

Detección de la negación y su ámbito en textos escritos en español

Loraine Monteagudo García

Grupo C211

L.MONTEAGUDO@ESTUDIANTES.MATCOM.UH.CU

Tutor(es):

Msc. Suilan Estévez Velarde, *Universidad de La Habana*

Msc. Alejandro Píad Morffis, *Universidad de La Habana*

Dr. Yudivián Almeida Cruz, *Universidad de La Habana*

Resumen

La detección automática de la negación y de las palabras que son afectadas por esta es una importante tarea que podría beneficiar otras tareas del Procesamiento de Lenguaje Natural tales como la Extracción de Información, Análisis de Sentimientos y Respuestas a Preguntas. En el presente trabajo se propone una solución al problema de la negación basado en técnicas de aprendizaje supervisado. La esencia de la propuesta es detectar tanto las partículas de la negación como su ámbito, todo ello en textos en español. Se trabaja en dos fases: en la primera las partículas son identificadas y en la segunda el alcance completo de estas es determinado.

Abstract

The automatic detection of negation and the words they affect is an important task that could benefit other Natural Language Processing tasks such as Information Extraction, Sentiment Analysis and Question Answering. In the present work a solution of the negation problem is proposed based in supervised machine learning techniques. The essence is the detection of the negation signals and their scope, all of it in spanish texts. The work is done in two phases: in the first one the signals are identified and in the second the whole scope of this signals is determined.

Palabras Clave: Detección de la negación, identificación del ámbito de la negación, análisis de sentimientos

Tema: Inteligencia Artificial, Procesamiento de Lenguaje Natural.

1. Introducción

En el Procesamiento de Lenguaje Natural (NLP) se realizan tareas importantes como Análisis de Sentimientos, Extracción de Información y Respuesta a preguntas en los que es de gran importancia el tratamiento de textos que incluyen información cargada de negación o subjetividad. En el análisis de sentimientos la negación constituye un elemento clave, dado que son muy abundantes las opiniones negativas expresadas en términos positivos negados y viceversa. La oración “No me gusta la carcasa del teléfono” es un claro ejemplo de una opinión negativa con un término positivo (gusta) negado.

La negación es un fenómeno lingüístico que es usado para cambiar el valor de verdad de una unidad lingüística (oración, frase o palabra) [18]. Distintos idiomas expresan la negación mediante distintas partículas o marcadores, siendo los más comunes la utilización de palabras como “no”, “nunca”, “nadie”, pero también puede ser expresado mediante prefijos “(in)dispensable”, “(im)pensable”, “(des)interesado” y frases hechas como “en la vida”, entre otros. El ámbito o alcance de la negación incluye todas esas palabras que son afectadas por la negación y típicamente el marca-

dor de esta será incluida en él. En el siguiente ejemplo el marcador se denota en negrita y el ámbito en paréntesis:

Una cámara de fotos que (**no** es una maravilla).

A primera vista, se podría pensar que la negación puede ser reducida a encontrar palabras negativas, detectar su ámbito usando un análisis sintáctico y revertir su polaridad. Pero, en realidad, es más complicado que esto. La negación juega un papel importante para entender el lenguaje y trae consigo varios desafíos.

La identificación de los marcadores de la negación puede resultar complicada: algunas palabras que acostumbra a expresar negación no siempre lo hacen de acuerdo al contexto, como, por ejemplo, en la oración “El coche lo compré para viajar, **no**?” *no* tiene simplemente un valor enfático. Además, existen negaciones que pueden consistir en más de una palabra, algunas de las cuales son discontinuas, como “**No** me gustó **en absoluto**”.

En el presente trabajo se propone una solución al problema de la negación basado en técnicas de aprendizaje supervisado. La esencia de la propuesta es detectar tanto las partículas de la negación como su ámbito, todo ello en textos en español. Se trabaja en dos fases: en la primera las partículas son identificadas y en

la segunda el alcance completo de estas partículas es determinado.

Aunque la negación ha sido largamente estudiada, no ha sido lo suficientemente analizada en varios idiomas como el español. Según nuestro conocimiento no ha sido llevado a cabo ningún sistema que realice un enfoque de aprendizaje supervisado en este idioma. Por lo tanto, el propósito de este trabajo será la realización de un modelo capaz de tratar la negación, a través de una solución simple que pretenda analizar las características más importantes para identificar las palabras negadas en una oración.

El resto del artículo se organiza como sigue: en la sección 2 se presenta los trabajos relacionados con el tema de detección del ámbito de la negación. La sección 3 describe el algoritmo propuesto para detectar las partículas de la negación y su alcance. Posteriormente, la sección 4 detalla el marco experimental realizado para comprobar la efectividad de la propuesta, los recursos utilizados y los resultados obtenidos. Por último, en la sección 5 se exponen las conclusiones y la línea de trabajo a seguir.

2. Trabajos Relacionados

La negación ha sido un área bastante estudiada en NLP por su influencia en otras tareas de gran interés, pero la mayor parte de los trabajos realizados han sido sobre textos escritos en inglés.

Los primeros sistemas desarrollados fueron típicamente basados en reglas. Por ejemplo, el algoritmo NegEx [6], que determina si una enfermedad o síntoma mencionado en un reporte médico está presente o ausente usando reglas basadas en expresiones regulares. En otros trabajos se plantean métodos basados en la definición de reglas a partir de árboles sintácticos, como los de Apostolova [3], De Albornoz [2] y Ballesteros [4].

Sin embargo, la mayoría de los trabajos en este campo están basados en enfoques de aprendizaje supervisado. Un notable ejemplo es aquel desarrollado por Morante [19]. Su sistema consiste en 4 clasificadores: tres de los cuales predicen si un token es el primero, último o si no está en un ámbito y un cuarto usa estas predicciones para determinar a qué clase pertenece cada token del ámbito. Otro trabajo fue aquel desarrollado por Agarwal [1], que detecta las frases que expresan negación/especulación y su alcance usando Conditional Random Fields (CRF) como algoritmo de aprendizaje. Un método similar fue el usado por Chowdhury [7], destacándose por su utilización de características contextuales aparte de las específicas del token.

Se caracterizan por su utilización de información sintáctica sistemas como el desarrollado por Council [8], que es entrenado y evaluado en un corpus de críticas de productos concluyendo que, como es esperado, el rendimiento mejora significativamente la clasificación de la polaridad introduciendo la detección del ámbito. Otros trabajos notables fueron aquellos desarrollados por Lapponi [16], [17] y Enger [11] que usan un amplio

conjunto de características léxicas y sintácticas así como etiquetas que capturan el comportamiento a nivel de tokens del ámbito usando Support Vector Machines (SVM) y CRF como algoritmos de aprendizaje para la detección de las partículas de negación y su ámbito respectivamente. Cruz Díaz [10] presenta un sistema para el tratamiento de la negación y de la especulación en textos médicos y opiniones, y el primer corpus de opiniones etiquetado a nivel de negación y especulación. El modelo introducido por Reitan [21] estudia el efecto del tratamiento de la negación para la clasificación de sentimientos desarrollando un sistema sofisticado para determinar el ámbito de la negación.

Otros enfoques utilizados para la detección del ámbito de la negación son aquellos realizados por Basile [5] y Packard [20] que se basan en las representaciones semánticas del texto a través de *deep parsing*. Además se destaca el desarrollado por Fancellu [12] donde se demuestra que las redes neuronales son una válida alternativa para la tarea de detectar el ámbito de la negación, ofreciendo un análisis detallado de su rendimiento en textos de diferentes géneros y conteniendo distintos tipos de negación.

Con respecto al tratamiento de la negación en documentos en español se destacan sistemas basados en reglas como el desarrollado por Costumero [9]. Este consiste en una adaptación del algoritmo NegEx para aplicarlo en la detección de la negación en documentos médicos escritos en español. Ellos traducen la lista de términos anteriormente identificados en Inglés por NegEx y lo enriquecen con sinónimos y términos extraídos en textos médicos en Español.

Jimenez-Zafra [13] propone un modelo que se basa en la estructura sintáctica del texto para la resolución del ámbito de la negación en español. Utiliza distintas reglas basadas en el árbol de dependencias para el tratamiento de determinadas partículas de la negación. Además, se hace uso de la identificación del ámbito de la negación en un sistema de clasificación de la polaridad, evaluando sus contribuciones en esta tarea.

Como se ve, la mayor parte de las investigaciones realizadas hasta el momento sobre el tratamiento de la negación se centran en textos escritos en inglés. A pesar de las similitudes de los conceptos relacionados con la negación los mecanismos para expresarlas varían dependiendo del lenguaje, por que cada uno de estos requiere una manera específica de tratarla. Los pocos trabajos llevados a cabo para su análisis en español se realizan mediante sistemas basados en reglas. Por lo tanto, uno de los propósitos de este proyecto será la realización de un enfoque basado en aprendizaje supervisado para la detección del ámbito de la negación en español.

3. Propuesta

La tarea de identificar las palabras afectadas por la negación se divide en dos sub tareas de clasificación consecutivas, en ambos se aplica un enfoque de aprendizaje automático supervisado.

En la primera fase se determina si una determinada oración está negada o no, identificando las palabras que determinan la negación, estas predicciones son usadas en la segunda fase, en donde se determina el alcance de la negación.

Primero, el texto es procesado y se obtienen de él información léxica (tokens, part-of-speech tags y lemmas) y sintáctica (análisis de dependencias). Estos datos son convertidos para la entrada del detector de la negación. Luego de obtener su predicción estas son usadas para la determinación del alcance de la negación. Las siguientes subsecciones explican el proceso más detalladamente.

3.1 Detección de la Negación

En esta primera fase se determinan las palabras que expresan negación, que no son más que los elementos lexicales que indican la negación en una oración.

La detección de la negación es resuelta usando Logistic Regression, presentándose todas las palabras al clasificador y determinando este cuáles palabras expresan negación.

Características	Descripción
Token_i	Token actual
Lemma_i	Lemma del token actual
POS_i	Part-of-speech del token actual
Lemma_{i-1}	Lemma del token anterior
POS_{i-1}	Part-of-speech del token anterior
Lemma_{i+1}	Lemma del token siguiente
POS_{i+1}	Part-of-speech del token siguiente
Lemma_{i-2}	Lemma del token $_{i-2}$
Lemma_{i+2}	Lemma del token $_{i+2}$

Cuadro 1: Características en la fase de detección de la partícula

Se introducen características basadas en el contexto con el objetivo de tratar marcadores complejos (compuestos por más de una palabra), principalmente complejos continuos (las palabras por las que está compuesto están seguidas) como “ni siquiera”. Sin embargo, las características más informativas han probado ser aquellas basadas solo en el token.

3.2 Identificación del Ámbito de la Negación

Una vez las partículas de la negación son identificadas, la siguiente tarea es detectar el alcance de estas partículas. Para ello un clasificador, Logistic Regression, predice si un token pertenece al ámbito de un marcador o no.

Las instancias que son presentadas al modelo representan un par: una señal de negación y un token de

una oración. Esto significa que todas las palabras de la oración se analizan tantas veces como negaciones haya en ella, analizando las características de todas las palabras con respecto a esta partícula, tomándose la decisión de si pertenece a su ámbito o no. Las partículas de la negación usadas son aquellas que fueron detectadas en la fase anterior. Solo las oraciones que expresan negación son usadas en esta fase.

Características	Descripción
Léxicas	
Token_i	Token actual
Lemma_i	Lemma del token actual
POS_i	Part-of-speech del token actual
Lemma_{i-1}	Lemma del token anterior
POS_{i-1}	Part-of-speech del token anterior
Lemma_{i+1}	Lemma del token siguiente
POS_{i+1}	Part-of-speech del token siguiente
Sobre la partícula	
Distancia derecha	Distancia hacia la derecha de la partícula de negación
Distancia izquierda	Distancia hacia la izquierda de la partícula de negación
POS del marcador	Part-of-speech del marcador
Sintácticas	
Relación de dependencia	Tipo de relación de dependencia del token actual
Camino del grafo dirigido de dependencias	Camino de longitud mínima del antecesor del marcador al token
Longitud del camino del grafo dirigido	Cantidad de aristas atravesadas en el camino de longitud mínima
Camino del grafo de dependencias	Camino del token actual al marcador
Longitud del camino	Cantidad de aristas atravesadas en el camino anterior

Cuadro 2: Características en la fase de resolución del ámbito

Estas características pueden ser divididas en léxicas, sintácticas y en aquellas relacionadas con las partículas. Las características empleadas están basadas en aquellas usadas en trabajos previos de la resolución de la negación.

Basados en observaciones hechas por Enger [11] y con el propósito de evitar características que pueden tomar muchos valores diferentes todas aquellas que adquieren distintos valores (las distancias con respecto a la partícula y la de los caminos de los grafos) se hicieron discretas. Para las longitudes de los grafos se eligieron dos límites: 3 y 7. Se plantea que cada token que tenga una distancia menor que 3 es un vecino in-

mediato del vértice inicial, si esta es menor que 7 es cercano y cualquier otro token que tenga una distancia mayor es lejano. La distancia con respecto al marcador es la cantidad de tokens que hay entre la palabra y la partícula que se están analizando y para hacerla discreta se puso un límite de 7 a la izquierda y de 12 a la derecha.

Las características extraídas mediante la representación del grafo de dependencias intenta modelar la relación sintáctica entre el token y la partícula que se está analizando y ha probado ser muy beneficioso en la tarea de determinar el ámbito de la negación. Cada oración a la que se realizó el análisis de dependencias es llevada a su representación a grafo.

Se usan los índices de los tokens como identificadores inequívocos de los nodos del grafo, y las relaciones de dependencias como las aristas. Representamos el grafo de dependencias como un conjunto V y dos conjuntos de aristas: A y E el primero conteniendo solo aristas dirigidas (arcos) y el segundo no dirigidas.

Del grafo $G = (V, A)$ se extrae el camino más corto del antecesor de la partícula de la negación hacia todos los vértices del grafo. Empezamos desde el antecesor porque al ser los grafos dirigidos acíclicos esencialmente árboles, las partículas de la negación son encontradas frecuentemente en las hojas, por lo que no habría camino de la negación hacia ningún otro vértice. Del grafo $G' = (V, E)$ extraemos el camino más corto de los vértices hacia la partícula de la negación que se está analizando.

Luego, se usan estos caminos de distancia mínima para guardar las longitudes y el camino de estos grafos como una característica. Este último representa el camino de cada token hacia la partícula de la negación, representando las relaciones de dependencias en el camino. Por ejemplo, el camino del grafo no dirigido entre *fotos* y *no* en la Figura 1 es descrita como “*nmod/acl/advmod/*”.

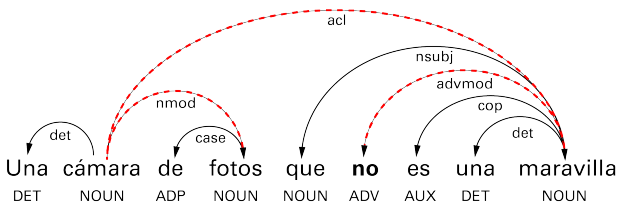


Figura 1: Ejemplo de un grafo de dependencias representando el camino entre *fotos* y *no*

4. Marco Experimental

Para la realización de las tareas de tokenización, el análisis léxico y el análisis de dependencias se ha utilizado spaCy. Se trata de una biblioteca que proporciona una amplia gama de herramientas para NLP en varios idiomas, entre ellos el español. Para la representación del grafo y su tratamiento se ha utilizado NetworkX.

Para probar la funcionalidad de la propuesta desa-

rollado se ha utilizado un corpus en español SFU Review_{SP-NEG} [14], que consiste en 400 comentarios de carros, hoteles, lavadoras, libros, móviles, ordenadores, música, libros y películas, extraído de la página Ciao.es. Las críticas que componen el corpus están escritas por usuarios de la web, lo que aumenta la dificultad de la tarea ya que los textos pueden contener errores gramaticales e incluso expresiones informales. De las 9,455 oraciones que contiene el corpus 3,022 están anotadas con negación, de las cuales 2,143 solo contienen un marcador de la negación, y 879 presentan más de uno. Solo el 6.31 % de los marcadores son complejos continuos (están compuestas por más de dos palabras consecutivas), siendo la mayoría simples y discontinuos complejos (compuestas por varias palabras discontinuas), representando el 70.78 % y 22.91 % respectivamente. Además es interesante mencionar que hay más de 400 distintos tipos de partículas de negación, pero las 10 más frecuentes constituyen más del 70 % de los casos de negación, siendo el marcador ‘no’ el más frecuente, representando el 52.14 % de los casos.

4.1 Experimentos

Para la realización de los experimentos el corpus fue dividido de la siguiente forma: el 80 % lo constituye el conjunto de entrenamiento, 10 % de desarrollo y el otro 10 % para evaluación.

Se probaron distintas configuraciones con respecto a las características de tanto el clasificador de los marcadores como del ámbito de la negación en el conjunto de desarrollo, sin embargo solo los resultados obtenidos del conjunto final son mostrados.

Para determinar los óptimos valores para los parámetros se usó grid search y una validación cruzada de 10 iteraciones en el conjunto de entrenamiento con los coeficientes de penalización L1 y L2 de los Logistic Regression, C con un espacio de $[0.001, 0.01, 0.1, 1, 5, 10]$ y max_iter con un espacio de $[50, 100, 150, 200]$. Las configuraciones óptimas fueron: para el clasificador de la detección de la negación $\text{penalty} = \text{L1}$, $C = 1$, $\text{max_iter} = 150$ y para el de la resolución del ámbito de la negación: $\text{penalty} = \text{L1}$, $C = 1$, $\text{max_iter} = 100$

4.2 Resultados

Los resultados obtenidos fueron medidos usando validación cruzada de 10 iteraciones. Por cada iteración, una partición del documento fue usada.

Como medidas de evaluación del sistema se ha usado: Precisión, Recall, F1 y Accuracy; siendo las tres primeras las más usadas para evaluar el rendimiento tanto de la detección de la partícula como del ámbito de la negación. Además, la medida F1 es bien establecida para conjuntos de datos que no están balanceados.

En la fase de detección de la negación, un token es correctamente clasificado si ha sido identificado como marcador de la negación correctamente. En la tarea de la resolución del alcance un token es clasificado debi-

damente si es determinado apropiadamente como parte del ámbito de alguna negación. Es decir, el sistema evalúa por cada marcador si un token pertenece a su ámbito o no. Este resultado es procesado de manera tal que si un token se clasifica como parte de algún ámbito entonces se identifica como palabra negada, existiendo así tokens que pueden ser clasificados más de una vez y ámbitos que se mezclan. Estos son interpretados como un solo ámbito que posee más de una partícula de la negación, uno de dichos marcadores haciendo función de refuerzo o modificador de la negación.

Aunque F1-score es bastante popular y adecuado para tratar problemas de clases no balanceadas, se centra solo en la clase positiva. Por lo tanto, Accuracy es usado como medida adicional, ya que es un buen indicador del desempeño general del sistema.

En la tabla 3 se muestran los resultados

	Prec	Rec	F1	Acc
Partícula	97.05	90.06	93.70	99.71
Ámbito	84.46	78.54	81.34	96.06

Cuadro 3: Resultados obtenidos

En la detección de las partículas de la negación el recall alcanza menor valor que la precisión. Esto es debido principalmente a la dificultad de identificar los marcadores que constituyen más de una palabra, solucionados en la mayoría de los sistemas desarrollados con post procesamientos.

Los resultados presentados aquí no pueden ser directamente comparados con investigaciones previas ya que, según se conoce, no hay investigaciones relacionadas con el reconocimiento de la negación en el corpus utilizado. Además, los pocos trabajos realizados en español utilizan distintos enfoques y los resultados del tratamiento de la negación no son presentados en ellos. A pesar de esto, la comparación con algunos trabajos en otros idiomas como el inglés sería un buen medidor de los resultados expuestos previamente. Luego, en la tabla ?? se presenta el rendimiento de algunos sistemas descritos en la sección 2.

	Cue			Scope		
	P	R	F1	P	R	F1
Enger	90.15	93.56	91.82	85.49	80.28	82.80
Chowdhury	93.42	91.29	92.34	81.53	82.44	81.98
Lapponi	89.17	93.56	91.31	82.25	82.16	82.20

Cuadro 4: Results of negation cue and scope detection of the approaches developed by related work.

Como se ve, a pesar de las diferencias entre los dominios e idiomas utilizados, se obtienen resultados competitivos con respecto a otros trabajos previamente desarrollados.

5. Conclusiones

En este trabajo se ha presentado una primera aproximación al tratamiento de la negación. La propuesta realizada se basa en técnicas de aprendizaje supervisado, detectando tanto las partículas de la negación como su ámbito en textos en español.

La propuesta anterior puede ser mejorada de varias maneras. Solo se tuvo en cuenta las palabras que determinan negación, ignorando algunos prefijos que también la expresan (in-, im-, dis-, etc), ya que en el corpus utilizado no se encontraba anotado este tipo de negación. Un tratamiento especial de los marcadores que están formados por más de una palabra también se cree que mejoraría el rendimiento, ya que estos constituyen la mayor parte de los errores en esta fase. Además, en otros trabajos la tarea de resolución del ámbito de la negación ha sido resuelta como una tarea de etiquetar secuencias, usando CRF como algoritmo de aprendizaje, y se cree que la aplicación de un enfoque similar mejorará los resultados obtenidos.

Referencias

- [1] Shashank Agarwal, Hong Yu. *Detecting Hedge Cues and their Scope in Biomedical Literature with Conditional Random Fields*. 2010.
- [2] Jorge Carrillo de Albornoz, Laura Plaza, Alberto Díaz, Miguel Ballesteros. *UCM-I: A Rule-based Syntactic Approach for Resolving the Scope of Negation*. 2012.
- [3] Emilia Apostolova, Noriko Tomuro, Dina Demner-Fushman. *Automatic Extraction of Lexico-Syntactic Patterns for Detection of Negation and Speculation Scopes*. 2011.
- [4] Miguel Ballesteros, Alberto Díaz, Virginia Francisco, Pablo Gervás, Jorge Carrillo de Albornoz and Laura Plaza. *UCM-2: a Rule-Based Approach to Infer the Scope of Negation via Dependency Parsing*. 2012.
- [5] Valerio Basile, Johan Bos, Kilian Evang, Noortje Venhuizen. *UGroningen: Negation detection with Discourse Representation Structures*. 2012.
- [6] Wendy W. Chapman, Will Bridewell, Paul Hanbury, Gregory F. Cooper. *A Simple Algorithm for Identifying Negated Findings and Diseases in Discharge Summaries*. 2001.
- [7] Faisal Mahbub Chowdhury. *FBK: Exploiting Phrasal and Contextual Clues for Negation Scope Detection*. 2012.
- [8] Isaac G. Council, Ryan McDonald, Leonid Velikovich. *What's Great and What's Not: Learning to Classify the Scope of Negation for Improved Sentiment Analysis*. 2010.

- [9] Roberto Costumero, Federico Lopez, Consuelo Gonzalo-Martín, Marta Millan, Ernestina Menasalvas. *An Approach to Detect Negation on Medical Documents in Spanish*. 2014.
- [10] Noa P. Cruz Díaz. *Negation and Speculation Detection in Medical and Review Texts*. 2014.
- [11] Martine Enger, Erik Velldal, Lilja Øvrelid. *An open-source tool for negation detection: a maximum-margin approach*. 2017.
- [12] Federico Fancellu, Adam Lopez, Bonnie Webber. *Neural Networks For Negation Scope Detection*. 2016.
- [13] Salud M. Jiménez Zafra, Eugenio Martínez Cámara, M. Teresa Martín Valdivia, M. Dolores Molina González. *Tratamiento de la Negación en el Análisis de Opiniones en Español*. 2015.
- [14] Salud M. Jiménez-Zafra, Mariona Taulé, M. Teresa Martín-Valdivia, L. Alfonso Ureña-López, M. Antonia Martí. *SFU Review_{SP}-NEG: a Spanish corpus annotated with negation for Sentiment Analysis*. 2016.
- [15] Salud M. Jiménez-Zafra, M. Teresa Martín-Valdivia, L. Alfonso Ureña-López, M. Antònia Martí, Mariona Taulé. *Problematic Cases in the Annotation of Negation in Spanish*. 2016. ExProM 2016, 42.
- [16] Emanuele Lapponi, Jonathon Read, Lilja Øvrelid. *Representing and Resolving Negation for Sentiment Analysis*. 2012.
- [17] Emanuele Lapponi, Erik Velldal, Lilja Øvrelid, Jonathon Read. *UiO 2 : Sequence-Labeling Negation Using Dependency Features*. 2012.
- [18] M. Antònia Martí, M. Teresa Martín-Valdivia, Mariona Taulé, Salud M. Jiménez-Zafra, Montserrat Nofre, Laia Marsó. *La negación en español: análisis y tipología de patrones de negación*. 2016. Procesamiento del Lenguaje Natural, 57, 41-48.
- [19] Roser Morante, Walter Daelemans. *A metalearning approach to processing the scope of negation*. 2009.
- [20] Woodley Packard, Emily M. Bender, Jonathon Read, Stephan Oepen, Rebecca Dridan. *Simple Negation Scope Resolution through Deep Parsing: A Semantic Solution to a Semantic Problem*. 2014.
- [21] Johan Reitan, Jørgen Faret, Björn Gambäck, Lars Bungum. *Negation Scope Detection for Twitter Sentiment Analysis*. 2015. Proceedings of the 6th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis.