Universidad Nacional de
Colombia
Cl. 80 #65223, Medellín,
Colombia
jaguzman@unal.edu.co

Dixon Salcedo
Universidad de la Costa
Cra. 55 #58-66, Barranquilla,
Colombia
dsalcedo2@cuc.edu.co

Resumen: La minería de opiniones (MO) ha mostrado una alta tendencia de investigación en los últimos años debido a la producción en gran escala de opiniones y comentarios por parte de usuarios activos en Internet. Las empresas y organizaciones en general están interesadas en conocer cuál es la reputación que tienen de sus usuarios en las redes sociales, blogs, wikis y otros sitios web. Hasta ahora, la gran mayoría de trabajos de investigación involucran sistemas de MO en el idioma inglés. Por este motivo, la comunidad científica está interesada en trabajos diferentes a este lenguaje. En este artículo se muestra la construcción de un sistema de minería de opiniones en español sobre comentarios dados por clientes de diferentes hoteles. El sistema trabaja bajo el enfoque léxico utilizando la adaptación al español de las normas afectivas para las palabras en inglés (ANEW). Estas normas se basan en las evaluaciones que se realizaron en las dimensiones de valencia, excitación y el dominio. Para la construcción del sistema se tuvo en cuenta las fases de extracción, preprocesamiento de textos, identificación del sentimiento y la respectiva clasificación de la opinión utilizando ANEW. Los experimentos del sistema se hicieron sobre un corpus etiquetado proveniente de la versión en español de *Tripadvisor*. Como resultado final se obtuvo una precisión del 94% superando a sistemas similares.

Palabras clave: Minería de opinión, Análisis de sentimiento, lexicón, PLN, ANEW.

Abstract: Recently, the Opinions Mining (OM) has shown a high tendency of research due to large-scale production of opinions and comments from users over the Internet. Companies and organizations, in general terms, are interested in knowing what is the reputation they have in social networks, blogs, wikis and other web sites. So far, the vast majority of research involving systems MO in English. For this reason, the scientific community is interested in researching different to this language. This article is about the construction of a mining system views in Spanish based on comments given by different clients and hotels. The system works on the lexical approach using Spanish adaptation of affective standards for English words (ANEW). These standards are based on evaluations conducted in the dimensions of valence, arousal and dominance. For the construction of the system took into account the phases of extraction, preprocessing of texts, identification of feelings and the respective ranking of the opinion using ANEW. System experiments were made on labeling a corpus from the spanish version of *Tripadvisor*. As a result, precision exceeding 94% was obtained at similar systems.

Keywords: Opinion mining, Sentiment analysis, lexicon, NLP, ANEW.

1 Introducción

Hoy en día la cantidad de datos producidos a nivel mundial es muy alta. Por ejemplo, en Internet se producen millones de datos debido a la utilización masiva de las redes sociales, servicios de mensajería, blogs, wikis, comercio electrónico, entre otros.

Toda esta gama de datos es atractiva para diferentes estamentos comerciales, industriales y académicos, pero la extracción y su respectivo procesamiento, hace que esta tarea sea muy compleja y difícil si se hace de forma manual. Sumado a esto, las personas del común participan activamente en Internet dejando sus propios comentarios, opiniones y hasta reseñas, en todo tipo de temas, usando su lenguaje nativo.

Debido a lo anterior, existen grandes frentes de trabajo para encontrar modelos, técnicas y herramientas que permitan el análisis de los textos de forma automática. Es allí donde tendencias actuales de inteligencia artificial como las técnicas de procesamiento de lenguaje natural (PLN) son una gran alternativa de investigación. Dentro del área de PLN existe una temática que ha llamado la atención en los últimos años: la minería de opiniones (MO). La MO busca analizar las opiniones, sentimientos, valoraciones, actitudes y emociones de las personas hacia entidades como productos, servicios, organizaciones, individuos, problemas, sucesos, temas y sus atributos (Liu, 2012).

La (MO) ha mostrado una gran tendencia de investigación en los últimos años debido a la producción a gran escala de opiniones y comentarios por parte de usuarios activos en Internet. Las empresas y organizaciones en general están interesadas en conocer cuál es la reputación que tienen de sus usuarios en las redes sociales, blogs, wikis y otros sitios web. Adicionalmente a esto, la gran mayoría de trabajos de investigación involucran sistemas de MO en el idioma inglés (Vilares, Alonso, y Gómez-Rodríguez, 2013). Por este motivo, la comunidad científica está interesada en trabajos diferentes a este lenguaje.

Este artículo pretende mostrar la construcción de un sistema de minería de opiniones en español

sobre comentarios dados por clientes de diferentes hoteles. El sistema trabaja bajo el enfoque léxico utilizando la adaptación al español de las normas afectivas para las palabras en inglés (ANEW) (Redondo et al., 2007).

El resto del artículo está organizado de la siguiente manera. En la sección 2 se abordan los antecedentes y trabajos similares. La sección 3 describe la metodología utilizada. La sección 4 muestra los experimentos y resultados, y finalmente las conclusiones.

2 Antecedentes y trabajos relacionados

La MO recibe en la literatura diferentes nominaciones o términos: el análisis de sentimientos, el análisis de subjetividad, análisis de la emoción, la computación afectiva y la extracción de la evaluación, entre otras. Las más usadas en la literatura son el análisis de sentimientos (AS) y la minería de opiniones (MO). Según Pang y Lee, (2008) son dos conceptos similares que denotan el mismo campo de estudio, que en sí mismo puede ser considerado como un sub-campo del análisis de la subjetividad.

Para la construcción de un sistema de MO se debe tener en cuenta varios aspectos. Primero la extracción de la opinión y luego la clasificación del sentimiento. Para la extracción se elige un conjunto de datos, que van desde redes sociales, hasta sitios web donde abundan opiniones y comentarios en línea (Henríquez y Guzman, 2015). Para la clasificación, normalmente positiva o negativa, se utilizan en su gran mayoría técnicas basadas en aprendizaje de máquinas (ML) y basadas en léxico (LEX). Las diferencias fundamentales radican en que la primera utiliza algoritmos para aprender y la segunda utiliza diccionarios y léxicos que ya vienen catalogados con su sentimiento. Las técnicas LEX a menudo pierden la lucha contra las ML porque dependen en gran medida de la calidad recursos lingüísticos, especialmente diccionarios de sentimientos (Medhat, Hassan, y Korashy, 2014). Las ML pueden lograr eficacia razonable pero la construcción de datos de marcado es a menudo costosa y necesita mucho trabajo humano (Brody y Elhadad., 2010).

El trabajo aquí presentado utiliza como recurso basado en LEX la adaptación al español de ANEW. La adaptación consistió en la traducción al español de las 1.034 palabras ANEW originales (Bradley y Lang, 1999) realizada por un filólogo profesional especializado en el idioma Inglés. A continuación un total de 720 participantes evaluaron las palabras en tres dimensiones: valencia, excitación y dominio en una escala de calificación de 9 puntos para cada dimensión (Redondo et al., 2007). Para la adaptación de ANEW al español se utilizó una medida pictográfica no verbal: el Maniquí de Autoevaluación (SelfAssessment Manikin, SAM). El rango del SAM, en la dimensión de valencia, se extiende desde lo agradable (representado por una figura sonriente) a lo desagradable (representado por una figura ceñuda) (Fraga et al., 2005).

Dentro de los trabajos relacionados, se encuentra en la literatura muchos aportes en el idioma inglés como: Martínez, García, y Sánchez, (2011) donde se emplean técnicas de PLN junto con procesos de análisis sentimental y tecnologías de la Web semántica para analizar sentimientos en el dominio de las películas de cine. En LI y Desheng, (2010) se desarrolla un algoritmo para analizar automáticamente la polaridad emocional de un texto sobre un conjunto de datos adquiridos de foros deportivos. En Rill et al., (2014) se presenta un sistema llamado *PoliTwi* diseñado para detectar emergentes temas políticos en *Twitter* antes que otros canales de información estándar. En Anjaria y Guddeti, (2014) se estudia la tarea de predicción de sentimiento sobre *Twitter* utilizando técnicas de aprendizaje automático, considerando los retweets. Otros trabajos se dedican al área de la salud como (Greaves et al., 2013), finanzas (Dong, Zou, y Guan, 2012) y revisión de opiniones en automóviles (Turney, 2002).

En cuanto a los aportes en el idioma español bajo la técnica ML se encuentra a Salas-Zárate et al., (2014) que muestra experimentos para estudiar la eficacia de la clasificación de las opiniones en cinco categorías: muy positiva, muy negativa, positiva, negativa y neutro, utilizando

la combinación de las características psicológicas y lingüísticas de LIWC (LIWC, 2015). En Valdivia et al., (2012) se propone el uso de meta-clasificadores que combinan técnicas de aprendizaje de máquina con el fin de desarrollar un sistema de clasificación de sentimiento para críticas de cine.

Dentro de las técnicas LEX se encuentran algunos trabajos como: (Molina-González et al., 2013) que genera un nuevo léxico al traducir al español un lexicón existente en inglés. En Rodríguez, (2013) se propone un método para cuantificar el interés de un usuario en un tema, por la cual desarrollan la herramienta *Tom* que utiliza un léxico creado de forma semiautomática mediante la traducción de un léxico existente en inglés.

Dentro de los aportes en el mismo dominio (opiniones de hoteles) del trabajo propuesto, se encuentran: Fiol et al., (2012) que analiza las relaciones que existen entre la imagen, conocimiento, lealtad de marca, calidad de marca y el valor del cliente medido a través de las opiniones de los turistas en *Tripadvisor*. En Moreno, Castillo, y García, (2010) se realiza un análisis de valoraciones de usuarios de hoteles usando un sistema conocido como *Sentitext* que permite un sistema de análisis de sentimiento independiente del dominio. En García, Gaines, y Linaza, (2012) que usa un lexicón propio para el análisis de 100 reseñas de usuarios en español para los sectores de alimentos, bebidas y alojamiento. En González, Cámara, y Valdivia, (2015) se presenta la clasificación de polaridad de opiniones en español (1816 comentarios) utilizando un recurso léxico adaptado al dominio turístico (*eSOLHotel*).

En cuanto al uso de ANEW en el idioma inglés, en Gökçay, İşbilir, y G.Yıldırım, (2012) se realiza una investigación entre la relación entre palabras y frases con valencia y excitación (ANET y ANEW) para abordar el problema del reconocimiento de emociones a partir de texto. El estudio es preliminar tratando de entender cuánto éxito se puede lograr en este esfuerzo sin usar complicados análisis sintácticos y semánticos. Por su parte en Hutto y Gilbert, (2014) se presenta el desarrollo del sistema VADER (Valence Aware Dictionary for

Sentiment Reasoning) que combina métodos cualitativos y cuantitativos para construir y validar empíricamente un conjunto de características léxicas junto con su medida asociada de intensidad de sentimiento. La construcción se basa en un banco de palabras de sentimiento existente LIWC, ANEW, y GI (General Inquirer, 2015).

En cuanto al uso de ANEW en español, se encuentra del-Hoyo-Alonso et al., (2015) que presenta una integración de varias técnicas de análisis de opinión. Las aproximaciones utilizadas son máquinas de aprendizaje, diccionario afectivo DAL (Whissell, 2009) y ANEW. Este enfoque es aplicado para el análisis de opiniones en *Twitter*.

3 Metodología

La construcción del sistema se basó en un sencillo modelo presentado en la Figura 1. Este modelo consiste en cuatro fases: extracción, preprocesamiento, identificación de sentimientos y clasificación de la opinión.

En la primera fase, se realiza el proceso de extracción del texto que contiene las opiniones. Este proceso puede ser la recuperación de un párrafo en línea desde la Web, redes sociales, blogs, micro blogs o un corpus previamente definido.

La siguiente fase busca la aplicación de diferentes técnicas para obtener un texto más limpio. Dentro de estas técnicas están desde el borrado de palabras y signos sin sentido, corrección de ortografía, normalización y lematización entre otras.

La fase identificación de sentimientos trabaja con el texto ya normalizado, y lo que busca es identificar dentro de la opinión las características esenciales que permitan descubrir seguidamente el sentimiento asociado. Aquí se etiqueta morfológicamente cada palabra con su respectiva categoría gramatical (verbo, adjetivo, adverbio, nombre etc.). Una vez hecho el etiquetado se toman para la clasificación todas aquellas palabras en las categorías de verbo, adjetivo o adverbio.



Figura 1. Modelo para minería de opiniones.

La última fase, toma las palabras identificadas anteriormente para asignar una polaridad, es decir, definir el sentimiento asociado. Para determinar la polaridad se tiene en cuenta de ANEW únicamente la dimensión de valencia. El rango de valencia va desde 1(agradable) y 9 (desagradable). El primer paso consiste en buscar las palabras etiquetadas en el recurso para obtener su valencia. De este proceso se calcula el promedio de los datos encontrados y se asigna la polaridad por cada opinión dependiendo de su promedio así: negativa (valencia menor que 4), neutra (entre 4 y 6) y positivas (valencia mayor igual que 6). Si ninguna palabra es hallada en ANEW, se procede a tomar las palabras y buscarlas en un diccionario de sentimientos. Allí se cuentan las palabras con inclinación positiva y negativa y se asigna la polaridad así: Mas palabras positivas: positiva, más palabras negativas: negativa, igual número de palabras: neutra. Si ninguna palabra de la opinión es encontrada ni en ANEW ni en el diccionario se considera opinión no procesable.

Con respecto al diccionario de sentimientos utilizado se utiliza como base 71 adjetivos positivos y 52 negativos del trabajo de Jiménez, Vázquez, y Hernangómez, (1998) y estos se complementan manualmente con sinónimos y antónimos hasta llegar a 600 términos en total. En última instancia en esta fase, hay un manejo básico de negación. Primero se analiza si la opinión tiene o no inclinación negativa y luego se afecta el promedio calculado que afecta a la polaridad. Para la primera parte se toma como

diccionario de sentimientos hecho de forma manual. No se maneja la negación.

Experimento 4: Se analizan las palabras usando ANEW complementando con un diccionario de sentimientos hecho de forma manual. Con manejo de negación.

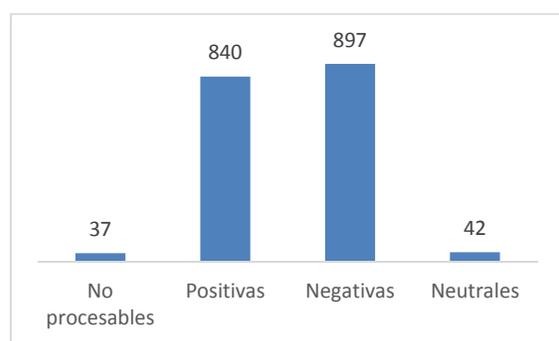


Figura 4. Histograma resultado de clasificación de las opiniones

Para la validación de los experimentos se ha escogido la medida de precisión. Esta se calcula entre el número de ejemplos positivos clasificados correctamente dividido por el número de ejemplos etiquetados por el sistema como positivos (Sokolova y Lapalme, 2009). Los resultados se muestran en la Tabla 1.

Experimento	Precisión
1	89
2	92
3	93
4	94

Tabla 1. Resultados precisión del sistema

De acuerdo a los resultados obtenidos se nota que la utilización única del recurso ANEW dio como resultado una precisión máxima de 92%. Resultado no despreciable ya que no se tuvo en cuenta el dominio en el análisis.

Por otra parte, se ha aumentado la precisión en casi dos puntos (94%) al incluir a la clasificación un diccionario manual con palabras que representan sentimientos positivos y

negativos no incluidos en ANEW. Con respecto al manejo negativo, el algoritmo utilizado para reconocer este tipo de opiniones ha funcionado eficazmente ya que ha logrado aumentar en 3% los resultados.

Para la verificación del sistema se ha establecido una comparación con trabajos similares de la literatura. Se ha tomado como base el idioma español, enfoque léxico de sentimiento y el dominio de hoteles. Para esto se han escogidos los trabajos: (Moreno, Castillo, y García, 2010), (García, Gaines, y Linaza, 2012) y (González, Cámara, y Valdivia, 2015). En la Tabla 2 se muestra la comparación.

Trabajo	Opiniones	Léxico	Precisión
(Moreno, Castillo, y García, 2010)	100	<i>Sentitext</i>	84.8
(García, Gaines, y Linaza, 2012)	994	<i>Léxico propio</i>	80.0
(González, Cámara, y Valdivia, 2015)	1816	<i>eSOLHotel</i>	84.7
Propuesta	1816	Anew	94.4

Tabla 2. Comparación con sistemas de MO en el dominio de hoteles.

Como se puede apreciar el enfoque aquí presentado alcanza mejores resultados que los sistemas comparados en cuanto a la medida de precisión. Cabe destacar que el sistema que más se asemeja es (González, Cámara, y Valdivia, 2015) ya que hace el análisis sobre el mismo corpus, superando casi en 10 puntos su precisión. De igual forma pasa con (Moreno, Castillo, y García, 2010) a pesar de que se tienen un número más alto de opiniones procesadas.

Finalmente se considera que este sistema es una base firme para nuevos trabajos de minería de opinión que utilicen estos recursos u otros existentes para el análisis de textos en el idioma español independiente del dominio.

5 Conclusiones

Se ha logrado un sistema más que aceptable alcanzando una precisión del 94% para el dominio de turismo, específicamente hoteles, teniendo en cuenta como base la utilización del recurso lingüístico independiente ANEW. Cabe destacar la utilización de varios algoritmos complementarios para el cálculo de la polaridad.

El uso de herramientas de PLN potentes como FREELING permite construcción de sistemas de minería de opiniones muy robustos.

La calidad del recurso ANEW permitirá realizar nuevos experimentos de análisis de sentimientos enfocados a diferentes dominios empleando la metodología propuesta.

Bibliografía

- Anjaria, M. y R. M. Guddeti. 2014. A novel sentiment analysis of social networks using supervised learning. *Social Network Analysis and Mining*, 4(1), 1-15.
- Bradley, M. M. y P. J. Lang. 1999. Affective norms for English words (ANEW): Instruction manual and affective ratings. *Technical Report C-1, The Center for Research in Psychophysiology*, 1 - 45.
- Brody, S. y N. Elhadad. 2010. An unsupervised aspect-sentiment model for online reviews. *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*.
- del-Hoyo-Alonso, R., M. D. Rodrigalvarez-Chamarro, J. Veá-Murguía, y R. M. Montañes-Salas. 2015. Algoritmo de ensamble con introducción de la estructura morfosintáctica para la mejora del análisis de opinión. *TASS 2015*.
- Fraga, I., M. Perea, J. Redondo, y M. Vila. 2005. Estudio normativo del valor afectivo de 478 palabras españolas. *Psicológica: Revista de metodología y psicología experimental*, 317-326.
- García, A., S. Gaines, y M. T. Linaza. 2012. A Lexicon Based Sentiment Analysis Retrieval System for Tourism Domain. *e-Review of Tourism Research*, 39(10), 35-38.
- General Inquirer. 2015. Recuperado el 05 de 2015, de <http://www.wjh.harvard.edu/~inquirer/>
- Gökçay, D., E. İşbilir, y G. Yıldırım. 2012. Predicting the sentiment in sentences based on words: An Exploratory Study on ANEW and ANET. *Cognitive Infocommunications (CogInfoCom), 2012 IEEE 3rd International Conference on. IEEE, 2012*.
- Greaves, F., D. Ramirez-Cano, C. Millett, A. Darzi, y L. Donaldson. 2013. Use of Sentiment Analysis for Capturing Patient Experience From Free-Text Comments Posted Online. *Journal of medical Internet research*, 15(11).
- Henriquez, C. y J. Guzmán. 2015. Extracción de información desde la web para identificar acciones de un modelo de dominio en planificación automática. *Ingeniare*, 23(3), 439-448.
- Hunter, J. 2000. *JDOM*. Recuperado el 01 de 06 de 2015, de <http://www.jdom.org/>
- Hutto, C. y E. Gilbert. 2014. VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. *Eighth International AAAI Conference on Weblogs and Social Media. 2014*.
- Jiménez, F., C. Vázquez, y L. Hernangómez. 1998. Adjetivos en castellano de contenido depresivo autorreferente y de contenido neutral: Normas de emocionalidad y frecuencia subjetiva de uso. *Revista de psicopatología y psicología clínica*, 199-215.
- Jiménez-Zafra, S. M., E. M. Martínez-Cámara, M. T. Martín-Valdivia, y M. D. Molina-González. 2015. Tratamiento de la Negación en el Análisis de Opiniones en Español. *Procesamiento del Lenguaje Natural*, 37-44.
- LI, N. y D. D. Wu. 2010. Using text mining and sentiment analysis for online forums hotspot detection and forecast. *Decision Support Systems*, 48(2), 354 - 368.

- Liu, B. 2012. Sentiment Analysis and Opinion Mining. *Synthesis Lectures on Human Language Technologies*.
- LIWC. 2015. Recuperado el 06 de 2015, de <http://liwc.wpengines.com/>
- Martin-Valdivia, M. T., E. M. Martínez-Cámara, J. M. Perea-Ortega, y L. A. Ureña-López. 2013. Sentiment polarity detection in Spanish reviews combining supervised and unsupervised approaches. *Expert Systems with Applications*, 4(10), 3934–3942.
- Medhat, W., A. Hassan, y H. Korashy. 2014. Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 1093-1113.
- Molina-González, M. D., E. Martínez-Cámara, M. Martín-Valdivia, y J. M. Perea-Ortega. 2013. Semantic orientation for polarity classification in Spanish reviews. *Expert Systems with Applications*, 40(18), 7250-7257.
- Molina-González, M. D., E. Martínez-Cámara, M. Martín-Valdivia, y L. Ureña-López. 2014. Cross-Domain Sentiment Analysis Using Spanish Opinionated Words. *19th International Conference on Applications of Natural Language to Information Systems*, 8455, págs. 214-219. Montpellier.
- Molina-González, M. D., E. Martínez-Cámara, M. Martín-Valdivia, y S. M. Jiménez-Zafra. 2015. eSOLHotel: Generación de un lexicón de opinión en español adaptado al dominio turístico. *Procesamiento del Lenguaje Natural*, 54, 21-28.
- Moreno, A., F. Pineda Castillo, y R. H. Hidalgo García. 2010. Análisis de Valoraciones de Usuario de Hoteles con Sentitext. *Procesamiento del Lenguaje Natural*, 45, 31-39.
- Padró, L. y E. Stanilovsky. 2012. FreeLing 3.0: Towards Wider Multilinguality. *Proceedings of the Language Resources and Evaluation Conference (LREC 2012) ELRA*.
- Pang, B. y L. Lee. 2008. Opinion Mining and Sentiment Analysis. *Foundations and trends in information retrieval*, 2(1-2) 1-135.
- Peñalver-Martínez, I. P., F. García-Sánchez, y R. Valencia-García. 2011. Minería de Opiniones basada en características guiada por ontologías. *Procesamiento del Lenguaje Natural* (46), 91-98.
- Redondo, J., I. Fraga, I. Padrón, y M. Comesaña. 2007. The Spanish adaptation of ANEW. *Behavior Research Methods*, 3, 39.
- Rill, S., D. Reinell, J. Scheidt, y R. V. Zicari. 2014. PoliTwi: Early detection of emerging political topics on twitter and the impact on concept-level sentiment analysis. *Knowledge-Based Systems*, 69, 24-33.
- Rodríguez, F. 2013. *Cuantificación del interés de un usuario en un tema mediante minería de texto y análisis de sentimiento*. Diss. Universidad Autónoma de Nuevo León.
- Salas-Zárate, M., E. López-López, R. Valencia-García, N. Aussenac-Gilles, Á Almela, y G. Alor-Hernández. 2014. A study on LIWC categories for. *Journal of Information Science*, 40(6), 749-760.
- Sokolova, M. y G. Lapalme. 2009. A systematic analysis of performance measures for classification tasks. *Information Processing and Management*, 427 - 437.
- Turney, P. D. 2002. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. *Proceedings of the 40th annual meeting on association for computational linguistics*. Stroudsburg, PA, USA.
- Vilares, D., M. A. Alonso, y C. Gómez-Rodríguez. 2013. Clasificación de polaridad en textos con opiniones en español mediante análisis sintáctico de dependencias. *Procesamiento del lenguaje natural*, 13-20.
- Whissell, C. 2009. Using the revised dictionary of affect in language to quantify the emotional undertones of samples of natural language 1, 2. *Psychological reports*, 105(2), 509-521.