

PROCESAMIENTO DE LENGUAJE NATURAL PARA EL ANÁLISIS DE LENGUAJE SUBJETIVO*

NATURAL LANGUAGE PROCESSING FOR THE ANALYSIS OF SUBJECTIVE LANGUAGE

Horacio Saggion**

Resumen

Describimos la aplicación de la tecnología de procesamiento de lenguaje natural (NLP) al análisis del lenguaje subjetivo. En particular, nos concentramos en la problemática de la clasificación de opinión de material textual extraído de fuentes de datos relacionados con negocios. Estudiamos la derivación de los valores de opiniones de palabras a partir del recurso léxico SentiWordNet y utilizamos estos valores para la interpretación de texto con el objetivo de obtener la valoración de una opinión a partir de sus palabras y frases. Utilizamos características de las palabras para inducir un clasificador basado en el uso de Máquinas de Vectores de Soporte que alcanzan resultados acordes con el estado del arte. También mostramos experimentos preliminares en los que el uso de resúmenes de opiniones ofrece ventaja competitiva para el problema de clasificación respecto del uso de documentos completos cuando los documentos son extensos y contienen material tanto subjetivo como no-subjetivo.

Palabras clave: procesamiento del lenguaje natural, minería de opiniones, sistemas de clasificación de textos, resumen automático, SentiWordNet.

Summary

We describe the application of natural language processing (NLP) technology to the analysis of subjective language. In particular we concentrate on the problem of opinion classification of textual material extracted from business-related data-sources. We study the derivation of sentiment values for words from the SentiWordNet lexical resource and use them for text interpretation to produce word, sentence, and text-based sentiment features for opinion classification. We use word-based and sentiment-based features to induce a classifier based on the use of Support Vector Machines achieving state of the art results. We also show preliminary experiments where the use of summaries before opinion classification provides competitive advantage over the

* Este trabajo ha sido posible gracias al apoyo de la Royal Society (Reino Unido), que financia el proyecto "NLP Tools for Discourse Analysis in Psychology" (JP090069) y del Ministerio de Ciencia e Innovación (España) a través del programa Ramón y Cajal 2009.

** Universitat Pompeu Fabra (España)/Universidad de Sheffield (Reino Unido). Dirección: C/Tanger 122. E-mail: horacio.saggion@upf.edu

use of full documents when the documents are long and contain both subjective and non-subjective material.

Key words: natural language processing, opinion mining, text classification, text summarization, Sentiwordnet.

1. Introducción

Se acepta generalmente el hecho de que la opinión pública tiene gran impacto en la toma de decisiones de compañías y del gobierno (Saggion and Funk, 2009). La Web se ha convertido en una fuente de información importante en todas las áreas de sociedad; en particular en el campo de Business Intelligence¹, los analistas de negocios se están remitiendo a la Web para obtener información tanto factual como de orden más sutil y subjetiva (opiniones) sobre compañías y productos con el fin de monitorear las opiniones de consumidores y del público en general. Sin embargo sin las herramientas apropiadas, la identificación y rastreo de puntos de vista y opiniones públicas, sobre temas particulares dista de ser trivial.

El campo de análisis de las opiniones se ha convertido en un tema de investigación importante para el procesamiento de lenguaje natural (NLP) (Pang and Lee, 2008) y estudia la opinión o la opinión expresada en un discurso particular en el que una opinión se puede considerar como un enunciado subjetivo sobre un tema definido. La investigación en esta área ha disparado gracias a iniciativas de evaluación las Conferencias de Recuperación de Información (TREC) (Ounis et al., 2008) con la tarea de clasificación de opiniones, La Conferencia de Análisis de Texto (TAC) de generación de Resúmenes de Texto² de opiniones, y el programa de “Defi Fouille de Textes” interesado en la clasificación de opinión (Grouin et al., 2009).

El análisis de la opinión se enfrenta a diversos problemas:

1. Identificar si un fragmento de texto expresa una opinión o no;
2. Identificar quién es la entidad que expresa la opinión;
3. Identificar el objeto o tema de opinión;
4. Identificar la polaridad de la opinión (positiva, negativa).

Tabla 1. Ejemplos de fragmentos con y sin manifestación de opinión tomados de críticas de cine

Text Fragment	Opinion?
(1) Soul Survivors tells the story of four college-bound friends... Driving back from a creepy, gothic party, the quartet gets into a horrific car crash...	NO

¹ Nota del traductor: El término *Business Intelligence* no aparece traducido en el texto pues el término se usa generalmente en inglés.

² <http://www.nist.gov/tac/tracks/2008/summarization>

Text Fragment	Opinion?
(2) There's <u>no captivating</u> dialogue, <u>no character chemistry</u> exists anywhere... Soul Survivors is <u>so awful</u> I feel compelled to knock on doors and <u>warn</u> people about it.	YES
(3) A young Northern Iranian schoolboy, Ahmed, wishes to return a school notebook to his classmate, who is threatened with expulsion from school if he does not bring in his notebook to school for tomorrow's class.	NO
(4) ...but this is still <u>great</u> cinema, something American filmmakers would be wise to observe... <u>What more can you ask for?</u>	YES

En la Tabla 1 mostramos ejemplos de fragmentos de texto que expresan y no expresan opiniones extraídos de un corpus de Críticas de Cine³. Estos fragmentos ejemplifican el problema acerca de si el texto expresa o no una opinión (problema 1). Todas las críticas están hechas por un crítico quien es la entidad expresando la opinión (problema 2), sin embargo vale notar que existen casos más sutiles que estos, como en el texto: “*La compañía manifestó que las medidas eran inapropiadas*” en que la entidad que expresa la opinión es una compañía. Todos los textos de la Tabla 1 son sobre cine, y por lo tanto el texto contiene enunciados subjetivos u opiniones sobre el film comentado, sin embargo las opiniones pueden referirse a diferentes características del film, como los actores, el argumento, el director, etc. (problema 3). Respecto del problema 4 el texto (2) es negativo y el texto (4) es positivo. La cualidad de la opinión puede variar desde muy negativa a muy positiva. La identificación de las sutilezas en las opiniones se considera un problema difícil (Pang and Lee, 2008).

Nótese que no es trivial para una máquina detectar qué fragmentos textuales contienen opiniones, ya que las palabras individuales son generalmente ambiguas y pueden ser utilizadas en situaciones completamente neutrales o fácticas, como en “*el chico vivía en un barrio pobre*”, o para expresar una opinión tal como “*la calidad del servicio es pobre*”.

En este trabajo nos interesamos por el problema de clasificación de opiniones tanto en una escala binaria (positivo/negativo) como en una escala matizada desde muy negativo a muy positivo. Hemos desarrollado un conjunto de herramientas para interpretar opiniones y para utilizar modelos estadísticos de Máquinas de Vectores de Soporte para entrenar sistemas de clasificación para bases de datos en inglés. Los aparatos analíticos y recursos son implementados y accesibles mediante el sistema GATE, una plataforma de procesamiento del lenguaje natural para análisis robusto de texto.

Este trabajo se organiza de la siguiente manera: la Sección 2 presenta trabajos relacionados con el área de análisis de opiniones. En la Sección 3 introducimos nuestro

³ <http://www.cs.cornell.edu/people/pabo/movie-review-data/>

estudio de casos sobre clasificación de opiniones, y en la Sección 4 describimos los instrumentos que utilizamos para análisis de documentos textuales e implementamos el sistema de clasificación. La Sección 5 describe los experimentos y los resultados, y la Sección 6 cierra el trabajo con las conclusiones y los trabajos futuros.

2. Trabajos relacionados

La clasificación de reseñas en categorías positivas o negativas es un problema común en minería de opiniones, y una variedad de técnicas han sido utilizadas para encarar este problema, incluyendo aprendizaje automatizado supervisado (Saggion & Funk, 2009) y no supervisado (Zagibalov, 2007). Algunos trabajos apuntan a desarrollar listas de palabras para categorías como positivo o negativo, a veces las palabras de estas listas tienen ponderaciones asociadas, que indican cuán positivas o negativas son. Estas listas son utilizadas para análisis de textos en un proceso de búsqueda en diccionario, y luego las palabras encontradas son usadas para derivar puntajes para oraciones y/o textos, agregando las ponderaciones o la cantidad de las palabras encontradas. Hatzivassiloglou & McKeown (1997) observaron contextos sintácticos donde aparecen adjetivos con igual/opuesta orientación (“interesante y útil” vs. “hermoso pero aburrido”) y desarrollaron un método no supervisado capaz de separar adjetivos en dos series, tal que los elementos de una misma serie tienen similar orientación, mientras que los elementos de series opuestas tendrán orientaciones opuestas. Una heurística les permite establecer cuál de las series contiene los adjetivos con orientación positiva y cuál los de orientación negativa.

El método de Turney (2002) de derivación de listas de palabras consiste en determinar la orientación semántica de patrones lexicales, al calcular su información mutua específica (Point-wise Mutual Information (PMI)) basado en la probabilidad de colocaciones (Church and Hanks, 1990) de las palabras de referencia *excelente* y *pobre*. Las palabras son entonces usadas para la interpretación y clasificación de textos, obteniendo un sistema de clasificación de un 74% de exactitud para la distinción entre positivo/negativo.

El recurso lingüístico SentiWordNet ha sido desarrollado recientemente para la comunidad de investigación (Esuli & Sebastiani, 2006; Baccianella et al., 2010). Es un recurso lexical que agrega información de opiniones a los sentidos de la palabra en la base de datos WordNet (Fellbaum, 1998). En SentiWordNet cada sentido de palabra tiene tres puntajes numéricos para la objetividad, subjetividad y neutralidad del sentido de la palabra. Cada puntaje se encuentra entre 0 y 1, y su suma resulta en 1. La dificultad en utilizar la información de opiniones de SentiWordNet reside en que una palabra encontrada en un texto tiene que ser desambiguada antes de que sus puntajes puedan ser usados. Un adjetivo como *bueno* tiene 24 interpretaciones o significados diferentes en SentiWordNet (por ejemplo, *bueno* como sinónimo de *bien* o como sinónimo de *justo*). Entonces, desambiguar la ocurrencia de la palabra *bueno* en

el texto sería una tarea difícil. Para el lector interesado, en (Saggion and Funk, 2010) hemos propuesto una solución práctica a este problema.

En el dominio de las noticias financieras, Devitt y Ahmad (2007) están interesados en dos problemas relacionados con las noticias financieras: identificar la polaridad de un artículo de noticias, y clasificar un texto en una escala fina de 7 puntos (desde muy positivo hasta muy negativo). Proponen un clasificador básico para la distinción entre positivo/negativo, que tiene una exactitud de un 46%, y tienen también un clasificador más sofisticado, basado en la cohesión lexical y en SentiWordNet, alcanzando un 55% de exactitud.

Dave et al. (2003) presentan numerosas técnicas para crear características (palabras o términos) y puntajes asociados, a partir del entrenamiento de corpora para una tarea de clasificación que consiste en tamizar declaraciones positivas y negativas asociadas a reseñas de productos. Su clasificador combina puntajes de características para oraciones y basa la clasificación en el signo del puntaje combinado. Ghose et al. (2007) investigan la posibilidad de generar, de manera objetiva, un lexicon de expresiones para opiniones positivas y negativas, al correlacionar las ganancias de la compañía con las reseñas.

3. Estudio de casos

En esta investigación nos interesa contabilizar la reputación de entidades de negocios tales como compañías, investigando y teniendo en cuenta información cualitativa y cuantitativa, donde la información cuantitativa consiste en declaraciones fácticas sobre compañías, y la información cualitativa se refiere a las opiniones del público acerca de los productos, servicios y de la compañía en sí misma. No estamos describiendo aquí la identificación de información fáctica de la compañía, sino más bien el trabajo que fue realizado sobre información cualitativa, más específicamente en el desarrollo de un clasificador de opiniones. Para armar nuestro estudio hemos seleccionado y recolectado reseñas de tres fuentes de datos:

- DataSet I: reseñas positivas y negativas acerca de productos, servicios y compañías, del foro *click2complaints*. 92 reseñas: 67% positivas y 33% negativas.
- DataSet II: reseñas detalladas en una escala de 5 puntos (estrellas) acerca de productos, servicios y compañías, del foro *pricegrabber*. 7.300 reseñas con la siguiente distribución de clase: 7,8% 1*, 2,3% 2*, 3,2% 3*, 18,9% 4*, y 67,9% 5*.
- DataSet III: reseñas detalladas en una escala de 5 puntos (estrellas) acerca de productos, servicios y compañías, del foro *ciao*. 89 reseñas largas con la siguiente distribución de clase: 19% 1*, 12% 2*, 10% 3*, 32% 4*, y 27% 5*.

Ejemplos de la reseña se muestran en las Tablas 2 y 3.

Tabla 2. Reseñas detalladas

Reseñas	Clase
Brillante. Simplemente brillante. Ordené justo antes del tiempo de cierre a las 5:30 pm, seleccioné la opción más básica de entrega (City-link) y mi paquete (procesador, ram, discos duros, cd drive, motherboard, tarjeta gfx, etc.) llegaron a las 9 am del día siguiente! Extremadamente satisfecho con el servicio, la atención al cliente es muy amigable y competente!	5 *
Realmente proveen buenos productos a buenos precios. No obstante, he descubierto que el servicio post-venta es extremadamente pobre, y yo ahora uso Tekheads solo cuando no hay otras opciones. De hecho, felizmente pagaré un peso o dos más para conseguir mejor servicio en otro lado.	2 *

Tabla 3. Reseñas positivas y negativas

Reseñas	Clase
El sitio <i>web</i> básicamente ahorró mucho tiempo, era rápido y fácil para enviar los trabajos y los comerciantes que vinieron a evaluar eran mayormente locales. Hay una sección de comunicación de opiniones y entonces los constructores saben que tendrán un buen comentario para que otros clientes vean cuando hacen un buen trabajo.	pos
Tomaron mal nuestra dirección en los diarios financieros, el esposo tuvo que hacer un viaje de ida y vuelta de 60 millas o más para firmar un nuevo acuerdo! No lo querían enviar por correo, ni por envío certificado o por courier. Qué pérdida de tiempo... Entonces, quizás cuando el Sr Ex-Eastenders+ nos hable de las maravillosas ofertas que se obtienen en la TV, él debería mencionar también el servicio más bien pobre que acompaña los productos.	neg

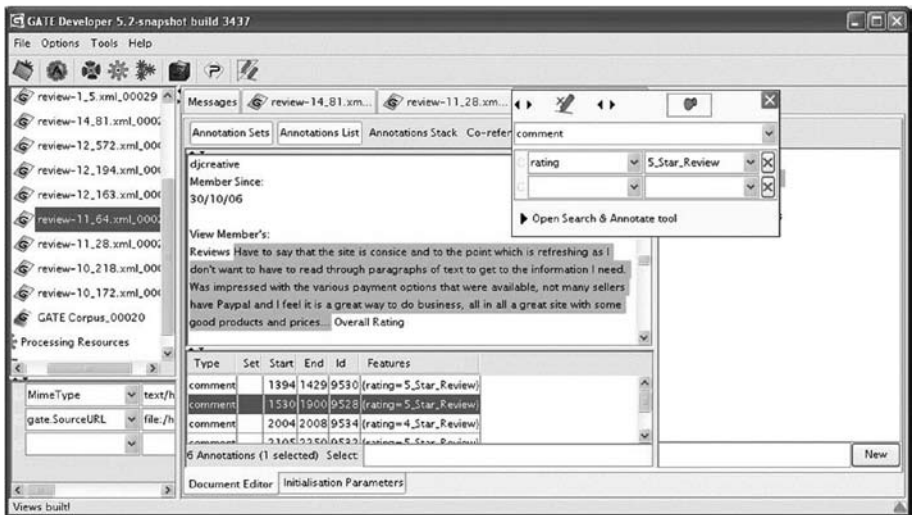
Porque estos grupos de datos están, por un lado, sesgados (reseñas), y, por el otro lado, ya están calificados con clases positivas y negativas o escalas minuciosas, decidimos encarar el problema como de clasificación supervisada. El enfoque adoptado consiste en transformar cada texto junto con su clase en una instancia de aprendizaje, definir un clasificador a partir de los documentos, y luego aplicarlo a un conjunto de documentos no revisados para verificar la exactitud del clasificador (el ratio de casos correctamente clasificados sobre el número total de casos).

4. Aparato de análisis de textos

Como herramienta analítica para procesar los documentos utilizamos componentes del sistema GATE (http://gate.ac.uk), una plataforma para el desarrollo y despliegue de sistemas de procesamiento del lenguaje natural. Una de las reseñas puede ser vista en el Cuadro 1. De este instrumento utilizamos componentes muy simples y estandarizados tales como:

- Un tokenizador para segmentar el texto en palabras y derivar numerosas características ortográficas.
- Un divisor de oraciones para identificar oraciones.
- Un marcador de partes del discurso para asociar categorías lexicales (sustantivos, verbos, etc.) a las diferentes unidades de palabras.
- Un analizador morfológico para computar raíces de palabras o lemmas.

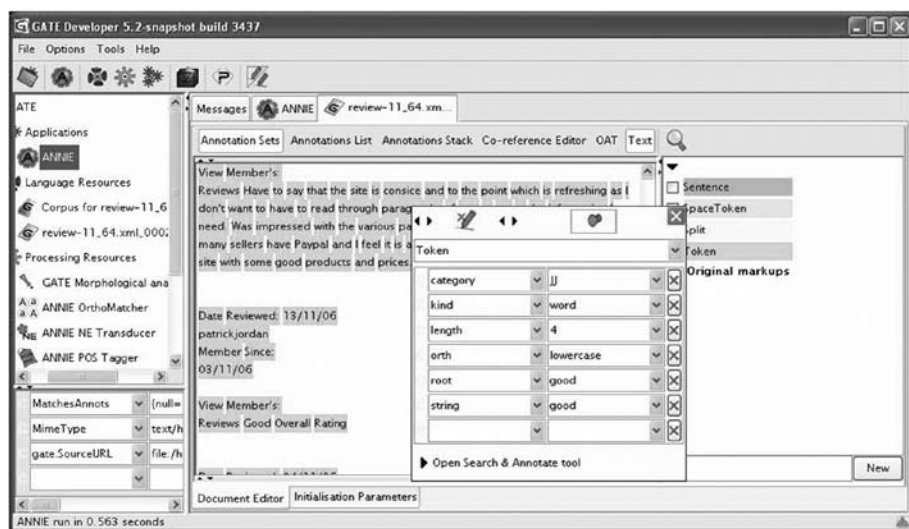
Cuadro 1. Reseña con su clase en la Interface de Usuario del GATE



El Cuadro 2 muestra comentarios lingüísticos asociados a la reseña. Además de los componentes lingüísticos básicos, relevamos información de opiniones para cada palabra en la reseña, utilizando nuestra propia interpretación de los valores de positividad o negatividad de las palabras en el SentiWordNet (Ver Saggion & Funk (2010) para detalles de este procedimiento). Debido a que realizar la desambiguación del sentido de una palabra antes de usar SentiWordNet sería impráctico, relevamos la opinión general de cada palabra en SentiWordNet: para cada palabra nos interesa saber si es generalmente “positiva” (p. e., la palabra tiene más entradas en la base de datos con una interpretación positiva que con una interpretación negativa), generalmente “negativa” (p. e., la palabra

tiene más entradas negativas que positivas), o te otro modo “neutral”. Este procedimiento ha sido implementado como un componente en el sistema GATE y puede ser utilizado para asociar características positivas, negativas o neutrales de cada palabra en un documento. Con este enfoque, una palabra como *bueno* tendrá un valor positivo para el rasgo SentiWordNet, y una palabra como *inútil* tendrá una interpretación negativa. En adición a este rasgo de opinión estamos interesados en relevar rasgos de opiniones en los niveles del la oración y el texto. Para cada oración contamos el número de palabras positivas y negativas que contiene; consideramos una oración *positiva* si contiene más palabras positivas que negativas, y la consideramos *negativa* si contiene más palabras negativas que positivas. En el nivel del texto nos interesa el número de oraciones positivas, negativas y neutrales que contiene y también en un valor total, teniendo en cuenta el número de palabras positivas y negativas encontradas en toda la reseña. Si hay más palabras positivas que negativas, el puntaje es positivo; si hay más palabras negativas que positivas, el puntaje es negativo.

Cuadro 2. Análisis lingüístico de la reseña

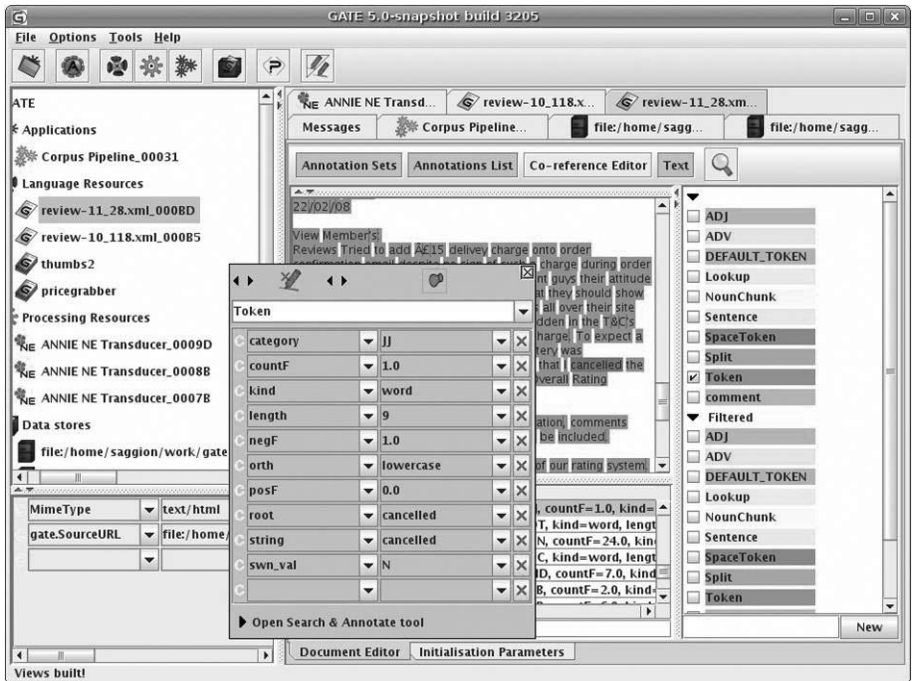


Además del procesamiento básico del texto y del relevamiento de rasgos de opiniones incluimos también un nivel de análisis que identifica adjetivos, adverbios, y sus combinaciones binarias porque han sido señaladas como relevantes para la clasificación de opiniones (Turney, 2002). Este procedimiento nos permite extraer rasgos tales como “bueno”, “malo”, “bastante”, “suficientemente”, “no”, “bastante bueno”, “muy bueno”, “no malo”, etc. que pueden ayudar a identificar orientaciones positivas o negativas. Toda esta información se muestra en el Cuadro 3.

4.1. Herramienta de aprendizaje automatizado

Para desarrollar un sistema de clasificación nos basamos en la tecnología disponible de aprendizaje automatizado en el sistema GATE. En particular, hemos utilizado una biblioteca de Máquinas de Vectores de Soporte (MVSs) disponible en GATE para entrenar y testear un sistema de clasificación de textos en nuestros grupos de datos. Los MVSs son algoritmos de aprendizaje supervisado muy competitivos, los cuales son particularmente apropiados para problemas en que cada instancia es representada en un enorme espacio multi-dimensional (Joachims, 1998), por ejemplo en la clasificación de textos, un espacio consistente en un enorme vocabulario. Durante el entrenamiento, dado un conjunto de instancias positivas y negativas en un espacio vectorial, el algoritmo encuentra un hyper-plano para separarlos. Durante la prueba, una nueva instancia oculta es mapeada en el espacio vectorial y es clasificada como positiva o negativa, dependiendo del lado del hyper-plano en que fue comprendida. La implementación de MVSs en el GATE nos oculta las complejidades asociadas a la extracción de rasgos y la representación de instancias. Detalles del sistema pueden encontrarse en la guía para el usuario de GATE (<http://gate.ac.uk>).

Cuadro 3. Interpretación de la opinión en la reseña



5. Sistemas de clasificación y experimentos

Se testearon dos sistemas de clasificación diferentes. Un clasificador utiliza información lexical obtenida del análisis lingüístico básico aplicado a cada reseña (palabra, raíz, partes del discurso e información ortográfica). El segundo clasificador utiliza información derivada del recurso lexical SentiWordNet en el nivel de las palabras, las oraciones y el texto, y de los rasgos adjetivos y adverbiales. Los experimentos consistieron en 10 sesiones de experimentos de validación cruzada registrando el promedio de exactitud de los diez experimentos.

5.1. Clasificación basada en palabras

En el DataSet I el mejor clasificador fue obtenido por la combinación de información de raíz con ortografía, este sistema alcanzó el 80% de exactitud de clasificación (80% de positivos y 80% de negativos).

En el DataSet II, el mejor clasificador fue obtenido utilizando solo información de raíz. El clasificador obtuvo en este caso 74% de exactitud de clasificación con buen desempeño para las categorías extremas (80% para 5 estrellas y 75% para 1 estrella) pero un pobre desempeño para categorías de clase media.

5.2. Clasificación basada en opinión

En el DataSet I el uso de rasgos de opiniones proporcionó un clasificador con 76% de exactitud, con mejor performance que el clasificador basado en las palabras en la categoría positiva (82% de exactitud) y peor performance en la categoría negativa (61%). Mientras que globalmente la performance es menor que la del clasificador basado en las palabras, hay mucha variación en los diferentes experimentos como para concluir que un clasificador es mejor que el otro.

En el DataSet II, el uso de los rasgos de opinión proporcionó un clasificador con un 72% de exactitud global y con una mejor performance que el clasificador basado en las palabras para las categorías de rango medio. Nuevamente, no se pueden sacar conclusiones respecto de cuál clasificador es mejor.

5.3. Clasificador de opiniones basado en resumen

En el DataSet III todos los clasificadores que testearon tuvieron una pobre performance (niveles de exactitud entre 30% y 40%). Atribuimos esto al tamaño de las reseñas, que eran muy largas, y la escala de clasificación detallada. Decidimos entonces investigar si aplicando las técnicas de resumen antes de la clasificación podría ayudar al algoritmo de aprendizaje automatizado.

El resumen de textos apunta a proveer a los lectores de versiones condensadas de documentos, eliminando información innecesaria y manteniendo lo que es esencial en el documento. Aunque el trabajo en resúmenes automáticos de textos comenzó a finales

de los cincuenta, la última década ha mostrado un interés renovado en este problema. Numerosos e importantes avances han sido observados, especialmente en el área de la evaluación de los resúmenes de textos.

La evaluación extrínseca de los resúmenes de textos consiste en testear si un resumen puede ser utilizado en lugar del texto completo, para cumplir con una tarea específica. Los resúmenes automáticos han demostrado ser útiles en tareas tales como la clasificación o categorización de textos, respuesta a preguntas, y co-referencias de documentos cruzados. Aquí investigamos si el resumen de textos puede ayudar a detectar la opinión que un texto expresa. En otras palabras, queremos verificar si reduciendo el ruido mediante los resúmenes de los textos en documentos extensos puede ayudar a un clasificador de textos a llevar a cabo la tarea de clasificación de opiniones.

Para este fin, hemos desarrollado marcos de investigación para testear diferentes tipos de resúmenes como posibles documentos sucedáneos en la clasificación de opiniones. Los experimentos completos son detallados en la investigación de Lloret et al. (2010). Aquí mencionamos los resultados positivos que obtuvimos con el uso de resúmenes basados en consultas, que son conjuntos de oraciones del documento, tal que contienen una mención de la entidad que está siendo revisada (p. e., un banco en la DataSet III). Para crear resúmenes experimentales, hemos utilizado el sistema SUMMA (Saggion, 2008), que provee un conjunto de recursos de lenguaje y procesamiento para la creación de resúmenes que pueden ser usados con el sistema GATE. Un clasificador que utiliza tanto los rasgos lexicales como los de opinión alcanza un 56% de exactitud en el DataSet III cuando usa resúmenes (en una tasa de 10% de comprensión) comparado con un 41% de exactitud cuando utiliza los textos completos. Otros experimentos de resúmenes siguen la misma tendencia.

6. Conclusión

Hemos presentado nuestro trabajo en el desarrollo y uso de un conjunto de instrumentos para la clasificación de opiniones basadas en el texto, basada en el uso de técnicas de aprendizaje automatizado supervisado. Nuestro trabajo se basa en el análisis lingüístico de textos utilizando herramientas básicas del sistema GATE, y en la interpretación de opiniones en el nivel de las palabras, las oraciones y del texto siguiendo la agrupación de valores de positividad y negatividad en SentiWordNet. Nuestros experimentos muestran que podemos clasificar textos cortos en inglés de acuerdo al índice (el valor positivo o negativo de las opiniones) alcanzando resultados comparables al estado del arte.

En general, hemos notado que un simple clasificador basado en palabras tiene mejor comportamiento (promedio de exactitud más alto) que el clasificador basado en las opiniones en los sets de datos que hemos utilizado. No obstante, los niveles de exactitud varían mucho y no podemos concluir que un clasificador sea mejor que el

otro. Mientras que un clasificador basado en las palabras tiene mejor comportamiento desde el punto de vista global, un clasificador basado en las opiniones tiene mejor comportamiento en las categorías difíciles. Un clasificador basado en las opiniones que utilice información de un lexicon externo puede también proveer información sobre ítems no revisados en el set de entrenamiento.

Dada la dificultad de clasificación de reseñas largas en clases detalladas, introdujimos y estamos aún investigando un marco para testear el valor del resumen de textos para la clasificación de opiniones. Hemos obtenido resultados positivos con el uso de resúmenes basados en la entidad y en la consulta. Estamos trabajando actualmente en una forma diferente de medir la exactitud de los clasificadores con el método de los cuadrados mínimos que hemos establecido como las medidas más apropiadas para clasificación en una escala ordinal. En un mediano a largo plazo planeamos extender nuestros experimentos y análisis a otros sets de datos disponibles en diferentes dominios, tales como reseñas de libros y de películas, para verificar si los resultados pueden verse influenciados por la naturaleza del corpus.

Bibliografía

Baccianella, S.; Esuli, A.; Sebastiani, F. “SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining”. In *Proceedings of LREC 2010*.

Churck, K. & Hanks, P. (1990). Words associations norms, mutual information and lexicography. *Computational Linguistics*, 16(1), pp. 22-29.

Dave, K.; Lawrence, S. & Pennock, D.M. “Mining the Peanut Gallery: Opinion Extraction and Sentiment Classification of Product Reviews”. In *Proceedings of the WWW 2003*.

Devitt, A. & Ahmad, K. (2007). “Sentiment Polarity Identification in Financial News: A Cohesion-based Approach”. In *Proceedings of the ACL 2007*.

Esuli, A. & Sebastiani, F. (2006). “SentiWordNet: A Publicly Available Lexical Resource for Opinión Mining”. In *Proceedings of LREC 2006*.

Fellbaum, C. (1998). *WrodNet - An Electronic Lexical Database*. The MIT Press.

Ghose, A.; Ipeirotis, P.G. & Sundararajan, A. (2007). “Opinion Mining using Econometrics: A Case Study on Reputation Systems”. In *Proceedings of the ACL 2007*.

Grouin, C.; Hurault-Plantet, M.; Paroubek, P. & Berthelin, J.-B. (2009). “DEFT’07: une campagne d’évaluation en fouille d’opinions”. *RNTE E-17*, pp. 1-24.

- Hatzivassiloglou, V. & McKeown, K. (1997). "Predicting the semantic orientation of adjectives". In *Proceedings of EACL 1997*.
- Lloret, E.; Saggion, H. & Palomar, M. (2010). "Experiments on Summary-based Opinion Classification". In *Proceedings of the NAACL Workshop on Computational Approaches to the Analysis and Generation of Emotion in Text*.
- Ounis, I.; McDonald, C. & Soboroff, I. (2008). "Overview of the TREC-2008 Blog Track". In *Proceedings of TREC 2008*.
- Pang, B. & Lee, L. (2008). "Opinion mining and sentiment analysis". *Foundations and Trends in Information Retrieval*, 2(1-2), pp. 1-135.
- Saggion, H. "SUMMA: A Robust and Adaptable Summarization Tool". *TALN*, 49(2), pp. 103-125.
- Saggion, H. & Funk, A. (2010). "Interpreting SentiWordNet for Opinion Classification". In *Proceedings of LREC 2010*.
- Saggion, H. & Funk, A. (2009). "Extracting Opinions and Facts for Business Intelligence". *RNTI E-17*, pp.119-146.
- Turney, P. (2002). "Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews". In *Proceedings of ACL 200*.
- Zagibalov, T. & Carroll, J. (2008). "Unsupervised classification of sentiment and objectivity in Chinese text". In *Proceedings of IJCNLP*.

Fecha de recepción: 15/12/09
Fecha de aceptación: 10/05/10