

Tratamiento sintáctico de la negación en análisis del sentimiento monolingüe y multilingüe*

Syntactic treatment of negation for monolingual and multilingual sentiment analysis

David Vilares, Miguel Alonso y Carlos Gómez-Rodríguez

Universidade da Coruña

Grupo LYS

Departamento de Computación, Facultad de Informática

Campus de Elviña, 15071 A Coruña (España)

{david.vilares, miguel.alonso, carlos.gomez}@udc.es

Resumen: Tratar adecuadamente la negación es importante para obtener sistemas de análisis del sentimiento de alto rendimiento. En este marco, presentamos un método de tratamiento de la negación en español que explota la información sintáctica obtenida a partir de un analizador de dependencias entrenado en el corpus Ancora junto con un conjunto de reglas específicas para el español. A continuación presentamos un sistema que generaliza este método a un entorno multilingüe en el que se explota el marco armonizado de anotación que proporcionan las Dependencias Universales, junto con un formalismo genérico para definir las reglas que determinan el alcance de la negación. Los resultados experimentales avalan la efectividad del método, tanto en entornos monolingües como multilingües.

Palabras clave: Negación, Análisis sintáctico, Análisis del sentimiento, Minería de opiniones, Multilingüismo

Abstract: Dealing with negation in a proper way is a relevant factor in order to obtain high-performance sentiment analysis systems. In this framework, we present a method for the treatment of negation in Spanish that exploits the syntactic information obtained from a dependency parser trained on the Ancora corpus, together with a set of specific rules. Next, we generalize this method to a multilingual environment in which the harmonized framework of annotation provided by the Universal Dependencies is exploited, along with a generic formalism to define language-independent rules that determine the scope of negation. The experimental results support the effectiveness of the method, both on monolingual and multilingual environments.

Keywords: Negation, Parsing, Sentiment analysis, Opinion mining, Multilingualism

1 Introducción

El análisis del sentimiento o minería de opiniones es un área de investigación centrada en determinar automáticamente si en un texto se opina o no, si la polaridad o sentimiento que se expresa en él es positiva, negativa o mixta; y en extraer automáticamente la percepción de un autor sobre aspectos concretos de un tema. Los enfoques más habituales de

análisis del sentimiento consideran un texto como una secuencia de palabras y realizan un procesado superficial de las oraciones que no tiene en cuenta las relaciones sintácticas entre palabras.

En contraste con las propuestas léxicas habituales, nuestro método de análisis del sentimiento se basa en explotar la estructura sintáctica de la frase. Para ello, primero se procede a segmentar el texto en oraciones y a dividir cada una de ellas en *tokens* (principalmente palabras, pero también signos de puntuación, números, etc.) para después realizar la etiquetación morfosintáctica de cada uno de los elementos del texto. El siguiente paso consiste en realizar un análisis sintáctico de dependencias mediante el cual se identifican

* Investigación parcialmente financiada por el European Research Council (ERC) en el marco del programa de investigación e innovación Horizon 2020 de la Unión Europea (FASTPARSE Starting Grant, grant agreement No 714150), por el Ministerio de Economía, Industria y Competitividad (proyecto FFI2014-51978-C2-2-R) y por el Ministerio de Educación, Cultura y Deporte (FPU13/01180).

relaciones binarias padre/dependiente entre los términos de una oración. Cada uno de esos vínculos binarios constituye una dependencia, que se anota con la función sintáctica que relaciona los dos términos. La estructura obtenida se denomina árbol de dependencias.

2 Detección del alcance de la negación en español

La manera más común y simple de negar un fragmento de texto en español es mediante el adverbio *no*, aunque también es frecuente utilizar términos como *sin* o *nunca*. Algunos tipos de oraciones requieren el uso de más de una palabra negativa para dar lugar a una oración negativa, como ocurre habitualmente con las palabras *nada*, *ninguno* o *nadie* que son comúnmente precedidas por *no*. Por otra parte, no siempre es fácil diferenciar entre el carácter negador y el carácter intensificador¹ de una determinada palabra. Por ejemplo, palabras como *apenas* o *casi* podrían ser clasificadas en cualquiera de estas dos categorías. En nuestro sistema de análisis del sentimiento hemos optado por considerar este tipo de expresiones como intensificadores y por lo tanto sólo consideramos explícitamente como negadores los adverbios *no*, *nunca* y *sin*, que cubren la mayor parte de frases negativas (Jiménez-Zafra et al., en prensa).

Hay varias maneras de tomar en cuenta el efecto de la negación. Por ejemplo, los métodos de procesamiento textual basados en aprendizaje automático tienden a unificar el negador y la palabra negada en un solo término (de esta manera *no es bueno* se convierte en *no_es_bueno*) (Sidorov et al., 2013). Otro recurso habitual es cambiar la orientación semántica de un cierto número de palabras después del negador (Yang, 2008; Fernández Anta et al., 2012). Nuestro enfoque es más completo puesto que se basa en utilizar la estructura sintáctica de la frase para determinar el alcance de la negación.

En el caso particular del español tomamos como base la estructura sintáctica definida por el corpus Ancora (Taulé, Martí, y Recasens, 2008). En lo que respecta al tratamiento de la negación, la estructura sintáctica utilizada en Ancora para representar un adverbio *sin* nos asegura que su nodo secundario debe

¹En el contexto de este trabajo englobaremos bajo el término *intensificadores* tanto aquellas palabras que aumentan la intensidad del significado como aquellas que lo disminuyen.

ser el ámbito de la negación, sin necesidad de analizar el tipo de dependencia, pero no podemos asumir lo mismo para los negadores *no* y *nunca*, ya que al ser representados habitualmente como nodos hoja, el ámbito candidato de negación siempre implica un nodo principal o una colección de nodos hermanos, por lo que se requiere un algoritmo más complejo para su tratamiento. En particular, extendemos el procedimiento de Jia, Yu, y Meng (2009) con el fin de ser capaces de identificar el alcance de cada negación a partir del árbol de análisis y una colección de reglas: cuando un token tiene un negador *no* o *nunca* como nodo hijo y es una dependencia de tipo **neg** o **mod**, probaremos las siguientes reglas heurísticas sintácticas en el orden en que se indican, teniendo en cuenta que sólo se utiliza la primera regla aplicable:

- *Regla del padre subjetivo*: Si el nodo padre de un término de negación tiene sentimiento, solo se niega ese nodo. Por ejemplo, en la frase *no elogia mi trabajo*, la negación *no* depende de *elogia*, que tiene una connotación subjetiva, por lo que identificaremos esta palabra como el alcance de la negación.
- *Regla del atributo/objeto directo*: Si una rama al mismo nivel que un nodo de negación está etiquetada con una dependencia de tipo atributo (**atr**) (por ejemplo, *buena* en *la comida no es buena*) u objeto directo (**cd**) (por ejemplo, *bien* en *la pantalla no se ve bien*), se identifica esa rama como el alcance de la negación.
- *Regla del complemento circunstancial*: Si un término de negación tiene una rama de complemento circunstancial (**cc**) al mismo nivel, se identifica esa rama como el alcance de la negación. Si hay más de un complemento circunstancial, sólo se utiliza el primero. Por ejemplo, en la frase *no funciona correctamente los viernes*, se toma el complemento circunstancial de modo *correctamente* como el alcance de la negación.
- *Regla por defecto*: Cuando ninguna de las reglas anteriores es aplicable, se considera que el ámbito de la negación lo constituyen las ramas hermanas del nodo negador.

Una vez que hemos identificado el alcance de la negación, podemos determinar el im-

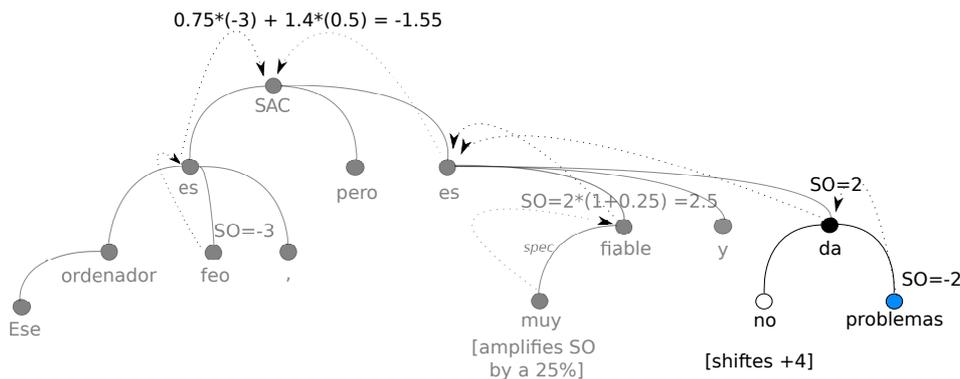


Figura 1: Ejemplo de evaluación de la orientación semántica de una oración con negación

pacto que esta negación tendrá en la determinación de la orientación semántica de esa parte de la oración.

3 Modificación de la orientación semántica por la negación

Ilustramos el proceso mediante un ejemplo. En la parte resaltada de la Figura 1 podemos observar que la partícula de negación *no* tiene como padre el verbo *da*. Primero se intenta aplicar la regla del padre subjetivo, lo cual no es posible ya que *da* no tiene carácter subjetivo, por lo que se pasa a la regla siguiente, la del objeto directo. Esta regla sí se puede aplicar dado que hay un objeto directo (identificado por la dependencia *cd*) al mismo nivel que *no*, por lo que se establece *problemas* como el alcance de la negación. Como consecuencia, el fragmento *no da problemas* pasa de tener una connotación negativa (indicada mediante el valor de orientación semántica -2 de la palabra *problemas*) a positiva (indicada mediante el valor de orientación semántica 2 que se asigna al fragmento). Podríamos pensar que lo que ha ocurrido es que la negación ha invertido el valor de la orientación semántica de la palabra *problemas*, sin embargo esto no es exactamente así.

Ciertamente la manera más simple de negar una palabra consiste en invertir su orientación semántica (OS). Por ejemplo, si $OS(\text{bueno})=2$ entonces $OS(\text{no bueno})=-2$. El principal inconveniente de este método es que no se corresponde con la intuición humana. Por ejemplo, si la orientación semántica de *fascinante* es 5, el sentimiento de *no fascinante* pasaría a ser -5 (muy negativo), cuando incluso podría considerarse que *no fascinante* es una expresión ligeramente positiva. Para evitar este inconveniente, en nuestro sistema hemos optado por *desplazar* la polaridad

en una cantidad fija: siguiendo la propuesta de Taboada et al. (2011), hemos dado un valor de desplazamiento 4 a los adverbios *no* y *nunca*. Así, *no fascinante* pasaría a tener un valor de 1 (ligeramente positivo) mientras que en el ejemplo de la Figura 1 *no da problemas* pasa a tener un valor de orientación semántica 2, resultado de sumar 4 al valor -2 de *problemas*. Para el adverbio *sin* se ha optado un valor de desplazamiento de 3.5, ya que este tipo de negación es menos potente.

Los detalles sobre la evaluación del resto de la oración de la Figura 1 (en particular el tratamiento de intensificadores y cláusulas adversativas) están disponibles en (Vilares, Alonso, y Gómez-Rodríguez, 2015).

4 Evaluación en español

La evaluación de nuestro sistema de análisis del sentimiento para español se ha realizado sobre el SFU Spanish Reviews Corpus (Brooke, Tofiloski, y Taboada, 2009), una colección de 400 reseñas en español sobre automóviles, hoteles, lavadoras, libros, teléfonos, música, ordenadores y películas, obtenidas del sitio web *ciao.es*. Cada categoría tiene un total de 25 reseñas favorables y 25 desfavorables. Los textos contienen palabras mal acentuadas, abreviaturas no reconocidas y frases agramaticales. Esto nos permite evaluar nuestra propuesta en un entorno real y complejo.

La Tabla 1 muestra el desempeño de nuestro sistema activando diferentes opciones. Es interesante señalar que una de las mejoras más importantes proviene del tratamiento de la negación. Como podemos ver, antes de incorporar esta mejora nuestro enfoque tendía hacia las clasificaciones positivas. Esto sucede probablemente como resultado de la tendencia humana a usar lenguaje positivo (Kennedy y Inkpen, 2006). En este sentido, es

Categoría	Exactitud en negativos	Exactitud en positivos	Exactitud global
Sistema base	0.310	0.925	0.618
+intensificación	0.450	0.870	0.660
+cláusulas adversativas	0.455	0.885	0.670
+negación	0.745	0.765	0.755
Sistema final	0.740	0.830	0.785

Tabla 1: Rendimiento del sistema para español en el SFU Spanish Reviews Corpus

común que se utilice la negación de frases positivas para expresar una opinión desfavorable. Por ejemplo, es habitual utilizar expresiones como *no es bueno* en lugar de *malo*, o *no me gusta* en lugar de *feo*. Incluso después de procesar los términos de negación, algunos sistemas léxicos (Taboada et al., 2011) incrementan en un 50% la orientación semántica final de cualquier expresión negativa para superar ese sesgo positivo, mejorando su rendimiento en torno al 6% con esta estrategia. Sin embargo, en nuestro caso la utilización de este proceso en el sistema final no aporta apenas beneficio. Esto parece sugerir que nuestro manejo de la negación es adecuado, al menos en un contexto general.

5 Negación en entornos multilingües

Una vez definido el manejo de la negación en español, nos propusimos definir un método que nos permitiese manejar la negación en un entorno multilingüe. El objetivo no es tener un sistema específico para cada idioma sino aplicar etiquetadores y analizadores sintácticos multilingües para de esta manera obtener, dado un texto en el que se entremezclen varios idiomas (Vilares, Alonso, y Gómez-Rodríguez, 2016a), una representación sintáctica única para todo el texto a la que se puedan aplicar una serie de reglas genéricas. Para ello hemos de recurrir a corpus anotados armónicamente en varias lenguas. En particular, consideramos que la mejor opción actualmente consiste en utilizar los treebanks desarrollados en el marco de las Universal Dependencies (McDonald et al., 2013), a partir de los cuales construimos analizadores sintácticos realmente multilingües siguiendo la metodología definida por Vilares, Alonso, y Gómez-Rodríguez (2016b).

Las reglas genéricas de gestión de la negación operan sobre el árbol sintáctico obtenido. Su genericidad se deriva del hecho de que son parametrizables mediante una función de

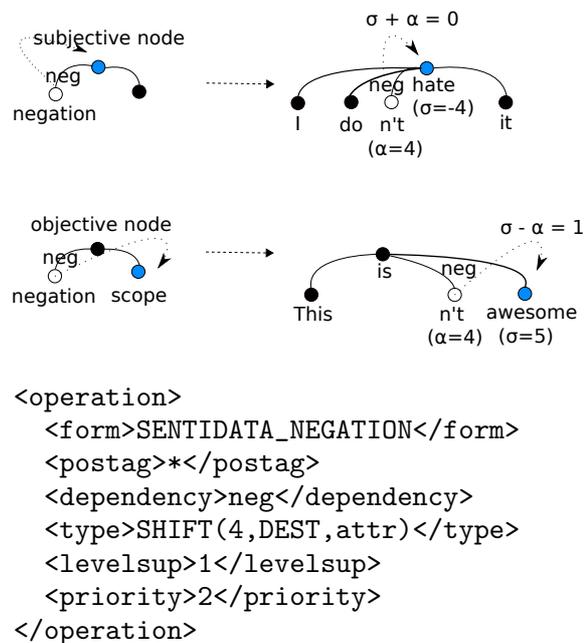


Figura 2: Arriba: esqueletos de las reglas genéricas de tratamiento de la negación, con un ejemplo para cada una de ellas (σ se refiere a la orientación semántica y α el valor de desplazamiento de σ por la negación). Abajo: fragmento XML que describe estas dos reglas, donde SHIFT es una τ que indica que se aplique $\alpha = 4$ a la σ del nivel superior.

transformación τ que indica el cambio que se ha de producir en la orientación semántica de un fragmento de texto; el número de niveles que debemos ascender en el árbol para aplicar τ ; y la función que determina el alcance de τ . Los detalles técnicos completos están disponibles en (Vilares, Gómez-Rodríguez, y Alonso, 2017). Por brevedad, ilustramos algunos casos relevantes de negación mediante la Figura 2. Las reglas genéricas que describen tal comportamiento se traducen en la práctica en fragmentos XML del archivo de configuración, como el que se muestra en la parte inferior de la Figura 2.

Como lexicones de subjetividad, usamos

los mismos diccionarios usados por SO-CAL (Taboada et al., 2011) tanto para el inglés (2 252 adjetivos, 1 142 sustantivos, 903 verbos, 745 adverbios y 177 intensificadores) como para el español (2 049 adjetivos, 1 333 sustantivos, 739 verbos, 594 adverbios y 165 intensificadores). Para el alemán, utilizamos los diccionarios alemanes de SentiStrength (Momtazi, 2012) (2 677 raíces y 39 intensificadores). Todos ellos son recursos disponibles libremente y son utilizados tal cual en nuestro sistema, sin realizar ningún cambio. La lista de emoticones de SentiStrength también se utiliza como recurso léxico. Si un término no aparece en ninguno de estos recursos, no tendrá ningún impacto en el cálculo de la orientación semántica. El sistema completo se puede descargar de <http://www.grupolys.org/software/UUUSA/>

6 Evaluación en análisis del sentimiento multilingüe

Comparamos nuestro sistema multilingüe con otros disponibles en inglés, español y alemán. Como medida de evaluación utilizaremos la exactitud o *accuracy*, ya que es la más adecuada para medir el desempeño de los clasificadores cuando los corpus elegidos son equilibrados y porque ya existen resultados publicados de esta métrica para los sistemas seleccionados.

La evaluación en el caso del inglés se realiza sobre los siguientes corpus estándar:

- Corpus de Taboada y Grieve (2004): Una colección de 400 reseñas largas (200 positivas, 200 negativas) sobre hoteles, películas, ordenadores o música, entre otros temas, extraídas de *epinions.com*.
- Corpus de Pang y Lee (2004): Una colección de 2 000 reseñas largas de películas (1 000 positivas, 1 000 negativas).
- Corpus de Pang y Lee (2005): Una colección de reseñas cortas (oraciones) de películas. En particular, se utilizó la división del corpus utilizada por Socher et al. (2013), eliminando los neutrales (tal y como indican ellos mismos) para la tarea de clasificación binaria (lo que da un total de 1 821 oraciones con opiniones).

Para mostrar las capacidades multilingües de nuestro sistema incluimos una evaluación para el español utilizando el corpus presentado por Brooke, Tofiloski, y Taboada (2009)

(200 reseñas positivas y 200 negativas de *ciao.es*), así como un conjunto de 2 000 reseñas (1 000 positivas y 1 000 negativas) en alemán extraídas de Amazon.

6.1 Comparación con otros sistemas no supervisados

La disponibilidad de corpus y otros sistemas de análisis del sentimiento no supervisados para el inglés y el español nos permite realizar una comparación más rica que en el caso del alemán, donde sólo disponemos de un corpus ad-hoc. En particular, comparamos nuestro sistema con dos de los sistemas más populares y ampliamente utilizados para inglés: SO-CAL (Taboada et al., 2011), un sistema dependiente del idioma disponible para inglés y español guiado por reglas léxicas a nivel morfológico; y SentiStrength (Thelwall et al., 2010), un sistema multilingüe que no aplica ningún etiquetado morfosintáctico, ya que se basa exclusivamente en un conjunto de lexicones de subjetividad, reglas superficiales y tratamiento de ciertos fenómenos no gramaticales (por ejemplo, replicación de caracteres). Adicionalmente, para la evaluación en español, también se tuvo en cuenta el sistema descrito en la sección 3.

La Tabla 2 compara el rendimiento de nuestro sistema con respecto a SentiStrength y SO-CAL en el corpus Taboada y Grieve (2004). Con respecto a SO-CAL, los resultados muestran que nuestro manejo de negación e intensificación proporciona mejores resultados (superando a SO-CAL en 3.25 puntos porcentuales en total). Con respecto a SentiStrength, nuestro sistema logra un mejor rendimiento en reseñas largas.

La Tabla 3 compara los tres sistemas no supervisados en el corpus Pang y Lee (2004), mostrando la robustez de nuestro sistema en diferentes dominios, ya que se desempeña de nuevo mejor que SO-CAL para la negación y la intensificación (aunque no se comporta tan bien cuando se trata de irrealis, probablemente debido a la necesidad de definir operaciones de composición más complejas para manejar este fenómeno), y también mejor que SentiStrength en reseñas largas de películas.

La Tabla 4 compara el desempeño de nuestro enfoque universal en español con respecto a: Spanish SentiStrength (Vilares, Thelwall, y Alonso, 2015), la versión para español de SO-CAL (Brooke, Tofiloski, y Taboada, 2009) y el sistema monolingüe para español

Rules	SentiStrength	SO-CAL	Nuestro sistema
Sistema base	N/A	65.50	65.00
+negación	N/A	67.75	71.75
+intensificación	66.00	69.25	74.25
+irrealis	N/A	71.00	73.75

Tabla 2: Exactitud (%) en el corpus Taboada y Grieve (2004). El hecho de que haya una sola fila para SentiStrength se debe que usamos la configuración estándar para inglés que ya incluye el manejo de la negación y la intensificación.

	SentiStrength	SO-CAL	Nuestro sistema
Sistema base	N/A	68.05	67.77
+negación	N/A	70.10	71.85
+intensificación	56.90	73.47	74.00
+irrealis	N/A	74.95	74.10

Tabla 3: Exactitud (%) en el corpus de test de Pang y Lee (2004).

definido en la sección 3. Utilizamos exactamente el mismo conjunto de operaciones de composición que se usó para el inglés (tan solo se reemplaza la lista de palabras de negación, intensificación y conjunciones adversativas por la correspondiente al español). Nuestro sistema multilingüe supera de nuevo a SentiStrength y SO-CAL en sus versiones en español.

Con el fin de comprobar la validez de nuestro enfoque para idiomas distintos del inglés y el español, hemos considerado el caso de alemán. Cabe señalar que los autores de este artículo no tenemos ninguna noción de alemán. A pesar de esto, hemos sido capaces de crear un sistema competitivo de análisis del sentimiento para este idioma mediante la integración de un léxico de sentimiento existente. Utilizamos el sistema alemán SentiStrength (Momtazi, 2012) para la comparación. El uso del diccionario alemán de SentiStrength nos permite mostrar cómo de robusto es nuestro sistema cuando se utilizan diferentes lexicones. Los resultados experimentales muestran una precisión del 72,75 % en el conjunto de datos de revisión de Amazon, mientras que SentiStrength informa un 69,95 %. Nuevamente, agregar la negación (72.05 %) y posteriormente la intensificación (72.85 %) produjeron mejoras relevantes sobre nuestro sistema base (69.85 %). Los resultados son comparables a los obtenidos para otras lenguas, utilizando un conjunto de datos de tamaño comparable, reforzando la robustez de nuestro enfoque en diferentes dominios, len-

guas y diccionarios.

6.2 Comparación con sistemas supervisados

En la evaluación no nos ceñimos únicamente a sistemas no supervisados, también realizamos una comparación con un enfoque supervisado de vanguardia, como es la red neuronal recursiva profunda presentada por Socher et al. (2013), entrenada en un treebank específico anotado con sentimiento, disponible solamente en inglés. Hasta donde sabemos, no existen métodos de este tipo para español y alemán. Los sistemas supervisados suelen ser imbatibles cuando se evalúan en la parte de test de los corpus con los que han sido entrenados. Sin embargo, en aplicaciones reales a menudo no se dispone de un corpus lo suficientemente grande que se corresponda con los textos objetivo en términos de género, estilo, longitud, etc. En este caso, el rendimiento de los sistemas supervisados desciende notablemente (Aue y Gamon, 2005).

La Tabla 5 compara nuestro sistema multilingüe sin supervisión con el sistema supervisado de Socher et al. (2013) en varios corpus: (1) la colección utilizada en la evaluación de Socher et al. (2013) (Pang y Lee, 2005); (2) un corpus del mismo dominio, es decir, películas (Pang y Lee, 2004); y (3) la colección Taboada y Grieve (2004).

Los resultados experimentales muestran que nuestro enfoque obtiene mejores resultados sobre los corpus (2) y (3). Vale la pena mencionar que nuestro enfoque de compo-

	Spanish SentiStrength	Spanish SO-CAL	Nuestro sistema multilingüe	Sistema monolingüe español
Sistema base	N/A	N/A	63.00	61.80
+negación	N/A	N/A	71.00	N/A
+intensificación	73.00	N/A	74.25	75.75
+irrealis	N/A	74.50	75.75	N/A

Tabla 4: Exactitud (%) en el corpus de test de Brooke, Tofiloski, y Taboada (2009).

Corpus	Socher et al. (2013)	Nuestro sistema
<i>Corpus original del modelo Socher et al. (2013):</i>		
Pang y Lee (2005)	85.40	75.07
<i>Otros corpus: Taboada y Grieve (2004)</i>	62.00	73.75
<i>Otros corpus: Pang y Lee (2004)</i>	63.80	74.10

Tabla 5: Exactitud (%) sobre diferentes corpus para (Socher et al., 2013) y nuestro sistema.

ción no supervisado superó al modelo supervisado no solo en un corpus fuera de su dominio, sino también en otro conjunto de datos del mismo dominio (películas) en el que la red neuronal fue entrenada y evaluada. Esto refuerza la utilidad de un enfoque no supervisado para las aplicaciones que necesitan analizar una serie de textos procedentes de diferentes dominios, estilos o fechas, cuando hay una carencia de datos etiquetados para entrenar clasificadores supervisados para todos ellos. Como era de esperar, Socher et al. (2013) es imbatible para un enfoque no supervisado en el conjunto de test del corpus donde fue entrenado. Sin embargo, nuestro algoritmo no supervisado también se comporta de manera muy robusta en este conjunto de datos.

7 Conclusiones

Hemos mostrado la efectividad de la información sintáctica a la hora de determinar el alcance de la negación en español en tareas de análisis del sentimiento y hemos visto que este enfoque es generalizable a un entorno multilingüe.

En un futuro nos proponemos aplicar el sistema multilingüe a lenguas no europeas, así como extenderlo a lenguas con pocos recursos lingüísticos.

Bibliografía

- Aue, A. y M. Gamon. 2005. Customizing sentiment classifiers to new domains: A case study. En *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP)*.
- Brooke, J., M. Tofiloski, y M. Taboada. 2009. Cross-linguistic sentiment analysis: From English to Spanish. En *Proceedings of RANLP 2009*, páginas 50–54, Bovorets, Bulgaria, Septiembre.
- Fernández Anta, A., P. Morere, L. Núñez Chiroque, y A. Santos. 2012. Techniques for sentiment analysis and topic detection of Spanish tweets: Preliminary report. En *TASS 2012 Working Notes*, Castellón de la Plana, Septiembre.
- Jia, L., C. Yu, y W. Meng. 2009. The effect of negation on sentiment analysis and retrieval effectiveness. En *CIKM'09 Proceeding of the 18th ACM conference on Information and knowledge management*, páginas 1827–1830, Hong Kong, Noviembre. ACM.
- Jiménez-Zafra, S. M., M. T. Martín-Valdivia, E. Martínez-Cámara, y L. A. Ureña López. En prensa. Studying the scope of negation for Spanish sentiment analysis on Twitter. *IEEE Transactions on Affective Computing*.
- Kennedy, A. y D. Inkpen. 2006. Sentiment classification of movie reviews using contextual valence shifters. *Computational Intelligence*, 22(2):110–125.
- McDonald, R., J. Nivre, Y. Goldberg, Y. Quirnbach-Brundage, D. Das, K. Ganchev, K. Hall, S. Petrov, H. Zhang, O. Täckström, C. Bedini, N. Bertomeu Castelló, y J. Lee. 2013. Universal dependency annotation for multilingual parsing. En *ACL 2013. 51st Annual Meeting of the Association for Computational*

- Linguistics. Proceedings of the Conference. Volume 2: Short Papers*, páginas 92–97, Sofia, Bulgaria, Agosto. ACL.
- Momtazi, S. 2012. Fine-grained german sentiment analysis on social media. En *LREC 2012*, páginas 1215–1220.
- Pang, B. y L. Lee. 2004. A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. En *Proceedings of the 42nd annual meeting on Association for Computational Linguistics*, páginas 271–278. ACL.
- Pang, B. y L. Lee. 2005. Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales. En *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, páginas 115–124. ACL.
- Sidorov, G., S. Miranda-Jiménez, F. Viveros-Jiménez, A. Gelbukh, N. Castro-Sánchez, F. Velásquez, I. Díaz-Rangel, S. Suárez-Guerra, A. Treviño, y J. Gordon. 2013. Empirical study of machine learning based approach for opinion mining in tweets. En I. Batyrshin y M. G. Mendoza, editores, *Advances in Artificial Intelligence*, volumen 7629 de *Lecture Notes in Artificial Intelligence*. Springer, Berlin and Heidelberg, páginas 1–14.
- Socher, R., A. Perelygin, J. Wu, J. Chuang, C. D. Manning, A. Ng, y C. Potts. 2013. Recursive deep models for semantic compositionality over a sentiment treebank. En *EMNLP 2013. 2013 Conference on Empirical Methods in Natural Language Processing. Proceedings of the Conference*, páginas 1631–1642, Seattle, Washington, USA, Octubre. ACL.
- Taboada, M. y J. Grieve. 2004. Analyzing appraisal automatically. En *Proceedings of AAAI Spring Symposium on Exploring Attitude and Affect in Text (AAAI Technical Report SS0407)*, Stanford University, CA, páginas 158–161. AAAI Press.
- Taboada, M., J. Brooke, M. Tofiloski, K. Voll, y M. Stede. 2011. Lexicon-based methods for sentiment analysis. *Computational Linguistics*, 37(2):267–307, Junio.
- Taulé, M., M. A. Martí, y M. Recasens. 2008. AnCorà: Multilevel Annotated Corpora for Catalan and Spanish. En N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis, y D. Tapias, editores, *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*, Marrakech, Morocco.
- Thelwall, M., K. Buckley, G. Paltoglou, D. Cai, y A. Kappas. 2010. Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 61(12):2544–2558, Diciembre.
- Vilares, D., M. A. Alonso, y C. Gómez-Rodríguez. 2015. A syntactic approach for opinion mining on Spanish reviews. *Natural Language Engineering*, 21(1):139–163, Enero.
- Vilares, D., M. A. Alonso, y C. Gómez-Rodríguez. 2016a. EN-ES-CS: An English-Spanish code-switching Twitter corpus for multilingual sentiment analysis. En N. C. C. Chair) K. Choukri, T. Declerck, M. Grobelnik, B. Maegaard, J. Mariani, A. Moreno, J. Odijk, y S. Piperidis, editores, *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, páginas 4149–4153, Portorož, Slovenia, Mayo.
- Vilares, D., M. A. Alonso, y C. Gómez-Rodríguez. 2016b. One model, two languages: training bilingual parsers with harmonized treebanks. En *ACL 2016. The 54th Annual Meeting of the Association for Computational Linguistics. Proceedings of the Conference, Vol. 2 (Short Papers)*, páginas 425–431, Berlin, Germany, Agosto. ACL.
- Vilares, D., C. Gómez-Rodríguez, y M. A. Alonso. 2017. Universal, unsupervised (rule-based), uncovered sentiment analysis. *Knowledge-Based Systems*, 118:45–55, Febrero.
- Vilares, D., M. Thelwall, y M. A. Alonso. 2015. The megaphone of the people? Spanish SentiStrength for real-time analysis of political tweets. *Journal of Information Science*, 41(6):799–813, Diciembre.
- Yang, K. 2008. WIDIT in TREC 2008 blog track: Leveraging multiple sources of opinion evidence. En E. M. Voorhees y L. P. Buckland, editores, *NIST Special Publication 500-277: The Seventeenth Text REtrieval Conference Proceedings (TREC 2008)*.