

# Efectos de la Negación, Modificadores, Jergas, Abreviaturas y Emoticonos en el Análisis de Sentimiento

Mario Amores, Leticia Arco, Abel Barrera

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  
marioamores89@gmail.com, leticiaa@uclv.edu.cu

**Resumen** Algunos problemas presentes en el tratamiento de opiniones son: el uso de lenguaje informal, irónico y sarcástico, las abreviaturas, los errores ortográficos y tipográficos, la semántica composicional, el nivel cultural y conocimiento del lenguaje. Estos problemas imponen mayor dificultad a la minería de opiniones respecto a la minería de textos en general. Por tanto, el objetivo de nuestra investigación es desarrollar soluciones computacionales que permitan resolver varios de estos problemas, contribuyendo a un mejor procesamiento de las opiniones y consecuentemente mayor efectividad en el cálculo de la polaridad. Como resultado de esta investigación se desarrollaron recursos para el manejo de jergas, emoticonos, palabras modificadoras y negaciones. Estos recursos son aplicables en cualquier sistema para minar opiniones en Inglés y Español. El estudio experimental a partir de la aplicación de los recursos propuestos arrojó valores de exactitud y F1 superiores a los obtenidos al calcular la polaridad sin incorporar dichos recursos.

**Palabras Claves:** Minería de opinión, Recursos léxicos, Negación, Emoticonos

**Abstract** Some problems present in the treatment of opinions are: the use of informal, ironic and sarcastic language, abbreviations, orthographic and typographic mistakes, semantic compositionality, the cultural level and knowledge of language. These problems impose greater difficulty on opinion mining than on text mining in general. Therefore, the aim of our research is to develop computational solutions to solve some of these problems, contributing to improve the processing of opinions and consequently to obtain more effective polarity detection. As a result of this research some resources were developed to manage jargons, emoticons, valence shifters and negations. These resources are applicable in any opinion mining system that requires them for mining opinions in Spanish or English. The experimental study from the application of the proposed resources showed values of accuracy and F1 higher than those obtained by calculating the polarity without incorporating those resources.

**Keywords:** Opinion Mining, Lexical Resources, Negation, Emoticons

## 1 Introducción

Las personas al prepararse para tomar alguna decisión siempre tienen en cuenta, además de su propia experiencia, lo que otras personas puedan aportar, por lo que las opiniones son una parte importante en la vida de los seres humanos. La creciente popularidad de sitios ricos en opiniones abre nuevas oportunidades y retos al poderse usar diferentes técnicas para entender y extraer las opiniones de otros.

Las opiniones son los estados subjetivos que reflejan los sentimientos y la percepción de una persona sobre un suceso o un objeto [10]. Según Liu [13], una opinión es una emoción sobre una entidad o un aspecto de la entidad expresada por un usuario. El sentimiento asociado a una opinión puede ser positivo, negativo o neutro, o expresarse con distinta fuerza o niveles de intensidad.

Con la llegada de la web 2.0, enormes volúmenes de textos con opiniones están disponibles. Para extraer el sentimiento acerca de un objeto, producto o servicio de esta enorme red, se necesitan sistemas de minería de opinión automatizados. El procesamiento automático de opiniones no es una tarea sencilla. Algunos de los problemas presentes en el tratamiento de las opiniones son: el uso de lenguaje informal, las abreviaturas, los errores ortográficos y tipográficos, el lenguaje irónico y sarcástico, la semántica composicional, el nivel de conocimiento del lenguaje, el nivel cultural, entre otros. Estos problemas, en comparación con el procesamiento de documentos en otras tareas de la minería de textos, imponen una mayor dificultad a la minería de opiniones. Por tanto, el objetivo de nuestra investigación es desarrollar soluciones computacionales que permitan resolver varias de las problemáticas existentes en las opiniones, contribuyendo a un mejor procesamiento de las mismas y consecuentemente mayor efectividad en el cálculo de la polaridad.

## 2 Recursos para manejar problemas presentes en las opiniones

Algunos de los problemas presentes en las opiniones se pueden manejar con la minería de opinión basada en aspectos y entidades; por ejemplo, identificar las entidades y aspectos y los valores de polaridad asociados a cada uno de ellos, o analizar aquellos textos que contienen otra entidad para referirse a una entidad [13]. Otros problemas requieren realizar análisis sintáctico y desambiguación semántica en aquellos casos donde sea necesario identificar la polaridad de las palabras dependiendo del contexto y de la etiqueta POS (*Part Of Speech*) [13]. El manejo de algunos problemas requiere la aplicación de técnicas de minería de opinión dependientes del dominio, por ejemplo, para manejar aquellos términos que en determinado dominio son considerados positivos y en otros pueden ser considerados negativos. El manejo de estos problemas no es objetivo de este trabajo, ya que requieren la aplicación de técnicas muy específicas del procesamiento del lenguaje natural, desambiguación semántica y minería de opinión basada en aspectos embebidas en el sistema de minería de opinión que se diseñe.

En este epígrafe presentaremos recursos generales para manejar la negación, las palabras modificadoras, las jergas, las abreviaturas y los emoticonos, con independencia del sistema de minería de opinión que se desee aplicar a la colección de opiniones.

## 2.1 Manejo de la negación

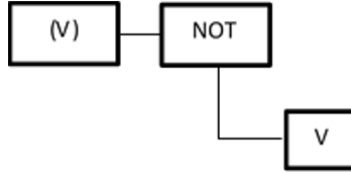
El tratamiento de la negación es un problema abierto dentro del procesamiento del lenguaje natural, y la minería de opinión en particular. Son muy abundantes las opiniones negativas expresadas con términos positivos negados y viceversa [9].

La negación es un elemento lingüístico muy particular de cada lengua, por lo que para un tratamiento efectivo se debe realizar un estudio pormenorizado de los distintos elementos lingüísticos que intervienen en el proceso de negación de un enunciado. La investigación existente sobre la influencia de la negación es escasa, no estudia en profundidad estos aspectos lingüísticos y se centra casi en exclusividad en textos escritos en Inglés [21], aunque existen algunas propuestas para otros idiomas.

Las primeras para manejar las negaciones comenzaron en el año 2001 [6,16,17,11,22]. Entre los trabajos más recientes cabe destacar el sistema basado en reglas propuesto en [8]. En [18] marcan como negadas todas aquellas palabras que se encuentren después de una partícula negativa hasta llegar a un signo de puntuación, un conector o ciertas palabras pertenecientes a una determinada categoría gramatical, e introducen una nueva forma de tratar la negación que consiste en reducir el valor de polaridad de las palabras negadas en lugar de invertirlo. En [7] se presenta un sistema para el tratamiento de la negación y de la especulación, y el primer corpus de opiniones etiquetado a nivel de negación y especulación [12]. El primer trabajo que conocemos para el Español es el propuesto en [4] que consiste en la adaptación al Español del sistema SO-CAL [19]. En [20] manejan la negación teniendo en cuenta la estructura sintáctica del texto, obteniendo así mejores resultados que los sistemas puramente léxicos.

Nuestra propuesta para manejar la negación se basa en la unión de varios de los enfoques antes descritos. Primeramente tomamos las reglas presentadas en [15] para la negación en el idioma Inglés y las modificamos descartando aquellos términos que no aportan en el cálculo de la polaridad y que pudieran ser eliminados en una aplicación de minería de opinión en la etapa de preprocesamiento. El rango de la negación está definido por las reglas. Por ejemplo, dada la oración “I don’t like this laptop” la rama de la regla que se muestra en la Fig. 1 se aplica como se explicará a continuación. Cuando se detecta el verbo “do” seguido del término “not” se activa el árbol de reglas para el término “not”. El siguiente término que aparece en la oración es un verbo, por tanto, se recorre el árbol buscando esta entrada. Si aparece la entrada, como es el caso de este ejemplo que se muestra en la rama derecha que sale del término “not” en la Fig. 1, la polaridad de este término también se niega. A continuación, el próximo término a analizar es un sustantivo, por tanto, se busca en la rama del árbol esta entrada, como no existe, se termina la negación. De manera general, la activación de una

regla consiste en identificar la regla a aplicar y para cada término buscar su clase gramatical e ir moviéndose por las ramas de la regla identificada. Cada término con la clase gramatical de la ramificación correspondiente será negado. Negar un término consiste en consultar su polaridad e invertirla.



**Figura 1.** Rama de la regla de negación para el término “not”.

Estas reglas están creadas para los términos “not”, “never” y “no” del idioma Inglés. Para el idioma Español adaptamos las reglas descritas en [20] para los términos “no”, “nunca” y “sin”. Las reglas solamente involucran los tres términos más utilizados para negar (not, never y no); sin embargo, existen otros términos que expresan negación, de ahí la necesidad de utilizar otros enfoques para manejar la negación. Por ello, se crearon listas de palabras negadoras en ambos idiomas a partir de diversas listas ya publicadas. El rango de negación que consideramos para estas palabras se obtuvo a través de la experimentación con rangos de uno, dos y tres términos después de la palabra negadora, siendo el de mejores resultados el rango que considera negar los próximos dos términos. También se tuvo en cuenta el uso de sufijos y prefijos para ambos idiomas, un ejemplo para el Inglés lo constituye el prefijo dis- (dislike) y para el Español el prefijo in- (incomible). Para evitar que el análisis de los prefijos y sufijos produjera falsos positivos, se tuvo en cuenta que el término modificado tuviera un alto valor de polaridad positiva. En todos los casos la negación se trata invirtiendo los valores de polaridad de cada término negado.

La aplicación de nuestra propuesta para manejar la negación considera el uso de uno u otro enfoque dependiendo del término que se presente en el procesamiento textual. Por ejemplo, si aparece la palabra “no” directamente se activa el enfoque basado en reglas, si el término a analizar presenta un prefijo que expresa negación entonces seguimos el enfoque basado en prefijos y si aparece otro término que expresa negación y que no es manejado desde las reglas, entonces se consulta la lista de palabras negadoras.

## 2.2 Manejo de las palabras modificadoras

Las palabras modificadoras son aquellas que pueden incrementar, reducir o cambiar la polaridad de otro término dentro de la oración. Especial atención tienen las palabras intensificadoras las cuales aumentan la fuerza de otros términos y las atenuadoras que reducen la fuerza de otros términos.

En gramática, un intensificador es una palabra que hace hincapié en otra palabra o frase. También conocido como un refuerzo o un amplificador. Por ejemplo, algo no sólo es bueno, sino muy bueno, o incluso, terriblemente bueno. Los adjetivos intensificadores modifican a los sustantivos; los adverbios de intensificación modifican comúnmente verbos, adjetivos graduables, y otros adverbios.

Hasta cierto punto, un intensificador actúa como una señal: anuncia que la siguiente palabra que se lleva a cabo debe ser entendida como insuficiente. Algunos ejemplos de palabras intensificadoras en el idioma Inglés aparecen en las siguientes oraciones: “The woodwind has a slightly greater scope than the violin”, “A really good detective never gets married” y “Maggy knew that something surpassingly important had happened to him”.

Otras palabras modificadoras son las palabras atenuadoras, como por ejemplo, el término “less” en Inglés o “poco” en Español. Los atenuadores (también mitigadores) han sido descritos como elementos que se utilizan para mitigar o relativizar el impacto de las afirmaciones, en una estrategia propia de la cortesía [5]. Se dice también que “desactivan” la fuerza elocutiva de un acto [14]. Pragmáticamente, la atenuación sirve para regular la relación interpersonal y social entre interlocutores.

Los diversos autores que estudian en forma sistemática este tipo de fenómenos se refieren alternativamente a estrategias, recursos, operadores, formas o estructuras de atenuación [19]. Desde el punto de vista semántico-pragmático, la atenuación se opone a la intensificación. En ambos casos los recursos pueden corresponder a ámbitos categoriales diferentes, como adverbios, adjetivos y frases.

Para tratar estas palabras modificadoras, como resultado de esta investigación, se construyeron dos listas (una para el idioma Inglés y otra para el Español) de términos intensificadores y atenuadores a partir de las consultas manuales a diccionarios, análisis de listas de palabras modificadoras ya existentes y consultas a expertos lingüistas. Como estas palabras pueden influir tanto en los términos precedentes como en los posteriores, se les agregó a cada uno una etiqueta para identificar donde ejercía éste su influencia. El formato consiste en la frase o el término modificador seguido de un valor entero en el dominio  $\{1, 2\}$ , donde uno significaba que es una palabra intensificadora y dos que es una palabra atenuadora. A continuación aparece un segundo valor entero en el dominio  $\{1, 2, 3\}$  que señala a qué término afecta siendo uno el término antecesor, dos el sucesor y tres para ambos casos. Un fragmento de una lista de palabras modificadoras en Inglés se muestra en la Fig. 2. Sugerimos que aquellos sistemas que utilicen nuestra propuesta para el manejo de palabras modificadoras de la polaridad, le agreguen uno al valor de polaridad predominante del término predecesor o sucesor según sea la etiqueta que porte la palabra intensificadora, y reduzcan en un 50 % el valor de polaridad predominante del término predecesor o sucesor en el caso de las palabras atenuadoras. Por tanto, ante la presencia de intensificadores le asignamos a las palabras modificadas el mayor valor de polaridad, mientras que a aquellas modificadas por atenuadoras lo que se pretende es reducir su va-

lor de polaridad de forma tal que si era medianamente fuerte, será débil y si era fuerte, será medianamente fuerte.

1	a bit	2	1		
2	a little bit			2	3
3	a lot	1	1		
4	a lot of		1	1	
5	abruptly		1	3	
6	absolute		1	3	
7	absolutely		1	3	
8	abundantly		1	2	

**Figura 2.** Fragmento de lista de palabras modificadoras.

### 2.3 Manejo de jergas, abreviaturas y emoticonos

Las jergas de internet son una variedad de lenguajes cotidianos y abreviaciones en la escritura utilizadas por las diferentes comunidades en el ámbito de Internet [23]. Esos nuevos términos o formas de escritura a menudo se han originado con el propósito de abreviar pulsaciones de teclado, sobre todo por la dificultad y/o rapidez que a veces trae aparejado el uso de algunos dispositivos portátiles. Muchas personas utilizan las mismas abreviaturas en los mensajes de textos y en la mensajería instantánea, así como en las redes sociales [6]. Por ejemplo “u” para “you”, “cam” para “camera”, “pic” para “picture”, “f9” para “fine”, “b4” para “before”, “gud” para “good”, etc.

De igual forma, los acrónimos, símbolos del teclado, y abreviaturas, también son de uso frecuente y más o menos normalizado, y por tanto, integran lo que puede llamarse jerga de Internet. Por ejemplo “gr8” es una palabra del argot usado para “great”, “5n” o “fyn” para “fine”. Estas palabras no son una parte de los diccionarios tradicionales, pero se encuentran ampliamente en los textos en línea. Si estas jergas se pueden asignar a las palabras originales, los resultados de rendimiento de los analizadores de los sentimientos pueden ser mejorados.

Otro elemento que encontramos dentro de las jergas de internet son los emoticonos (emoticon) que son una secuencia de caracteres ASCII que, en un principio, representaba una cara humana y expresaba una emoción. Posteriormente, fueron creándose otros emoticonos con significados muy diversos.

Existen varias listas que incluyen jergas y emoticonos; sin embargo, cada una adolece de términos que presentan otras, sobre todo debido a la velocidad y dinámica con que surgen nuevos emoticonos y jergas. De ahí que para tratar con estos dos temas creamos dos listas, una para las jergas, y otra para los emoticonos, construidas a partir de varias listas previamente publicadas<sup>1</sup>. Cada emoticón o jerga tiene asignados sus correspondientes valores de polaridad positiva y negativa. Para obtener estos valores de polaridad de las jergas utilizamos

<sup>1</sup> [www.netlingo.com](http://www.netlingo.com), [www.noslang.com](http://www.noslang.com), [es.wikipedia.org/wiki/Anexo:Emoticonos](http://es.wikipedia.org/wiki/Anexo:Emoticonos)

su correspondiente significado en el idioma y sumamos los valores de polaridad de sus términos. Los valores de polaridad de los términos pueden obtenerse a partir de listas<sup>2</sup> o recursos léxicos como por ejemplo SentiWordNet [3] o Pos-Neg\_Lexicon [1,2]. En la Fig. 3 se muestra un fragmento de la lista de jergas y los valores de polaridad positiva y negativa asignados. Para los emoticonos es similar, solo que se le asigna a cada emoticón el valor de polaridad de la emoción que expresan. En ambos casos se sugiere que los valores de polaridad sean agregados al valor final de toda la opinión, pues estos elementos la mayoría de las veces no afectan directamente a un término sino al contexto en general.

np	"ningún problema" o "no hay problema"	0.6585	0.0115
sdr	"dulces sueños"	0.6486	0.0208
thx	"agradecer a usted" o "gracias"	0.6227	0.0
wombat	"pérdida de dinero, tiempo, y cerebro"	0.0159	0.6663

**Figura 3.** Fragmento de lista de jergas y los valores de polaridad asignados.

Nuestra propuesta sugiere que después de cargar el corpus textual es necesario localizar los emoticonos y las jergas para pre-procesarlos y determinar su influencia en el cálculo de la polaridad. Este es un paso que distingue la minería de opinión de otros enfoques de minería de textos donde en la etapa de pre-procesamiento son eliminados inicialmente los emoticonos y las jergas. Después de calcular los valores de polaridad asociados a las jergas y los emoticonos, estos deben ser eliminados, ya que no aparecerán directamente en los recursos léxicos a consultar, en el caso que estemos usando un enfoque no supervisado basado en recursos léxicos. El análisis de las palabras intensificadoras y atenuadoras debe realizarse cuando estemos asignando los valores de polaridad a los términos, al concluir este proceso es posible aplicar la estrategia híbrida definida para manejar la negación.

### 3 Evaluación de los recursos desarrollados

En este epígrafe se desea mostrar las potencialidades que ofrecen los recursos desarrollados para el manejo de jergas, emoticonos, negaciones y el uso de palabras modificadoras en el cálculo de la polaridad de las opiniones. Para ello, utilizamos el sistema PosNeg Opinion que permite detectar la polaridad de opiniones en Español y en Inglés siguiendo un esquema general que consta de seis etapas para el cálculo de la polaridad [1]. Ya seleccionado el sistema a utilizar, fue necesario conformar tres casos de estudios que permitan ilustrar la efectividad de nuestra propuesta. Los dos primeros están conformados por subconjuntos de opiniones del corpus Amazon<sup>3</sup> y el tercero coincide con el corpus Sentiment140<sup>4</sup>.

<sup>2</sup> <http://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html>

<sup>3</sup> <http://times.cs.uiuc.edu/~wang296/Data/>

<sup>4</sup> <http://cs.stanford.edu/people/alecmgo/trainingandtestdata.zip>

El corpus Amazon contiene comentarios sobre seis productos (cámaras, teléfonos móviles, TV, laptop, tablet y sistemas de video vigilancia) en Amazon.com. Cada producto contiene los siguientes atributos: ID del producto (único), nombre, características del producto, precio y el URL de la imagen. Cada comentario también contiene además los siguientes atributos: ID del comentario (único), autor, título, contenido, puntuación global y fecha. La puntuación global es brindada por el usuario quien le da una puntuación de cero a cinco estrellas, para este trabajo solo tomamos las puntuaciones de cero a dos como negativas, de cuatro a cinco como positivas, dejando fuera las puntuaciones de tres estrellas pues éstas muestran indecisión de los usuarios al clasificar sus opiniones, lo cual no asegura que sea una clasificación confiable. El otro corpus utilizado es Sentiment140 que incluye 800 000 tweets que contienen emoticonos positivos y 800 000 tweets que incluyen emoticonos negativos, aquellos mensajes que contenían ambos no fueron tomados en cuenta.

El cálculo de la polaridad de las opiniones puede ser analizado como un proceso de clasificación, donde se determina si la opinión pertenece a la clase positiva o negativa. Las medidas para evaluar el desempeño de los clasificadores pueden ser utilizadas para evaluar la calidad de la detección de la polaridad de las opiniones. Las medidas de evaluación utilizadas en la experimentación son: Exactitud (*Accuracy*; Ac) que calcula la proporción del número total de predicciones correctas; la Razón de Verdaderos Positivos (*True Positive Rate*; TPR) o *Recall* que calcula la proporción de casos positivos que fueron correctamente identificados; la Razón de Falsos Positivos (*False Positive Rate*; FPR) que calcula la proporción de casos negativos que han sido incorrectamente clasificados; la Razón de Verdaderos Negativos (*True Negative Rate*; TNR) que calcula la proporción de casos negativos que han sido correctamente clasificados; la Razón de Falsos Negativos (*False Negative Rate*; FNR) que calcula la proporción de casos positivos que fueron incorrectamente clasificados; la Precisión (*Precision*; P) que calcula la proporción de casos predichos positivos que fueron correctos y la medida-F (*F-measure*) que determina un valor único ponderado de la precisión y la exactitud.

Para verificar el impacto que tienen las palabras intensificadoras y atenuadoras en la clasificación de la polaridad tomamos del corpus Amazon aquellas opiniones que contuviesen dichas palabras quedando así una colección de 99 148 opiniones negativas y 645 678 opiniones positivas. En la Tabla 1 se muestra como al realizar el análisis de dichas palabras se logran mejores medidas de Precisión, Recall, Exactitud y F1, respecto al procesamiento omitiendo dichas palabras.

**Tabla 1.** Impacto de las palabras intensificadoras y atenuadoras.

PosNeg Opinion	Precisión	Recall	Exactitud	F1
Sin intensificadores y atenuadores	89 %	83 %	82 %	86 %
Con intensificadores y atenuadores	92 %	87 %	83 %	90 %



Para mostrar el impacto que tiene la negación en la clasificación de la polaridad utilizamos del corpus Amazon aquellas opiniones que utilizan palabras negadoras quedando así una colección de 65 578 opiniones negativas y 225 758 opiniones positivas. En la Tabla 2 se muestra como al realizar el análisis de la negación se logran mejores medidas de Precisión, Recall, Exactitud y F1.

**Tabla 2.** Impacto del manejo de las negaciones en la detección de la polaridad.

<b>PosNeg</b>	<b>Opinion</b>	<b>Precisión</b>	<b>Recall</b>	<b>Exactitud</b>	<b>F1</b>	<b>TPR</b>	<b>FPR</b>	<b>TNR</b>	<b>FNR</b>
Sin negación		88 %	86 %	80 %	87 %	86 %	51 %	49 %	14 %
Con negación		90 %	88 %	83 %	89 %	88 %	42 %	58 %	12 %

También es posible notar en la Tabla 2 que al realizar este análisis se logra una mayor proporción de casos positivos y negativos que fueron correctamente identificados (TPR y TNR) y una disminución de la proporción de casos negativos que han sido incorrectamente clasificados como positivos (FPR), así como una disminución de casos positivos que han sido incorrectamente clasificados como negativos (FNR).

Para verificar el impacto que tienen los emoticonos en la clasificación de la polaridad tomamos el corpus Sentiment140 ya que parte de éste fue construido a partir de la búsqueda de diferentes emoticonos. En la Tabla 3 se muestra como al realizar el análisis de este corpus se obtienen mejores resultados de Precisión, Recall, Exactitud y F1 al incluir el análisis de los emoticonos.

**Tabla 3.** Impacto del manejo de los emoticonos en la detección de la polaridad.

<b>PosNeg</b>	<b>Opinion</b>	<b>Precisión</b>	<b>Recall</b>	<b>Exactitud</b>	<b>F1</b>
Sin emoticonos		83 %	90 %	77 %	86 %
Con emoticonos		88 %	92 %	83 %	90 %

Para el análisis del impacto de las jergas también utilizamos el corpus Sentiment140 ya que este se encuentra construido sobre la red social Twitter donde es común la utilización de estos elementos. En la Tabla 4 se muestra cómo realizar el análisis de las jergas ayuda a la obtención de mejores resultados de Precisión, Exactitud y F1, y el valor de Recall se mantiene. El manejo de las jergas es complejo, ya que la jerga en sí puede generar otros de los problemas presentes en las opiniones.

**Tabla 4.** Impacto del manejo de las jergas en la detección de la polaridad.

<b>PosNeg</b>	<b>Opinion</b>	<b>Precisión</b>	<b>Recall</b>	<b>Exactitud</b>	<b>F1</b>
Sin jergas		83 %	90 %	77 %	86 %
Con jergas		87 %	90 %	80 %	88 %

## 4 Conclusiones y recomendaciones

Como resultado de esta investigación se desarrollaron recursos para tratar algunos problemas presentes en las opiniones. Estos recursos son aplicables en cualquier sistema para la detección de la polaridad de las opiniones en Inglés y Español, incluso, no solo para detectar la polaridad, sino para otras tareas de la minería de opinión que así lo requieran. Estos recursos fueron incorporados al sistema PosNeg Opinion 3.0 [1] para ilustrar las ventajas que ofrecen estos recursos en el cálculo de la polaridad de las opiniones de manera no supervisada. Al incorporar el manejo de jergas, emoticonos, palabras modificadoras y negaciones a PosNeg Opinion se obtuvieron valores promedios de 82 % de exactitud y 89 % de F1, superándose de esta forma los valores de calidad de la detección de la polaridad de las opiniones que ofrecía PosNeg Opinion antes de incorporarle estos recursos. De esta forma se mostró el impacto positivo que tiene el manejo de las palabras modificadoras, las negaciones, los emoticonos y las jergas en el cálculo de la polaridad mediante el uso de los recursos diseñados como resultado de esta investigación a tales efectos. Se recomienda continuar actualizando los recursos desarrollados a partir de la retroalimentación de los resultados de la experimentación y proponer soluciones para otros problemas como el sarcasmo y la ironía.

## Referencias

1. Amores, M.: Detección de la polaridad de las opiniones basada en nuevos recursos léxicos. Master en ciencia de la computación, Universidad Central "Marta Abreu" de Las Villas (2016)
2. Amores, M., Arco, L., Borroto, C.: Unsupervised Opinion Polarity Detection based on New Lexical Resources. *Computación y Sistemas* 20(2) (2016)
3. Baccianella, S., Esuli, A., Sebastiani, F.: SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. In: *LREC*. vol. 10, pp. 2200–2204 (2010)
4. Brooke, J., Tofiloski, M., Taboada, M.: Cross-Linguistic Sentiment Analysis: From English to Spanish. In: *RANLP*. pp. 50–54 (2009)
5. Caffi, C.: On mitigation. *Journal of pragmatics* 31(7), 881–909 (1999)
6. Das, S., Chen, M.: Yahoo! for amazon: Extracting market sentiment from stock message boards. In: *Proceedings of the Asia Pacific finance association annual conference (APFA)*. vol. 35, p. 43. Bangkok, Thailand (2001)
7. Díaz, N.P.C.: Detección de la Negación y la Especulación en Textos Médicos y de Opinión. Phd, Universidad de Huelva (2014)
8. Jia, L., Yu, C., Meng, W.: The effect of negation on sentiment analysis and retrieval effectiveness. In: *Proceedings of the 18th ACM conference on Information and knowledge management*. pp. 1827–1830. ACM (2009)
9. Jiménez Zafra, S.M., Martínez Cámara, E., Martín Valdivia, M.T., Molina González, M.D.: Tratamiento de la negación en el análisis de opiniones en español. *Procesamiento del Lenguaje Natural* (54), 37–44 (2015)
10. Kaur, A., Duhan, N.: A survey on sentiment analysis and opinion mining. *International Journal of Innovations & Advancement in Computer Science IJIACS* 4 (2015)

11. Kennedy, A., Inkpen, D.: Sentiment classification of movie reviews using contextual valence shifters. *Computational intelligence* 22(2), 110–125 (2006)
12. Konstantinova, N., De Sousa, S.C.M., Díaz, N.P.C., López, M.J.M., Taboada, M., Mitkov, R.: A review corpus annotated for negation, speculation and their scope. In: LREC. pp. 3190–3195 (2012)
13. Liu, B.: Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies* 5(1), 1–167 (2012)
14. Mallarino, M.d.R.U.: Los marcadores del discurso y la cortesía verbal en español. *Tintas. Quaderni di Letterature iberiche e iberoamericane* (1), 300–303 (2012)
15. Na, J.C., Khoo, C., Wu, P.H.J.: Use of negation phrases in automatic sentiment classification of product reviews. *Library Collections, Acquisitions, and Technical Services* 29(2), 180–191 (2005)
16. Pang, B., Lee, L., Vaithyanathan, S.: Thumbs up?: sentiment classification using machine learning techniques. In: *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10*. pp. 79–86. Association for Computational Linguistics (2002)
17. Polanyi, L., Zaenen, A.: Contextual valence shifters. In: *Computing attitude and affect in text: Theory and applications*, pp. 1–10. Springer (2006)
18. Taboada, M., Brooke, J., Tofiloski, M., Voll, K., Stede, M.: Lexicon-based methods for sentiment analysis. *Computational linguistics* 37(2), 267–307 (2011)
19. Taboada, M., Voll, K., Brooke, J.: Extracting sentiment as a function of discourse structure and topicality. Simon Fraser Univeristy School of Computing Science Technical Report (2008)
20. Vilares, D., Alonso, M.A., Gómez-Rodríguez, C.: Clasificación de polaridad en textos con opiniones en español mediante análisis sintáctico de dependencias. *Procesamiento del lenguaje natural* 50, 13–20 (2013)
21. Wiegand, M., Balahur, A., Roth, B., Klakow, D., Montoyo, A.: A survey on the role of negation in sentiment analysis. In: *Proceedings of the workshop on negation and speculation in natural language processing*. pp. 60–68. Association for Computational Linguistics (2010)
22. Wilson, T., Wiebe, J., Hoffmann, P.: Recognizing contextual polarity in phrase-level sentiment analysis. In: *Proceedings of the conference on human language technology and empirical methods in natural language processing*. pp. 347–354. Association for Computational Linguistics (2005)
23. Yin, Y.: World Wide Web and the Formation of the Chinese and English “Internet Slang Union”. *Computer-Assisted Foreign Language Education* 1 (2006)