

Tipo de artículo: Artículo original
Temática: Inteligencia Artificial
Recibido: 03/12/2014 | Aceptado: 19/01/2015

PosNeg opinion: Una herramienta para gestionar comentarios de la Web

PosNeg opinion: A tool for managing comments from the web

Mario Amores ^{1*}, Leticia Arco ¹, Michel Artiles ¹

¹ Departamento de Ciencia de la Computación, Universidad Central "Marta Abreu" de Las Villas. Carretera a Camajuaní km 5 1/2. Santa Clara. Villa Clara. Cuba. Correo-e: leticiaa.mae@uclv.edu.cu

* Autor para correspondencia: marioamores@uclv.cu

Resumen

Una de las principales tareas de la minería de opinión es la clasificación de la polaridad de la opinión, que consiste en determinar si la opinión es positiva o negativa con respecto a la entidad a la que se esté refiriendo. En este trabajo se propone la aplicación PosNeg Opinion que detecta de manera no supervisada la polaridad de opiniones siguiendo un esquema compuesto por cinco etapas: identificar tokens, desambiguar léxicamente cada token, obtener las acepciones de cada palabra, clasificar los tokens en positivo o negativo y evaluar la opinión. PosNeg Opinion detecta la polaridad de opiniones en español. El sistema propuesto mostró un buen desempeño al clasificar la polaridad de 200 opiniones provenientes de foros de discusión en Yahoo.es, obteniéndose una exactitud de 0.965 y una precisión de 0.970.

Palabras clave: análisis de sentimiento, detección de polaridad, minería de opinión, web semántica.

Abstract

One of the primary tasks of opinion mining is the classification of the polarity of the opinion, that consists in determining whether the opinion is positive or negative with respect to the entity to which it is referring. This work seeks design a general scheme consists in five stages that enable the polarity detection of opinions in an unsupervised manner. The five considered stages are: token identification, lexical disambiguation, get all meanings of each word, classify each token in positive or negative, and evaluate the opinion. PosNeg Opinion detects the polarity of the opinions in Spanish language. The proposed system showed good performance in classifying the polarity of 200 opinions from discussion forums Yahoo.es, yielding an accuracy of 0.965 and a precision of 0.970.

Keywords: opinion mining, polarity detection, semantic Web, sentiment analysis.

Introducción

La minería de opinión es un área de la minería de textos consistente en la clasificación de palabras, textos o documentos de acuerdo a las opiniones, sentimientos, emociones y subjetividades expresadas (Vinodhini, 2012; Sharma, 2014).

Una de las tareas de la minería de opinión es la detección de la polaridad de las opiniones, que consiste en ser capaces de determinar si una opinión es positiva o negativa (Chung, 2014, Veselovská, 2011; Martín-Wanton, 2010). Más allá de una polaridad básica, también se puede querer obtener un valor numérico dentro de un rango determinado, que de una determinada forma trate de obtener una clasificación objetiva asociada a una determinada opinión (Liu, 2010; Seerat, 2012).

Se dice que un término tiene polaridad u orientación cuando éste porta información subjetiva bien sea positiva o negativa. En este sentido, las colocaciones pueden ser explotadas puesto que determinados términos pueden adquirir o cambiar su polaridad dependiendo de si forma parte o no de una colocación. Por ejemplo, el adjetivo *alto*, un término que a priori no tiene polaridad de ningún tipo, adquiere polaridad negativa al formar parte de la colocación *precio alto*. El mismo adjetivo en colocación con *valor* tiene, por el contrario, polaridad positiva. El problema está en cómo asignar automáticamente la polaridad.

Los cálculos de polaridad en la minería de opinión se pueden estructurar en varias fases (Liu, 2010):

1. Detectar la subjetividad: determinar si una unidad textual tiene naturaleza objetiva (hecho) o subjetiva (opinión).
2. Clasificar la opinión: determinar su polaridad, es decir, si la opinión es negativa o positiva.
3. Determinar la fuerza de la opinión: expresar en qué medida es positiva o negativa.
4. Determinar la fuente de la opinión: identificar si la fuente de la opinión fue una persona o una institución, esta tarea requiere frecuentemente resolución de anáforas.
5. Determinar el objetivo de la opinión: determinar de quién se habla en la opinión, con quién se está de acuerdo o no.
6. Resumir las opiniones y/o visualizar gráficamente los resultados: puede ser agregando votos (índice de 1-5, estrellas), sobresaltando algunas opiniones, representando acuerdo/desacuerdo, etc.

La clasificación del sentimiento o clasificación de la polaridad es una tarea del aprendizaje automatizado que puede ser supervisada, no supervisada o semi-supervisada (Buche, 2013; Rashid, 2013). Las dos aproximaciones más utilizadas para resolver automáticamente la polaridad de un texto son: aprendizaje automatizado supervisado y

orientación semántica. Los clasificadores obtenidos a partir de la primera alternativa se caracterizan por conseguir un buen rendimiento base para el dominio en el que son entrenados. Sin embargo, presentan complicaciones para mejorar su precisión, están sujetos al sobre entrenamiento y son altamente dependientes de la calidad, tamaño y dominio de los datos de entrenamiento. Además, las soluciones desarrolladas empeoran drásticamente su rendimiento cuando se utilizan para analizar textos de un dominio diferente al del corpus con el que se entrenaron. La segunda alternativa permite una mejor adaptación a los diferentes dominios, contempla más aspectos del texto, se basa en recursos externos (por ejemplo: WordNet Affect (Strapparava, 2004) o SentiWordNet¹), pero actualmente existen pocos recursos disponibles y son mayormente dependientes del idioma, sobre todo del idioma Inglés.

Existen varias herramientas que permiten realizar minería de opinión, algunas de ellas las relacionamos a continuación (Angulakshmi, 2014). ITelligent² que es un sistema de minería de opinión para inteligencia comercial; OPAL³ plugin de Drupal que analiza comentarios realizados por los usuarios y detecta si el resultado es positivo, negativo o neutral; IIC-Lynguoque⁴ es un conjunto de herramientas que ayudan a extraer la opinión positiva, negativa y neutra de un texto; netOpinion⁵ permite conocer opiniones de productos o servicios en foros y redes sociales; WebOpinion⁶ gestiona mes a mes la evolución de la imagen de un usuario en la red y Sentitex⁷ que consiste en un conjunto de aplicaciones para el análisis de sentimientos en textos, entre otras. Desafortunadamente la mayoría de estas herramientas son propietarias y es necesario pagar la licencia para su utilización, ya que fueron concebidas con fines comerciales. Adicionalmente, cada una de ellas responde a objetivos específicos de la minería de opinión. De ahí que el objetivo de este trabajo consiste en desarrollar una aplicación que permita detectar de manera no supervisada la polaridad de opiniones en idioma Español, reutilizando las herramientas, módulos y bibliotecas que permiten realizar alguna etapa de detección de la polaridad de las opiniones.

Materiales y métodos

En esta sección haremos referencia a los materiales que tuvimos en cuenta para el diseño de la investigación, así como la descripción del método que se propone. Existen varias herramientas de software libre que contribuyen a la detección no supervisada de la polaridad de las opiniones. En este trabajo hacemos una selección de aquellas que, utilizadas de manera integrada, permiten realizar análisis de sentimiento. Proponemos un esquema general que permite detectar de manera no supervisada la polaridad de las opiniones. Este esquema hace uso de las herramientas existentes e incluye nuevos métodos para finalmente realizar el cálculo de la polaridad.

Herramientas que contribuyen a la detección de la polaridad de las opiniones

Como mencionamos anteriormente, existen varias herramientas de software libre que permiten realizar algunas de las etapas necesarias en la detección de la polaridad de las opiniones (Strapparava, 2004; Esuli, 2006; Miller, 1995; Stone, 1966; Schmid, 1994), facilitando de esta forma el desarrollo de aplicaciones con este objetivo. A continuación se describen las herramientas que se utilizan en la aplicación que se propone en este trabajo.

TreeTagger: El TreeTagger (Schmid, 1994) es una herramienta para anotar textos con información de part-of- speech y lema, desarrollada por investigadores del Instituto para lingüística computacional de la Universidad de Stuttgart. Ha sido utilizada con éxito para etiquetar textos en Alemán, Inglés, Francés, Italiano, Español, Griego, y Francés antiguo, y es fácilmente adaptable a otros lenguajes si se dispone de un lexicón y corpus marcado manualmente.

Esta herramienta nos brinda dos facilidades pues podemos lematizar los términos obtenidos, así como desambiguarlos lexicalmente, muy fácilmente. Además, es una herramienta totalmente compatible con Java que es el lenguaje de programación usado para desarrollar la aplicación que aquí se propone.

SentiWordNet: El SentiWordNet (Esuli *et. al.*, 2006) es un recurso léxico creado para la minería de opiniones, en el cual, a cada synset de WordNet se le asignan tres valores de sentimiento: Positivo, Negativo y Objetivo, cuya suma es 1. Por ejemplo, el synset#3 del adjetivo *happy* (*{felicitous, happy}* - (*marked by good fortune; "felicitous life"; "happy outcome"*)) está clasificado como: Positivo=0.875, Negativo=0.0 y Objetivo=0.125.

Este recurso fue creado mediante una mezcla de técnicas lingüísticas y clasificadores estadísticos. Por ser construido semiautomáticamente todos los resultados no fueron validados manualmente, por lo que, algunas clasificaciones pueden ser incorrectas. Por ejemplo, el synset#1 del sustantivo *flu* (*{influenza, flu, gripe}* - (*an acute febrile highly contagious viral disease*)) se clasificó como Positivo=0.75, Negativo=0.0, Objetivo=0.25, a pesar de tener varias palabras negativas en su glosario.

General Inquirer: El General Inquirer (GI) (Stone *et. al.*, 1966) es un diccionario en Inglés que contiene información sobre las palabras; esto incluye etiquetas de las categorías: positiva, negativa, negaciones, intensificadores, etc. Las categorías se nombran positiv, negativ y negate en GI. A partir de estas categorías se crearon las listas de palabras modificadoras, intensificadoras y negadoras de la polaridad.

Las listas creadas contienen algunas palabras que por sí mismas pueden aumentar, disminuir o negar la positividad (negatividad) de una palabra. Dentro de la lista de palabras intensificadoras se encuentran dos categorías, las positivas

y las negativas, éstas son palabras que siempre serán positivas o negativas independientemente del contexto donde se encuentren. Dentro de la lista de las modificadoras se encuentran palabras y terminaciones que precediendo o sucediendo otras palabras modifican éstas, como es por ejemplo, el término muy y dentro de la negaciones son aquellos términos que cambian a la palabra que preceden de positiva a negativa, o viceversa, como es el caso de *no*. Las listas de las palabras modificadoras, intensificadoras y negadoras de la polaridad son pequeñas y contienen palabras fáciles de traducir a otro idioma. Esto permite que el método propuesto pueda ser adaptado a otros idiomas sin mucho esfuerzo.

Índice intralingüístico: El índice intralingüístico permite desarrollar aplicaciones que sean independientes del idioma y también facilita el intercambio de información entre herramientas concebidas para idiomas diferentes.

A través de este índice intralingüístico es posible obtener todas las acepciones en Inglés del término analizado originalmente en Español. Por ejemplo, para la palabra *agresor* obtendríamos las siguientes acepciones: *attacker assailant aggressor assaulter aggressor robber*.

El formato que contiene este índice es primeramente el término en español seguido por un carácter de espacio, luego la etiqueta pos que puede tomar valores n, a, v, r (n- sustantivo, a- adjetivos, v- verbos, r- adverbios), a continuación un carácter de tabulación y seguido varios identificadores de las diferentes relaciones semánticas del término, seguido por otro carácter de tabulación y finalmente las acepciones de la palabra en Inglés separadas por un carácter de espacio. A continuación se muestra un ejemplo con la palabra *agresor* para su mejor comprensión:

agresor n 09195176 09158637 09848308 attacker assailant aggressor assaulter aggressor robber

Esquema general para la detección no supervisada de opiniones

Los resultados que en la próxima sección se presentarán se basan en el esquema general para la detección no supervisada de opiniones propuesto en (Amores, 2013). Esta propuesta determina la polaridad de las oraciones teniendo en cuenta la polaridad de todas las acepciones de las palabras que se analizan al traducirse al idioma Inglés mediante un índice intralingüístico. Este esquema consta de cinco etapas como se muestra en la Figura 1.

La Etapa 1 es la encargada de leer las opiniones que fueron especificadas en el XML de entrada y seleccionar los términos que aporten información útil, eliminando las palabras vacías. A continuación, en la Etapa 2, se parte de cada término que aporta información útil, éste se lematiza y se desambigua lexicalmente. Posteriormente, en la Etapa 3 se traducen los términos seleccionados en la Etapa 2, obteniendo todas las acepciones del término en Inglés. Esta etapa es necesario ya que se sugiere el uso del SentiWordNet para clasificar la polaridad de las palabras, y éste requiere que

los términos estén en idioma Inglés. En la Etapa 4, se propone un método para el cálculo de la polaridad de los términos, considerando la polaridad de cada una de las acepciones del término. Se suman todos los valores de polaridad positivos de las acepciones, así como todos los valores de polaridad negativos, y se toma el mayor valor, asignándolo como la polaridad del término, y por tanto, contribuyendo a la polaridad de la opinión. Así, en la Etapa 5, al terminar de analizar todos los términos y sus acepciones, la opinión cuenta con un valor positivo y otro negativo, los cuales son comparados, y se toma como polaridad de la opinión el mayor valor.

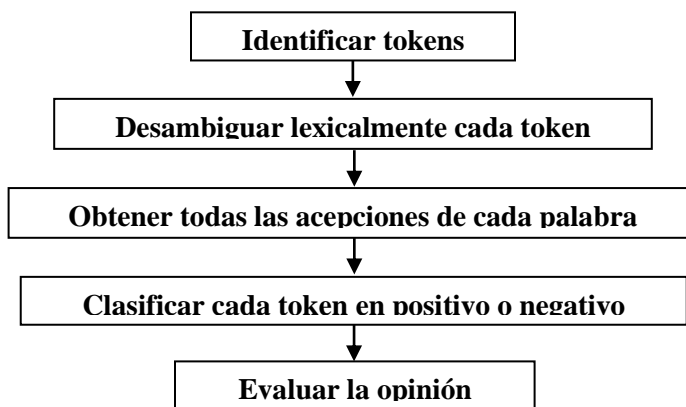


Figura 1. Esquema general para detectar la polaridad de las opiniones

Resultados y discusión

Como resultado de esta investigación se desarrolló el sistema PosNeg Opinion que permite detectar de manera no supervisada la polaridad de las opiniones. Este sistema utiliza las herramientas referenciadas en la sección anterior, y sigue el esquema general para detectar la polaridad de las opiniones propuesto en (Amores, 2013). En esta sección presentaremos una descripción general del sistema PosNeg Opinion, ilustraremos como se realiza la clasificación de una determinada opinión y finalmente mostraremos los resultados de la validación de PosNeg Opinion considerando la clasificación de 200 opiniones procedentes de varios foros de discusión en Yahoo.es.

PosNeg Opinion

PosNeg Opinion permite que el usuario analice un gran cúmulo de opiniones de manera sencilla, ya que se convierten los ficheros XML a texto plano y se analizan independientemente las opiniones almacenadas en una lista. Además, para el usuario es transparente el procesamiento de los datos, así como la teoría y los algoritmos que se aplican en el análisis. Esta aplicación puede utilizarse como un módulo de una aplicación más general de minería de opinión, pues

resuelve una de las fases de este proceso y es fácilmente reutilizable. Además, PosNeg Opinion se puede comunicar con otras aplicaciones mediante ficheros XML.

PosNeg Opinion fue desarrollada completamente en JAVA, por lo que es multiplataforma. Necesita como entrada un fichero XML con todas las opiniones a analizar y como salida muestra cuántas fueron positivas y cuántas negativas. A petición del usuario también retorna el porcentaje de las opiniones negativas y positivas así como una lista con las opiniones negativas y otra con las opiniones positivas. Adicionalmente, la aplicación destaca cuáles opiniones fueron las de mayor puntuación en cada caso (positivas/negativas).

PosNeg Opinion solo requiere que el Sistema Operativo tenga instalado Java Runtime Environment (JRE). Su archivo .jar ocupa 111,1 KB y se nombra OM.jar.

PosNeg Opinion requiere que junto con el archivo OM.jar se encuentren los siguientes archivos:

1. TreeTagger: brinda la posibilidad de lematizar los términos obtenidos, así como desambiguarlos lexicalmente, muy fácilmente.
2. spanish.par: permite especificarle al TreeTagger que el idioma a analizar es el Español.
3. Lib: contiene las bibliotecas de diferentes ficheros .jar necesarios para el proceso de filtrado de las opiniones y trabajo con los XML.
4. lucene-core-2.9.4.jar, org.annolab.tt4j-1.0.5.jar: Estas bibliotecas son determinantes para poder utilizar el analizador de Lucene el etiquetador y lematizador del TreeTagger.
5. es.txt: contiene la lista de todas las palabras gramaticales (o palabras vacías) que se van a eliminar (stopwords).
6. opiniones.xsd: es el esquema por el cual se van a validar los XML a ser procesados por la aplicación.
7. pneg.txt: contiene la lista de palabras intensificadora de la polaridad negativa.
8. ppos.txt: contiene la lista de palabras intensificadora de la polaridad positiva.
9. bonif.txt: contiene la lista de palabras modificadoras de la polaridad.
10. nega.txt: contiene la lista de palabras negadoras de la polaridad.
11. SentiWordNet_3.0.0.txt: es un recurso léxico creado para la minería de opiniones. A cada synset de WordNet se le asignan tres valores de sentimiento: Positivo, Negativo y Objetivo, cuya suma es 1.
12. translation_english.txt: contiene el índice intralingüístico con el cual obtenemos todas las acepciones del término analizado. Este índice es extremadamente útil ya que una palabra en Español puede tener varias acepciones en el idioma Inglés.

Tanto el ejecutable como las carpetas antes mencionadas requieren ser ubicados en un directorio donde se tenga permiso de escritura y lectura.

La interfaz visual de la aplicación, como se muestra en la Figura 2, se puede dividir en tres partes: la entrada que es donde se inicializan todos los archivos necesarios para comenzar a analizar las opiniones, una segunda parte que son los resultados globales y una tercera donde se muestra una parte de los resultados, especificando aquellas opiniones más positivas y más negativas.

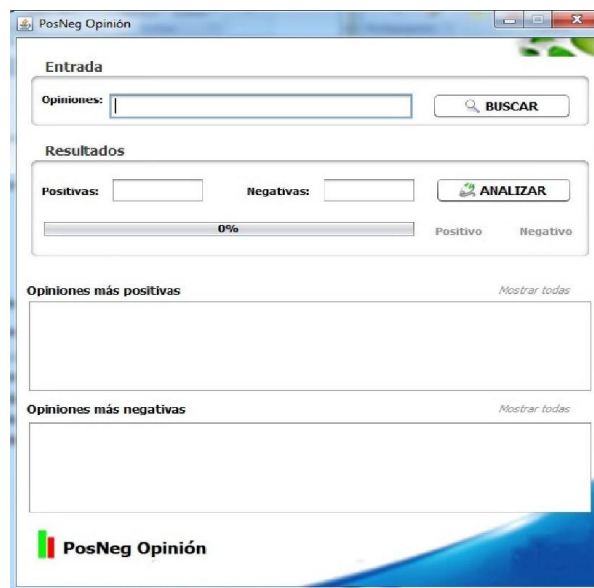


Figura 2. Interfaz de la aplicación PosNeg Opinión

Dentro del área de entrada, que se encuentra en la parte superior de la aplicación, existe la opción de cargar el fichero XML con las opiniones que se desean procesar. Este archivo XML debe tener la estructura que se muestra en la Figura 3.

```
<ROOT>
  <Opinion>
    <autor>
      <\autor>
    <contenido>
      <\contenido>
    <\Opinion>
  <ROOT>
```

Figura 3. Formato de los ficheros de entrada.

En el área del centro de la aplicación se muestran los resultados globales, se muestran cuántas opiniones resultaron positivas y cuántas resultaron negativas. Adicionalmente, se puede elegir mostrar el porcentaje de opiniones positivas o negativas mediante las etiquetas: Positiva o Negativa. En el área final se presentan las opiniones con mayores valores de positividad y negatividad, ofreciendo información al usuario para la futura toma de decisiones.

Ejemplo de una opinión clasificada como positiva

Dado un fichero XML con una opinión, lo primero que se debe hacer es obtener el contenido de la opinión a analizar. Ilustraremos el proceso de clasificación de la polaridad de las opiniones, tomando como ejemplo la siguiente opinión: *la hp 2000 tiene muy buena batería*. Al pasar este contenido por el analizador de Lucene obtenemos los siguientes términos: *muy, buena, batería*. Aquí es importante señalar que PosNeg Opinion elimina los términos *la, hp, 2000 y tiene* por ser palabras vacías (stopwords) y no ofrecen información para la detección de la polaridad de la opinión. Después de pasar por el proceso de desambiguación léxica y lematizado quedan clasificados los términos de la siguiente forma: *muy r, bueno a, batería n*; donde *r* es adverbio, *a* es adjetivo y *n* es sustantivo. En la Tabla 1 se muestran los términos con sus acepciones en Inglés junto a su valor de positividad y negatividad del SentiWordNet.

Tabla 1. Términos con sus acepciones y puntuación de la opinión positiva

Término	Acepciones	Puntuación	
		Positivo	Negativo
bueno	good	16.63	2.93
	right	23.63	3.68
	salutary	26.13	3.68
	glad	28.00	4.93
	fine	30.50	5.30
	well	38.29	6.01
	beneficial	38.92	6.01
	benevolent	41.04	6.01
	correct	44.04	7.26
batería	battery	44.04	7.51
	electric battery	44.04	7.51

Dentro de los términos analizados se encuentra *muy* que aparece en la lista de modificadores de la polaridad y a continuación le sigue el término *bueno* que tiene una polaridad positiva, por lo que se le adiciona 1 a la polaridad positiva dejando un saldo para esta opinión de 411.50 de polaridad positiva y 64.47 de polaridad negativa; considerándose así la opinión como positiva.

Validación de los resultados obtenidos con PosNeg Opinion

A continuación presentaremos las principales medidas para evaluar la calidad de la clasificación de la polaridad de las opiniones y mostraremos cómo se comportan estas medidas al clasificar con PosNeg Opinion la polaridad de comentarios tomados de la web. Para la validación del sistema desarrollado se creó una colección de 100 opiniones positivas y 100 opiniones negativas etiquetadas manualmente y procedentes todas de varios foros de discusión en Yahoo.es. Escogimos diferentes dominios para probar que puede funcionar para cualquier tipo de opinión. Algunos de los dominios elegidos son: productos de belleza, artículos informáticos, cine y televisión, y temas sociales.

La detección de la polaridad de las opiniones se reduce a un problema de clasificación de cada opinión en positiva y negativa. De ahí que podamos realizar una evaluación supervisada de PosNeg Opinion utilizando las medidas clásicas que permiten medir el desempeño de un clasificador: Exactitud y Precisión.

La Exactitud (Ac, del inglés Accuracy), como se muestra en la ecuación (1), es la proporción de la cantidad total de predicciones que fueron correctas.

$$Ac = \frac{a + b}{a + b + c + d} \quad (1)$$

La Precisión (P, del Inglés Precision), como se muestra en la ecuación (2), es la proporción de casos predichos positivos que fueron correctos.

$$P = \frac{d}{b + d} \quad (2)$$

En ambas expresiones, (1) y (2), a es la cantidad de predicciones correctas de que un caso es negativo (verdadero negativo), b es la cantidad de predicciones incorrectas de que un caso es positivo (falso positivo), c es la cantidad de predicciones incorrectas de que un caso es negativo (falso negativo) y d es la cantidad de predicciones correctas de que un caso es positivo (verdadero positivo).

En la Tabla 2 se presentan los resultados de las medidas de evaluación de la clasificación de la polaridad de las 200 opiniones de estudio con la aplicación PosNeg Opinión. De esta manera se evidencia el buen desempeño de PosNeg Opinion en la detección de la polaridad de comentarios de la Web.

Tabla 2. Resultados de la validación de PosNeg Opinion considerando Exactitud y Precisión

Medida	Resultado
Exactitud (Ac)	0.965
Precisión (P)	0.970

Conclusiones

Como resultado de este trabajo se desarrolló la aplicación PosNeg Opinion que permite la detección de manera no supervisada de la polaridad de opiniones siguiendo el esquema de cinco etapas propuesto. La aplicación se auxilia de las herramientas SentiWordNet, TreeTagger, General Inquirer e Índice intralingüístico, garantizándose de esta forma varias de las etapas de la minería de opinión.

Esta aplicación resultó ser muy efectiva al clasificar correctamente el 96.5% de 200 opiniones tomadas de foros de discusión de yahoo.es. De esta manera se sugiere el uso de la herramienta para minar los comentarios que se publiquen en la Web referidos a algún tema de interés.

Se pretende extender la aplicación a la minería de opiniones en otros idiomas mediante la definición de ficheros que traduzcan los términos de cualquier idioma al Inglés. Así como incorporar PosNeg Opinion a un sistema más general que integre el resto de las etapas de la minería de opinión.

Referencias

- AMORES, M. Detección no supervisada de la polaridad de las opiniones. Trabajo de Diploma para optar por el título de Licenciado en Ciencia de la Computación. Universidad Central “Marta Abreu” de Las Villas, Santa Clara, 2013.
- ANGULAKSHMI, G.; MANICKACHEZIAN, R. An analysis on opinion mining: techniques and tools. *International Journal of Advanced Research in Computer Communication Engineering*, 2014, 3(7): p. 7483-7487.
- BUCHE, A.; CHADAK, M. B. *et. al.* Opinion mining and analysis: a survey. *International Journal on Natural Language Computing (IJNLC)*, 2013, 2(3): 39-48.
- CHUNG, J. K.-C.; Wu, C.-E. *et. al.* Improve polarity detection of online reviews with bag-of-sentimental-concepts. En: *Proceedings of the 11th ESWC. Semantic Web Evaluation Challenge*. Crete, Greece: Springer, 2014.
- ESULI, A.; SABATIANI, F. Sentiwordnet: A publicly available lexical resource for opinion mining. En: *Proceedings of the 5th Conference on Language Resources and Evaluation (LREC'06)*. Genova: 2006, p. 417-422.
- LIU, B. Sentiment analysis and subjectivity. En: Indurkha, N. y Damerau, F. J. (editores). *Handbook of natural language processing*. United State of America: Chapman and Hall/CRC Press, 2010, p. 627-666.

- MARTÍN-WANTON, T.; PONS-PORRATA, A. *et. al.* Opinion polarity detection – using Word sense disambiguation to determine the polarity of opinions. En: Proceedings of the International Conference on Agents and Artificial Intelligence. Valencia, Spain: DBLP, 2010.
- MILLER, G. WordNet: a lexical database for English. Communications of the ACM, 1995, 38(11): p. 39-41.
- RASHID, A.; Anwer, N. *et. al.* A survey paper: areas, techniques and challenges of opinion mining. IJCSI International Journal of Computer Science Issues, 2013, 10(2): p. 18-31.
- SCHMID, H. TreeTagger: A language independent part-of-speech tagger. [En línea]. TC project at the Institute for Computational Linguistics of the University of Stuttgart. 1995. [Consultado el: 10 de diciembre de 2014] Disponible en: <http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/>
- SEERAT, B.; Azam, F. Opinion mining: Issues and challenges (a survey). International Journal of computer Applications, 2012, 49(9); p. 42-51.
- SHARMA, N. R.; CHITRE, V. D. Opinion mining, analysis and its challenges. International Journal of Innovations & Advancement in Computer Science, 2014, 3(1): p. 59-65.
- STONE, P; PHILIP, J. *et. al.* The General Inquirer: a computer approach to content analysis. Cambridge, Massachusetts, MIT Press, 1966. 651 p.
- STRAPPARAVA, C.; VALITUTI, A. WordNet Affect: an affective extension of WordNet. En: Proceedings of the 4th International Conference on Language Resources and Evaluation. Lisbon, Portugal: ELRA, 2004, p. 1083-1086.
- VESELOVSKÁ, K. Sentence-level polarity detection in a computer corpus. En: Proceedings of the 20th Annual Conference of Doctoral Students. WDS 2011 – Proceedings of Contributed Papers. Prague: MATFYZPRESS, 2011, p. 167-170.
- VINODHINI, G.; CHANDRASEKARAN, R. M. Sentiment analysis and opinion mining: a survey. International Journal of Advanced Research in Computer Science and Software Engineering, 2012, 2(6): p. 282-292.