

Tarea 2 del Taller NEGES 2018: Detección de Claves de Negación

NEGES 2018 Task 2: Negation Cues Detection

Salud María Jiménez-Zafra¹, Noa P. Cruz Díaz²,
Roser Morante³, María Teresa Martín-Valdivia¹

¹SINAI, Centro de Estudios Avanzados en TIC (CEATIC), Universidad de Jaén
{sjzafra, maite}@ujaen.es

²Savana Médica, Madrid
contact@noacruz.com

³CLTL Lab, Computational Linguistics, VU University Amsterdam
r.morantevallejo@vu.nl

Resumen: El XXXIV Congreso Internacional de la Sociedad Española para el Procesamiento del Lenguaje Natural (SEPLN 2108) albergó el taller de NEGación en ESpañol (NEGES 2018). La segunda tarea compartida estuvo dedicada a identificar las claves de negación. Este artículo presenta las especificaciones, el conjunto de datos y los criterios de evaluación de la tarea. Además, se proporciona una descripción general de los sistemas participantes y se resumen sus resultados.

Palabras clave: Negación, detección de claves, análisis de sentimientos

Abstract: The XXXIV International Conference of the Spanish Society for the Natural Language Processing (SEPLN 2108) hosted the workshop on Negation in Spanish (NEGES 2018). The second shared task was dedicated to identifying the negation cues. This article presents the specifications, the data set and the evaluation criteria of the task. Moreover, an overview of participating systems is provided and their results are summarised.

Keywords: Negation processing, cue detection, sentiment analysis

1 *Introducción*

La negación es un fenómeno lingüístico complejo que se ha estudiado ampliamente desde una perspectiva teórica (Morante y Sporleder, 2012). Su detección y tratamiento automático es relevante en una amplia gama de aplicaciones, como la extracción de información (Savova et al., 2010), la traducción automática (Baker et al., 2012) o el análisis de sentimientos (Cruz Díaz et al., 2012), donde es crucial conocer cuándo una parte del texto puede tener un significado diferente debido a la presencia de negación. Esta parte del texto es lo que se conoce como el alcance o ámbito.

La detección de la negación se está convirtiendo en una tarea importante en el Procesamiento del Lenguaje Natural (PLN). De hecho, en los últimos años, varios desafíos y tareas compartidas han incluido la extracción de este elemento del lenguaje (Morante y Blanco, 2012). Sin embargo, la mayor parte de la investigación sobre la negación se ha realizado para inglés. En español, en-

contramos sistemas como los propuestos por Costumero et al. (2014) y Stricker, Iacobacci, y Cotik (2015) orientados a identificar de forma automática la negación en el ámbito clínico mediante la adaptación del popular algoritmo basado en reglas, NegEx (Chapman et al., 2001).

Aunque en menor medida, la detección de la negación en español se ha aplicado también al análisis de sentimientos como medio para determinar la polaridad de los sentimientos y la opinión. Por ejemplo, Vilares, Alonso, y Gómez-Rodríguez (2015) demostraron que tener en cuenta la estructura sintáctica del texto para el tratamiento de la negación, de la intensificación y de las oraciones subordinadas mejora con respecto a los sistemas puramente léxicos. Por otro lado, Jiménez-Zafra et al. (2017) desarrollaron un sistema basado en reglas de detección de las claves de negación y su alcance para mejorar la clasificación de tweets en español. La evaluación de este enfoque en el corpus TASS (Villena-Román

et al., 2013) mostró cómo introducir la detección de la negación en un sistema de análisis de sentimientos aumenta considerablemente su precisión.

En un trabajo más reciente, Jiménez-Zafra et al. (2018a) llevaron a cabo varios experimentos orientados a la clasificación de opiniones en los que se combinaban diferentes aspectos de la información de la negación y distintos algoritmos de aprendizaje automático. La colección de documentos utilizada en este caso fue el SFU ReviewSP-NEG corpus, anotado con claves de negación, su alcance y la forma en que la negación afecta a la polaridad de las palabras de su ámbito (inversión, no efecto, incremento o reducción) (Jiménez-Zafra et al., 2018b). Los resultados obtenidos demostraron que considerar la negación mejora el rendimiento de la tarea de clasificación de sentimientos.

A pesar de que, como demuestran los trabajos descritos anteriormente, el campo de la detección de la negación en español es un área muy activa en los últimos años, son necesarios más esfuerzos que permitan avanzar en este área. Por ello, la tarea 2 del taller NEGES 2018 se propuso con el objetivo de promover el desarrollo y la evaluación de los sistemas de identificación de claves de negación en español en el área del análisis de opiniones.

El resto del documento está organizado de la siguiente manera. La tarea propuesta se describe en la Sección 2 y el corpus en la Sección 3. Los sistemas participantes y sus resultados se resumen en la Sección 4. Los enfoques utilizados por los sistemas participantes se describen en la Sección 5, así como el análisis de los resultados. Finalmente, la Sección 6 concluye el documento.

2 Descripción de la tarea

Como se ha especificado en la introducción, la tarea 2 del taller NEGES 2018¹ estuvo dedicada a la detección automática de las claves de negación en español en el dominio de las opiniones.

Por ejemplo, en la frase (1) los sistemas tenían que identificar la palabra *no* como clave de negación.

(1) Yo [no] recomiendo el libro.

¹<http://www.sepln.org/workshops/neges/index.php>

Los sistemas participantes recibieron un conjunto de datos de entrenamiento y desarrollo para construir sus sistemas durante la fase de entrenamiento. Todas las claves de negación utilizadas para las fases de entrenamiento, desarrollo y test se generaron a través de una exhaustiva anotación manual por expertos del dominio, siguiendo unas guías de anotación bien definidas (Jiménez-Zafra et al., 2018b). En una etapa posterior, se puso disponible un conjunto de test para el que los participantes enviaron sus predicciones que fueron evaluadas frente a las anotaciones manuales.

2.1 Métricas de evaluación

El script de evaluación fue el propuesto en el *SEM 2012 Shared Task - Resolving the Scope and Focus of Negation (Morante y Blanco, 2012) y se basa en los siguientes criterios:

- Los signos de puntuación son ignorados.
- Un verdadero positivo (VP) requiere que todos los elementos de la clave de negación sean identificados correctamente.
- Para evaluar las claves, las coincidencias parciales no se cuentan como falso positivo (FP), sólo como falso negativo (FN) con el objetivo de evitar penalizar más la identificación parcial de las claves de negación que la no identificación de las mismas.

Se utilizaron las medidas estándar de evaluación de precisión (P), cobertura (C) y Medida-F donde:

$$P = \frac{VP}{VP + FP} \quad (1)$$

$$C = \frac{VP}{VP + FN} \quad (2)$$

$$Medida - F = \frac{2 * P * R}{P + R} \quad (3)$$

En la evaluación propuesta, se cuenta como FN aquellas claves incorrectamente identificadas por el sistema, ya sea porque la coincidencia es parcial o porque directamente no hayan sido identificadas como claves de negación. Se considera FP aquellas palabras identificadas incorrectamente como claves de negación. Finalmente, se cuenta como VP aquellas claves de negación correctamente identificadas por el sistema (esto es, coincidencia exacta).

	Entrenamiento	Desarrollo	Test
Comentarios	264 (130-, 134+)	56 (34-, 22+)	80(36-, 44+)
Neg	2.511	594	836
Noneg	104	22	55
Contrast	100	23	52
Comp	18	6	6

Tabla 1: Distribución de comentarios y claves de negación en la colección de datos

3 Colecciones de datos

El corpus SFU ReviewSP-NEG (Jiménez-Zafra et al., 2018b) fue la colección de documentos utilizada para entrenar y probar los sistemas.

Los conjuntos de entrenamiento, desarrollo y test fueron generados de forma aleatoria manteniendo 33 comentarios por dominio en el entrenamiento, 7 por dominio en el desarrollo y 10 por dominio en el test. Las claves de negación se corresponden con las estructuras *neg* y las que no niegan con las demás estructuras (*noneg*, *contrast*, *comp*). En total, 221.866 palabras y 9.455 frases, de las cuales, 3.022 frases contienen al menos una estructura de negación. En la Tabla 1 puede verse la distribución de comentarios y claves.

Los datos se proporcionaron en formato CoNLL (Farkas et al., 2010) donde cada línea corresponde a un token, cada anotación se proporciona en una columna y las líneas vacías indican el final de la frase. El contenido de las columnas dadas es:

- Columna 1: dominio y nombre del archivo separados por barra baja.
- Columna 2: número de oración dentro del archivo.
- Columna 3: número de token dentro de la oración.
- Columna 4: palabra.
- Columna 5: lema.
- Columna 6: categoría gramatical.
- Columna 7: tipo de categoría gramatical.

Los sistemas tuvieron que generar un archivo por dominio con el contenido de las 7 columnas anteriores más las siguientes columnas (los archivos debían tener el mismo orden de oraciones y tokens que los archivos de test):

- Columnas 8 hasta el final: Si la oración no tiene negaciones, la columna 8 tiene un valor ***** y no hay más columnas. Si la oración tiene negaciones, la anotación para cada negación se proporciona en tres columnas. La primera columna contiene la palabra que pertenece a la clave de negación y la segunda y tercera columnas contienen -.

3.1 Anotación de claves de negación en el SFU ReviewSP-NEG corpus

El SFU ReviewSP-NEG corpus² consiste en 400 reseñas de coches, hoteles, lavadoras, libros, teléfonos móviles, música, ordenadores y películas extraídas del sitio web *ciao.es*. Cada dominio contiene 25 revisiones positivas y 25 negativas. La anotación se llevó a cabo a nivel de token para el lema y la categoría gramatical y a nivel de frase para las claves de negación, su alcance lingüístico y el evento.

Las etiquetas utilizadas para la anotación de la negación son las siguientes:

- `<review polarity="positive/negative">`. Describe la polaridad de la revisión, que puede ser positiva o negativa.
- `<sentence complex="yes/no">`. Esta etiqueta corresponde a una frase completa o fragmento de la misma en la que pueda aparecer una estructura negativa. Las frases sin claves negativas no son anotadas. Esta etiqueta lleva asociado un atributo que puede tomar los siguientes valores:
 - “yes”, si la frase contiene más de una estructura negativa (`<neg.structure>`).
 - “no”, si la frase contiene sólo una estructura negativa.

²<http://sinai.ujaen.es/sfu-review-sp-neg-2/>

- `<neg structure>`. Esta etiqueta corresponde con una estructura sintáctica en la que aparece una clave de negación. Tiene 4 atributos posibles, dos de los cuáles (*change* y *polarity_modifier* son mutuamente excluyentes).
 - *polarity*: indica la orientación semántica de la estructura de negación (*“positive”*, *“negative”* or *“neutral”*).
 - *change*: indica si la orientación o significado de la estructura de negación se ha modificado completamente a causa de la negación (*change=“yes”* o *change=“no”*).
 - *polarity modifier*: indica si la estructura negativa contiene un elemento que matiza su polaridad. Si hay un incremento en la intensidad del valor de polaridad, toma el valor *“increment”* y, en cambio, si hay una disminución del valor de polaridad, toma el valor *“reduction”*.
 - *value*: muestra el significado de la estructura de negación, esto es, *“neg”* si expresa negación, *“contrast”* si expresa contraste u oposición, *“comp”* si expresa comparación o desigualdad entre términos y *“noneg”* si no niega pese a contener una clave de negación.
- `<scope>`. Esta etiqueta delimita la parte de la estructura negativa que está dentro del alcance de la negación. Incluye tanto la clave de negación (`<negexp>`) como el evento (`<event>`).
- `<negexp>`. Esta etiqueta se corresponde con la(s) palabra(s) que indica(n) negación. Puede tener asociado el atributo *discid* si la negación está expresada por más de un elemento negativo discontinuo.
- `<event>`. Denota las palabras que están afectadas directamente por la negación.

4 Trabajos presentados y resultados

Dos sistemas participaron en esta tarea (Fabregat, Martínez-Romo, y Araujo, 2018; Loharja, Padró, y Turmo, 2018). No obstante, hubo participantes que mostraron interés

en la tarea y expresaron que no participaron debido a la falta de tiempo. Los resultados oficiales se muestran en las Tablas 2, 3 y 4, evaluados en términos de Precisión (Prec), Cobertura (Cob) y Medida F (F).

Equipo	Prec	Cob	F
UNED	79,45	59,58	67,97
UPC	91,48	82,18	86,45

Tabla 2: Resultados oficiales globales

Si comparamos los resultados globales de la Tabla 2, vemos cómo la precisión de ambos sistemas fue similar. Sin embargo, el nivel de cobertura que presentó el sistema desarrollado por el equipo de la UNED fue bajo, lo que provocó la caída de la medida-F. No obstante, hay que tener en cuenta que hubo un error al enviar los datos y el sistema no pudo ser evaluado sobre la subcolección de ordenadores, por lo que los resultados globales deben ser interpretados con precaución. El mejor resultado lo obtuvo el equipo de la UPC con una medida-F del 86,25 %.

Dominio	Prec	Cob	F
Coches	94,23	72,06	81,67
Hoteles	97,67	71,19	82,35
Lavadoras	92,00	66,67	77,31
Libros	79,52	66,27	72,29
Teléfonos móviles	93,33	73,68	82,35
Música	92,59	57,47	70,92
Ordenadores	-	-	-
Películas	86,26	69,33	76,87

Tabla 3: Resultados oficiales por dominio para el equipo de la UNED

Dominio	Prec	Cob	F
Coches	95,08	85,29	89,92
Hoteles	94,00	79,66	86,24
Lavadoras	94,74	78,26	85,72
Libros	84,19	84,52	84,35
Teléfonos móviles	89,80	77,19	83,02
Música	92,96	75,86	83,54
Ordenadores	91,36	91,36	91,36
Películas	89,68	85,28	87,42

Tabla 4: Resultados oficiales por dominio para el equipo de la UPC

Si analizamos los datos por dominio en las Tablas 3 y 4, hay subcolecciones como la de libros y móviles en la que ambos sistemas obtuvieron peores resultados en comparación con el resto de subcolecciones. El sistema desarrollado por la UNED obtuvo el mayor rendimiento en las subcolecciones de móviles y hoteles, mientras que el sistema de la UPC mostró una mejor detección de las claves de negación en la subcolección de ordenadores, en concreto, obtuvo una medida-F de 91,36 %.

5 Enfoques

Remitimos al lector a los artículos de los participantes (Fabregat, Martínez-Romo, y Araujo, 2018; Loharja, Padró, y Turmo, 2018) para obtener una descripción completa de los sistemas. En la sección anterior puede verse el rendimiento, tanto global como por dominio, de cada uno de estos sistemas.

Ambas propuestas utilizaron el esquema estándar de etiquetado *BIO* donde la primera palabra de una estructura de negación denota por *B* y las restantes por *I*. La etiqueta *O* indica que la palabra no se corresponde con una clave de negación.

El equipo de la UNED (Fabregat, Martínez-Romo, y Araujo, 2018) aplicó un modelo de aprendizaje profundo inspirado en las arquitecturas del reconocimiento de entidades nombradas y en modelos de detección del ámbito de la negación. En concreto, la propuesta está basada en el uso de varias redes neuronales junto a una LSTM (Long Short-Term Memory) bidireccional. Este enfoque supervisado se basa en word embeddings preentrenados para el español. Por su parte, el equipo de la UPC (Loharja, Padró, y Turmo, 2018) utilizó el algoritmo de aprendizaje automático Conditional Random Fields (CRFs) junto a una serie de atributos como la categoría gramatical de la palabra, información sobre cómo están escritas las palabras, etc.

Los recursos utilizados por los participantes son diversos. El equipo de la UNED hizo uso de la librerías Keras (Chollet, 2015) y TensorFlow (Abadi et al., 2016), además de word embeddings preentrenados para el español (Cardellino, 2016), y el equipo de la UPC utilizó NLTK (Loper y Bird, 2002).

6 Conclusiones

La tarea 2 del taller NEGES 2018 promovió el desarrollo de sistemas de detección automática de la negación en español en el ámbito del análisis de sentimientos, una tarea clave de PLN. En concreto, se centró en la detección de las claves de negación presentes en la colección de documentos SFU ReviewSP-NEG. Hasta donde conocemos, se trató de la primera tarea orientada a detectar la negación en español, de ahí la relevancia del taller. 2 fueron los sistemas presentados pero varios equipos mostraron su interés en la tarea, no pudiendo entregar sus propuestas a tiempo. Los participantes utilizaron enfoques supervisados basados en técnicas de aprendizaje automático y aprendizaje profundo, mostrando las primeras un mayor rendimiento.

En futuras ediciones del taller, las tareas podrían ir orientadas a detectar de forma automática las claves de negación en español en otros dominios como el clínico. Además, se podría incluir la detección del ámbito o alcance de dichas claves de negación.

Agradecimientos

Este trabajo ha sido parcialmente financiado por el Ministerio de Educación, Cultura y Deporte (MECD - ayuda FPU014/00983), el Fondo Europeo de Desarrollo Regional (FEDER) y el proyecto REDES (TIN2015-65136-C2-1-R) del Gobierno de España. Roser Morante ha sido financiada por la Organización Holandesa para la Investigación Científica (NWO) a través del premio Spinoza otorgado a Piek Vossen (SPI 30-673, 2014-2019).

Bibliografía

- Abadi, M., P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, y others. 2016. Tensorflow: a system for large-scale machine learning. En *OSDI*, volumen 16, páginas 265–283.
- Baker, K., M. Bloodgood, B. J. Dorr, C. Callison-Burch, N. W. Filardo, C. Piatko, L. Levin, y S. Miller. 2012. Modality and negation in simt use of modality and negation in semantically-informed syntactic mt. *Computational Linguistics*, 38(2):411–438.
- Cardellino, C. 2016. Spanish billion words corpus and embeddings.

- Chapman, W. W., W. Bridewell, P. Hanbury, G. F. Cooper, y B. G. Buchanan. 2001. A simple algorithm for identifying negated findings and diseases in discharge summaries. *Journal of biomedical informatics*, 34(5):301–310.
- Chollet, F. 2015. Keras.
- Costumero, R., F. López, C. Gonzalo-Martín, M. Millan, y E. Menasalvas. 2014. An approach to detect negation on medical documents in Spanish. En *International Conference on Brain Informatics and Health*, páginas 366–375. Springer.
- Cruz Díaz, N. P., M. J. Maña López, J. Mata Vázquez, y V. Pachón Álvarez. 2012. A machine-learning approach to negation and speculation detection in clinical texts. *Journal of the American society for information science and technology*, 63(7):1398–1410.
- Fabregat, H., J. Martínez-Romo, y L. Araujo. 2018. Deep Learning Approach for Negation Cues Detection in Spanish at NEGES 2018. En *Proceedings of NEGES 2018: Workshop on Negation in Spanish*, volumen 2174, páginas 43–48.
- Farkas, R., V. Vincze, G. Móra, J. Csirik, y G. Szarvas. 2010. The CoNLL-2010 shared task: learning to detect hedges and their scope in natural language text. En *Proceedings of the Fourteenth Conference on Computational Natural Language Learning—Shared Task*, páginas 1–12. Association for Computational Linguistics.
- Jiménez-Zafra, S. M., M. T. Martín-Valdivia, M. D. Molina-González, y L. A. Ureña-López. 2018a. Relevance of the SFU Review SP-NEG corpus annotated with the scope of negation for supervised polarity classification in Spanish. *Information Processing & Management*, 54(2):240–251.
- Jiménez-Zafra, S. M., M. Taulé, M. T. Martín-Valdivia, L. A. Ureña-López, y M. A. Martí. 2018b. SFU Review SP-NEG: a Spanish corpus annotated with negation for sentiment analysis. a typology of negation patterns. *Language Resources and Evaluation*, 52(2):533–569.
- Jimenez-Zafra, S. M., M. T. M. Valdivia, E. M. Camara, y L. A. Urena-Lopez. 2017. Studying the Scope of Negation for Spanish Sentiment Analysis on Twitter. *IEEE Transactions on Affective Computing*.
- Loharja, H., L. Padró, y J. Turmo. 2018. Negation Cues Detection Using CRF on Spanish Product Review Text at NEGES 2018. En *Proceedings of NEGES 2018: Workshop on Negation in Spanish*, volumen 2174, páginas 49–54.
- Loper, E. y S. Bird. 2002. NLTK: The natural language toolkit. En *Proceedings of the ACL-02 Workshop on Effective tools and methodologies for teaching natural language processing and computational linguistics-Volume 1*, páginas 63–70. Association for Computational Linguistics.
- Morante, R. y E. Blanco. 2012. * SEM 2012 shared task: Resolving the scope and focus of negation. En *Proceedings of the First Joint Conference on Lexical and Computational Semantics-Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation*, páginas 265–274. Association for Computational Linguistics.
- Morante, R. y C. Sporleder. 2012. Modality and negation: An introduction to the special issue. *Computational linguistics*, 38(2):223–260.
- Savova, G. K., J. J. Masanz, P. V. Ogren, J. Zheng, S. Sohn, K. C. Kipper-Schuler, y C. G. Chute. 2010. Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): architecture, component evaluation and applications. *Journal of the American Medical Informatics Association*, 17(5):507–513.
- Stricker, V., I. Iacobacci, y V. Cotik. 2015. Negated Findings Detection in Radiology Reports in Spanish: an Adaptation of NegEx to Spanish. En *IJCAI-Workshop on Replicability and Reproducibility in Natural Language Processing: adaptative methods, resources and software, Buenos Aires, Argentina*.
- 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.

Villena-Román, J., S. Lana-Serrano, E. Martínez-Cámara, y J. C. González-Cristóbal. 2013. TASS-Workshop on Sentiment Analysis at SEPLN. *Procesamiento del Lenguaje Natural*, 50:37-44.

