



Caracterizando a los usuarios a través de sus opiniones

Jesus Serrano-Guerrero, Francisco P. Romero, Jose Ruiz-Morilla, Jose A. Olivas

Departamento de Tecnologías y Sistemas de Información
Escuela Superior de Informática (UCLM), Ciudad Real, 13071, España
jesus.serrano@uclm.es, joseangel.olivas@uclm.es, franciscop.romero@uclm.es

Resumen—El análisis de sentimientos basados en aspectos permite detectar comportamientos de ciertos usuarios. Estos comportamientos necesitan ser modelado para poder entenderse la opinión final sobre un producto o servicio. Este artículo propone el modelado de cada opinión final en base a la agregación borrosa de las distintas opiniones sobre los distintos aspectos que conforman el producto opinado. Igualmente propone algunas aplicaciones para dicho aplicaciones.

Palabras clave—análisis de sentimientos basado en aspectos; operadores OWA; perfil de usuario

I. INTRODUCCION

Cada vez es más frecuente el uso de servicios, especialmente en Internet gracias a plataformas como Amazon o Tripadvisor, que permiten expresar opiniones a los usuarios sobre distintos productos. Detrás de estas opiniones se pueden ocultar distintos comportamientos que es necesario que sean caracterizados con el fin de entender los distintos usuarios y las situaciones en las cuales hacen un comentario determinado u otro.

El campo del Análisis de Sentimientos se centra principalmente en detectar posibles expresiones que sean susceptibles de ser interpretadas como opiniones y cuantificar el grado en el que cada opinión sea más o menos positiva, entre otras tareas. La detección de las opiniones se puede realizar a nivel de sentencia, de aspecto o de documento.

Este trabajo está más centrado en la detección a nivel de aspecto, porque permite ver cómo los usuarios pueden expresar sus preferencias acerca de determinadas características de un producto o servicio descartando otros. Por ejemplo, si un usuario que acaba de adquirir un teléfono hace una crítica feroz acerca de su batería, esto puede suponer que adquirió el teléfono pensando que la batería era uno de sus puntos fuertes y a la vez, se podría inferir que una de las preferencias del usuario era dicha batería. Sin embargo, si el usuario a pesar de hacer una queja sobre un aspecto como podría ser la cámara frontal del móvil no es demasiado efusivo, podría interpretarse que el producto no es bueno, pero tampoco era un detalle importante para el usuario a la hora de tomar su decisión final sobre comprar un móvil u otro.

Todos estos detalles hacen que para entender la opinión final de un usuario sobre un producto o servicio, sea necesario tener en cuenta la importancia de cada aspecto y el grado de positividad o negatividad emitido hacia cada uno de ellos. Es por ello, que se propone el modelado de una opinión final como la agregación borrosa de las opiniones parciales emitidas sobre cada aspecto.

La principal contribución de este artículo es una visión acerca de algunos posibles comportamientos de los usuarios cuando expresan opiniones, especialmente, sobre productos y servicios que compran o utilizan, así como una posible aproximación borrosa que permitiría el modelado de dichos comportamientos, y sus posibles aplicaciones.

El artículo está organizado en las siguientes secciones: la sección II se corresponde con el estado del arte, la sección III presenta la principal motivación del artículo, la sección IV presenta cuál sería la metodología para agregar las distintas opiniones y dar una conclusión final al usuario sobre un producto, la sección V presentará algunas aplicaciones de uso y finalmente, algunas conclusiones serán remarcadas.

II. ESTADO DEL ARTE

Según Cambria, el Análisis de Sentimientos puede ser considerado como un problema de Procesamiento de Lenguaje Natural (PLN) muy restringido, en el que sólo es necesario entender los sentimientos positivos o negativos respecto a cada frase y/o las entidades o temas objetivo [1]. Sin embargo, a pesar de ser un problema restringido, todos los trabajos en este campo, así como todos los trabajos en Recuperación de Información, siempre luchan con problemas de PLN no resueltos (manejo de la negación, reconocimiento de entidades con nombre, desambiguación del sentido de la palabra,...) que son esenciales para detectar claves del lenguaje como la ironía o el sarcasmo [2], y en consecuencia, para encontrar y valorar sentimientos.

Uno de los principales aspectos que debe abordar la PLN son los diferentes niveles de análisis. Dependiendo de si el objetivo del estudio es un texto completo o un documento, una o varias frases enlazadas, o una o varias entidades o aspectos de esas entidades, se pueden realizar diferentes tareas de PLN y Análisis de Sentimientos. Por lo tanto, es necesario distinguir

tres niveles de análisis: (i) nivel de documento, (ii) nivel de frase y (iii) nivel de entidad/aspecto.

A nivel de documento se considera que un documento es una opinión sobre una entidad o aspecto de la misma. Este nivel está asociado con la tarea llamada clasificación de sentimientos a nivel de documento [3]–[6]. Sin embargo, si un documento presenta varias frases que tratan de diferentes aspectos o entidades, entonces el nivel de oración es más adecuado. El nivel de la oración está relacionado con la clasificación de la subjetividad de la tarea; considera cada oración como una opinión positiva, negativa o neutral [7]–[10]. Y finalmente, cuando se necesita información más precisa, entonces surge el nivel de entidad/aspecto. Es el nivel de grano más fino, considera un objetivo sobre el que el ponente de opinión expresa una opinión positiva o negativa. Este último nivel es posiblemente el más complejo, ya que es necesario extraer con gran precisión muchas características, como las fechas o el lapso de tiempo, las diferentes características/espectáculos y entidades a tener en cuenta, así como las relaciones entre ellas, los formadores de opinión y sus características, etc. Está estrechamente relacionado con tareas como Opinion Mining y Opinion Summarization [9], [10].

Muchas tareas surgen vinculadas al Análisis de Sentimientos. Algunas de ellas están estrechamente relacionadas y es difícil separarlas claramente porque comparten muchos aspectos. Los más importantes son:

1. Clasificación de sentimientos: también llamada orientación de sentimientos, orientación de opinión, orientación semántica o polaridad de sentimientos [11]. Se basa en la idea de que un documento/texto expresa la opinión de un titular sobre la entidad y trata de medir el sentimiento de ese titular hacia la entidad. Por lo tanto, consiste principalmente en clasificar las opiniones en tres categorías principales: positivas, negativas o neutras. Parece una tarea simple; sin embargo, es una tarea compleja, especialmente cuando las opiniones provienen de múltiples dominios o idiomas [12], [13]. Esta tarea está estrechamente relacionada con la predicción de la valoración de los sentimientos, que consiste en medir la intensidad de cada sentimiento.
2. Clasificación de subjetividad. Consiste principalmente en detectar si una frase dada es subjetiva o no. Una frase objetiva expresa información objetiva, mientras que una frase subjetiva puede expresar otro tipo de información personal como opiniones, evaluaciones, emociones, creencias, etc. Además, las frases subjetivas pueden expresar sentimientos positivos o negativos, pero no todas lo hacen. Esta tarea puede ser vista como un paso previo a la clasificación de los sentimientos. Una buena clasificación de la subjetividad puede asegurar una mejor clasificación de los sentimientos [14]–[17].
3. Resumen de opiniones. Se centra especialmente en extraer las características principales de una entidad compartida dentro de uno o varios documentos y los sentimientos al respecto [18]. Por lo tanto, se pueden distinguir dos perspectivas en esta tarea: la integración de un solo documento y la integración de varios documentos. La integración de un solo documento consiste en analizar hechos internos presentes en el documento analizado, por ejemplo, cambios en la orientación de los sentimientos a lo largo del documento o vínculos entre las diferentes entidades/características encontradas, y principalmente mostrar aquellos textos que mejor los describen. Por otro lado, en la integración multidocumento, una vez detectadas las características y entidades, el sistema debe agrupar y/o ordenar las diferentes frases que expresan sentimientos relacionados con dichas entidades o características. El resumen final puede presentarse en forma de gráfico o texto que muestre las principales características/entidades y cuantifique el sentimiento con respecto a cada una de ellas de alguna manera, por ejemplo, agregando intensidades de sentimientos o contando el número de frases positivas o negativas [19], [20].
4. Recuperación de opiniones. Intenta recuperar documentos que expresan una opinión sobre una consulta determinada. En este tipo de sistemas, se requieren dos puntajes para cada documento, el puntaje de relevancia frente a la consulta y el puntaje de opinión sobre la consulta, y ambos se utilizan generalmente para clasificar los documentos [21], [22].
5. Sarcasmo e ironía. Se centra en detectar afirmaciones con contenido irónico y sarcástico. Esta es una de las tareas más complicadas en este campo, especialmente debido a la falta de acuerdo entre los investigadores sobre cómo se puede definir formalmente la ironía o el sarcasmo [2], [23], [24].
6. Otros. Además de las actividades anteriormente mencionadas, existen otras tareas relacionadas con el Análisis de Sentimientos, como por ejemplo, la detección de género o autoría, que trata de determinar el género o la persona que ha escrito un texto/opinión [25], [26], la detección de spam de opinión, que trata de detectar opiniones o reseñas que contienen contenidos no confiables publicados para distorsionar la opinión pública hacia personas, empresas o productos [27]–[29], o la identificación de temáticas en función de los términos referidos a sentimientos [30].

III. MOTIVACIÓN

Son muchas las aplicaciones y páginas web que permiten al usuario expresar sus opiniones. Analizando detenidamente las distintas opiniones de forma general, se puede observar que los comportamientos de los usuarios pueden ser muy diversos, pudiendo establecerse distintos tipos de usuario dependiendo de las formas en las que emiten sus opiniones.

De igual manera, son muchas las formas en las que un usuario puede expresar su opinión. Por ejemplo, entre las más típicas podemos encontrar el uso de conjuntos de estrellas como puede verse en muchas páginas webs de venta de productos como Amazon¹ o Joom²:



Fig. 1. Puntuación mediante estrellas

o mediante una barra que permita establecer el grado de conformidad con respecto a un aspecto concreto:



Fig. 2. Barra de puntuación

entre otros posibles mecanismos. Como puede observarse, estos mecanismos permiten de una manera fácil, rápida e intuitiva expresar una opinión, sin embargo, carecen de cierta expresividad que si provee el lenguaje natural. Por ejemplo, las estrellas vistas anteriormente podrían ir acompañadas de distintas etiquetas lingüísticas {No satisfecho, Poco satisfecho, Medianamente satisfecho, Satisfecho, Muy satisfecho}, que pueden resultar más intuitivas e interpretables que las estrellas o su representación numérica: 1 estrella, 2 estrellas, 3 estrellas, etc.

Esto se puede ilustrar mediante el siguiente ejemplo (Figura 3) sacado de la página web TripAdvisor³:



Fig. 3. Opinión con puntuación poco precisa

Si se observan las puntuaciones, la estancia parece haber sido perfecta porque son máximas respecto a todos los aspectos posibles (valor, localización, habitaciones, limpieza,). Sin embargo, si se observan los comentarios esto no es así. El tamaño de la habitación era bueno pero no excelente. Lo mismo ocurre con la calidad del sueño, fue buena pero no excelente. Se podría pensar que este usuario utilizaría la etiqueta lingüística “bueno” como símbolo para representar la máxima calidad, sin embargo, si se lee la primera frase, la localización era “excelente”, por lo que gradación de las puntuaciones podría llegar hasta ese extremo.

A. Algunos posibles tipos de usuarios

Ya hemos visto algún tipo de usuario como puede ser el usuario poco preciso de la imagen anterior, pero podemos ver más posibles usuarios. Por ejemplo, si miramos a la siguiente opinión que expresa de forma fehaciente que el hotel es fantástico:

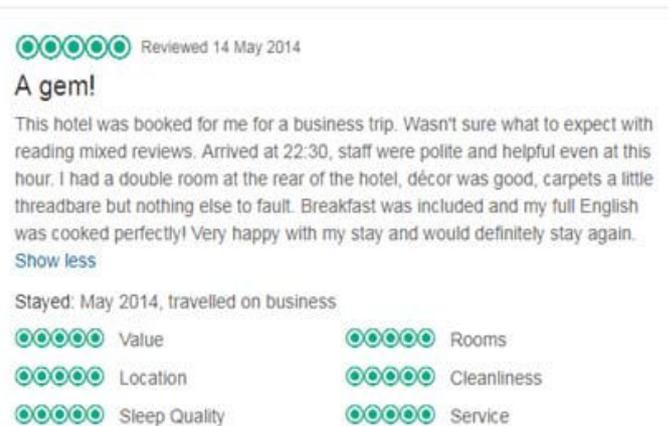


Fig. 4. Opinión excesivamente positiva

contrasta con la siguiente opinión emitida con una semana de diferencia, donde el usuario afirma que el mismo hotel es horrible:

¹ www.amazon.com

² www.joom.com

³ www.tripadvisor.com

Reviewed 20 June 2014

DISGUSTING ROOMS, NO POOL AND EXCEPTIONALLY RUDE MANAGER TRIED TO OVERCHARGE Me.

I booked a single room, rung to confirm my booking the day of arrival only to be told I'd booked a double. I tried to tell the woman (whom I now know to be the manager) there must be a mistake she argued with me, told me that her staff would not have made that mistake then told me I should be more careful with my credit card payments! After telling me that I couldn't have a single room originally, once I made it clear I wasn't paying extra for the double, she magically found a single room.

On arrival the receptionist admitted there had been a mistake when I phoned, that I had indeed booked a single room and they'd mistakenly changed it double. The single room I got must have been the last one available, it's absolutely disgusting. There are holes in the wall, hair in the toilet, the carpet on the stairs looks like the cleaner died in 1975 and it hasn't seen a vacuum since.

I booked this hotel instead the other ones in town for the pool, which again, I checked was available when booking. Arriving and intending to hear straight for the pool, of course it's out of order.

Disgusting hotel, awful manager who is happy to sarcastically tell you to me more careful with your credit card after trying to charge you extra. Wish I'd stayed at the Holiday Inn.
Show less

Stayed: June 2014

- Value
- Location
- Rooms
- Cleanliness
- Service

Fig. 5. Opinión excesivamente negativa

leyendo detenidamente ambos comentarios, puede verse cómo se comparan cosas como el personal o las instalaciones con perspectivas absolutamente distintas. En este caso, podría pensarse que los usuarios pueden ser falsos, es decir, un usuario expresa su opinión simplemente para denigrar la calidad de un hotel, lo que podría conocerse como un “hater”, o un usuario podría intentar mejorar la calidad de un hotel bien porque tiene algún interés escondido, en este caso sería un “lover”. Igualmente, podría ser, simplemente, que el estado de ánimo con el que se emitieron las opiniones no permite expresar de forma excesivamente realista cuál fue la situación real del hotel, o sí.

IV. PROPUESTA

Como puede observarse, modelar el comportamiento de los usuarios puede ser complejo y hay distintos factores a tener en cuenta. Entre estos factores pueden destacarse:

- la importancia que tiene para cada usuario un aspecto determinado dentro del producto del que se está opinando.

- La opinión que tiene el usuario de cada uno de los aspectos que conforman un objeto.
- La opinión global de un producto o servicio vista como el resultado de la agregación de cada uno de los aspectos en función de la importancia.

A. Opinión sobre un aspecto

Así pues, matemáticamente un producto podría definirse como un conjunto A de N aspectos:

$$A = \{A_1, A_2, A_3, \dots, A_N\}$$

Sobre cada aspecto es posible emitir una opinión que puede ser representada mediante un conjunto de etiquetas lingüísticas $\{S_1, S_2, S_3, \dots\}$. Estas etiquetas lingüísticas podrían ser definidas mediante número triangulares borrosos NTB, donde S_i se puede representar por el número $NTB_i = (a_i, b_i, c_i)$. La distribución de los parámetros puede dar lugar a un conjunto de etiquetas balanceado o no. Esto puede permitir tener más capacidad a la hora de filtrar a los usuarios y a los vendedores de los productos.

Véase el siguiente ejemplo. Si la valoración de las opiniones de cada aspecto fuera usando el siguiente conjunto balanceado de 5 etiquetas:

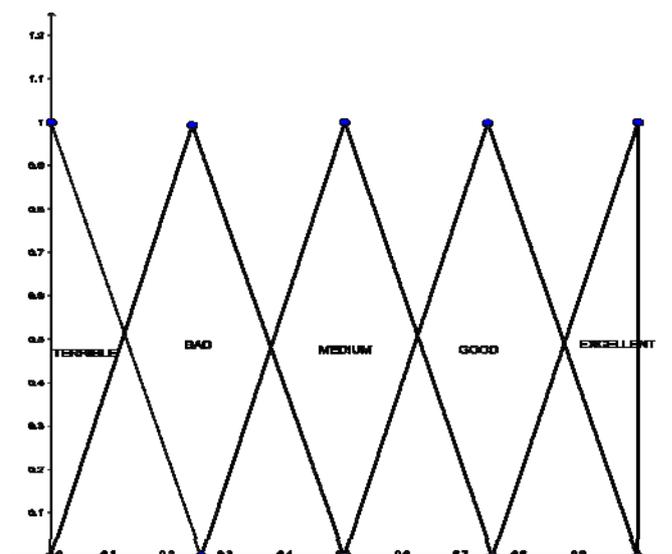


Fig. 6. Distribución de etiquetas balanceada

Podría corresponder intuitivamente con el sistema de estrellas visto en plataformas como Amazon (ver Figura 1), donde una opinión con una única estrella representa la peor valoración, es decir, el primer conjunto de la figura (TERRIBLE), y 5 estrellas representa la mejor valoración, en este caso el último conjunto (EXCELENTE).

Sin embargo, si se utilizara un conjunto de etiquetas como el visto en la siguiente figura, el usuario tendría un mayor poder discriminación, primero porque existen más etiquetas con las que valorar una opinión y segundo porque el soporte de las etiquetas que representan opiniones positivas es mucho menor, por lo que el usuario sería mucho más exigente a la

hora de decir si un producto es BUENO, MUY BUENO o EXCELENTE.

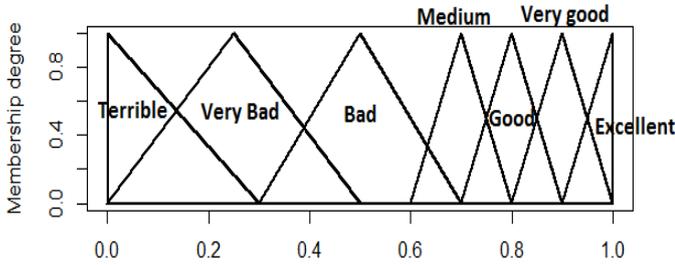


Fig. 7. Distribución de etiquetas no balanceada

B. Importancia de cada aspecto

Cada usuario puede considerar cada aspecto según su propio criterio. Suponer un teléfono móvil, ¿cuál es el aspecto más importante a la hora de comprarlo? Cada usuario pensará uno distinto, unos querrán que la cámara sea de alta definición para hacer fotos, a otros la cámara no les importará y preferirán que la batería le dure más tiempo, otros preferirán que tenga mucha memoria para almacenar muchos archivos, o la que conexión sea 5G, etc.

Por tanto, a la hora de emitir la opinión general sobre un producto, cada aspecto no tiene por qué pesar lo mismo. Así, es necesario modelar el peso de cada aspecto según el usuario.

Este peso puede ser calculado de muchas maneras, por ejemplo, a través de un valor numérico dentro de una escala [0,1]. Siguiendo con el ejemplo del teléfono se podría dar las siguientes importancias a estos posibles aspectos:

- Batería: 0.5
- Pantalla: 0.2
- Conectividad 5G: 0.1
- Cámara: 0.2

En este caso, como puede verse, todos los pesos suman 1, y el más importante sería en este caso el aspecto de la batería, mientras que la conectividad 5G no parece ser nada relevante para este usuario.

Siguiendo una aproximación más intuitiva, la importancia podría ser una variable representada mediante un conjunto de etiquetas $\{S_1, S_2, S_3, \dots\}$, las cuales nuevamente podrían ser definidas mediante número triangulares borrosos NTB, donde S_i se puede representar por el número $NTB_i = (a_i, b_i, c_i)$. La distribución puede ser igual balanceada o no, y como ejemplo podría darse el siguiente conjunto, en el cual el usuario podría seleccionar si un aspecto para él es irrelevante, simplemente importante o de vital importancia:

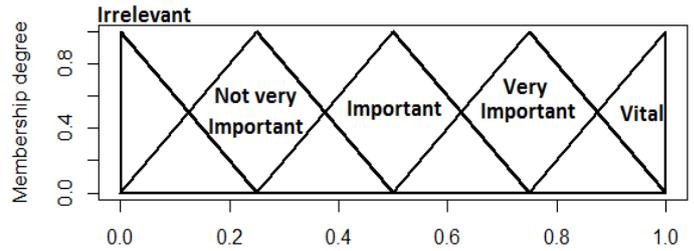


Fig. 8. Ejemplo de etiquetas para medir la importancia de un aspecto

Dependiendo de la forma en la que se mida la importancia, determinará cómo se calcula la opinión general que puede tener un usuario respecto a un producto.

C. Opinión general: Agregación de aspectos

Una vez que el usuario ha determinado cuál es la importancia de cada aspecto, y ha opinado sobre cada uno de ellos, es necesario agregar todas las opiniones individuales pero, teniendo en cuenta cuál es más importante que la otra.

Para ello, se puede proponer el uso de los operadores OWA [31]. Dependiendo de cómo se hayan obtenidos los distintos pesos y las valoraciones de cada aspecto, se podrán utilizar distintas familias de operadores OWA.

En el caso, por ejemplo, de medir a través de la calificación mediante estrellas (ver Figura 1) y que los pesos sean introducidos como valores en el rango [0,1] sumando en su conjunto 1, al ser todos los valores de naturaleza crisp, se podría utilizar un operador OWA clásico. Sin embargo, en el caso de que se utilizaran conjuntos borrosos para obtener los parámetros necesarios, serían necesarios acceder a familias como los OWA de Tipo-1 [32].

V. APLICACIONES

A través de la propuesta anterior, es posible modelar de una forma más fehaciente, cuál puede ser la opinión general de un usuario cuando valora un producto. Entre las aplicaciones que podría encontrarse a esta propuesta podrían mencionarse:

- Detección y clasificación de usuarios: Poder clasificar a los usuarios según sus opiniones puede resultar interesante desde distintos puntos vista. Por ejemplo, un usuario que siempre puntúa positivamente los productos de una marca y negativamente los de la competencia, no tiene por qué implicar que sus opiniones no sean válidas, simplemente es su forma de pensar. Así, un sistema de recomendaciones tendría más información para saber qué productos recomendarle y cuáles no.
- Asignación del grado de confianza de un usuario: Como consecuencia de lo anterior, las opiniones de todos los usuarios no parecen ser igualmente fiables, por lo que la descripción de cada usuario a través de un grado de confiabilidad podría permitir descartar las opiniones de ciertos usuarios. Este proceso hoy es imposible en webs como Amazon o Trivago.

VI. CONCLUSIONES

En este artículo se pone de manifiesto la necesidad de representar las opiniones de los usuarios de forma distinta a la utilizada por muchas plataformas web debido a su falta de expresividad en muchas ocasiones. El uso de aproximaciones borrosas puede ayudar a paliar estos defectos como se ha comentado a lo largo del artículo.

Seguindo el modelo propuesto, sería más fácil descartar opiniones que pueden no tener valor para un usuario, bien porque en sí pudieran ser contradictorias o de poco valor, o porque estuvieran emitidas por usuarios con el propósito de influir en otros usuarios de forma malintencionada.

AGRADEMIENTOS

Este artículo ha sido financiado por FEDER y la Agencia Española de Investigaciones del Ministerio de Economía y Competitividad bajo el proyecto TIN2016-76843-C4-2-R (AEI/FEDER, UE).

BIBLIOGRAFÍA

- [1] E. Cambria, B. Schuller, Y. Xia, and C. Havasi, "New Avenues in Opinion Mining and Sentiment Analysis," *IEEE Intell. Syst.*, vol. 28, no. 2, pp. 15–21, 2013.
- [2] A. Reyes, P. Rosso, and D. Buscaldi, "From humor recognition to irony detection: The figurative language of social media," *Data Knowl. Eng.*, vol. 74, pp. 1–12, 2012.
- [3] C. Zhang, D. Zeng, J. Li, F.-Y. Wang, and W. Zuo, "Sentiment analysis of Chinese documents: From sentence to document level," *J. Am. Soc. Inf. Sci. Technol.*, vol. 60, no. 12, pp. 2474–2487, 2009.
- [4] R. Moraes, F. Valiati, J. and W. P. Gavião Neto, "Document-level sentiment classification: An empirical comparison between SVM and ANN," *Expert Syst. Appl.*, vol. 40, no. 2, pp. 621–633, 2013.
- [5] A. Duric and F. Song, "Feature selection for sentiment analysis based on content and syntax models," *Decis. Support Syst.*, vol. 53, no. 4, pp. 704–711, 2012.
- [6] Y. He and D. Zhou, "Self-training from labeled features for sentiment analysis," *Inf. Process. Manag.*, vol. 47, no. 4, pp. 606–616, Jul. 2011.
- [7] T. Wilson, J. Wiebe, and P. Hoffmann, "Recognizing contextual polarity in phrase-level sentiment analysis," in *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing (HLT '05)*, 2005, pp. 347–354.
- [8] T. Wilson, J. Wiebe, and P. Hoffmann, "Recognizing Contextual Polarity: An Exploration of Features for Phrase-Level Sentiment Analysis," *Comput. Linguist.*, vol. 35, no. 3, pp. 399–433, 2009.
- [9] A. Agarwal, F. Biadys, and K. R. McKeown, "Contextual phrase-level polarity analysis using lexical affect scoring and syntactic N-grams," in *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics (EACL '09)*, 2009, pp. 24–32.
- [10] R. Remus and C. Hähnig, "Towards well-grounded phrase-level polarity analysis," in *Proceedings of the 12th international conference on Computational linguistics and intelligent text processing (CICLing'11)*, 2011, pp. 380–392.
- [11] L.-C.- Yu, J.-L. Wu, P.-C. Chang, and H.-S. Chu, "Using a contextual entropy model to expand emotion words and their intensity for the sentiment classification of stock market news," *Knowledge-Based Syst.*, vol. 41, pp. 89–97, 2013.
- [12] Y. He, C. Lin, and H. Alani, "Automatically extracting polarity-bearing topics for cross-domain sentiment classification," in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (HLT '11)*, 2011, pp. 123–131.
- [13] M. Dragoni and G. Petrucci, "A fuzzy-based strategy for multi-domain sentiment analysis," *Int. J. Approx. Reason.*, vol. 93, pp. 59–73, 2018.
- [14] S. Raaijmakers and W. Kraaij, "A Shallow Approach to Subjectivity Classification," in *Proceedings of the Second International Conference on Weblogs and Social Media (ICWSM '08)*, 2008, pp. 216–217.
- [15] K. Sarvabhotla, P. Pingali, and V. Varma, "Sentiment classification: a lexical similarity based approach for extracting subjectivity in documents," *Inf. Retr. Boston.*, vol. 14, no. 3, pp. 337–353, 2011.
- [16] A. Montoyo, P. Martínez-Barco, and A. Balahur, "Subjectivity and sentiment analysis: An overview of the current state of the area and envisaged developments," *Decis. Support Syst.*, vol. 53, no. 4, pp. 675–679, 2012.
- [17] I. Maks and P. Vossen, "A lexicon model for deep sentiment analysis and opinion mining applications," *Decis. Support Syst.*, vol. 53, no. 4, pp. 680–688, 2012.
- [18] D. Wang, S. Zhu, and T. Li, "SumView: A Web-based engine for summarizing product reviews and customer opinions," *Expert Syst. Appl.*, vol. 40, no. 1, pp. 27–33, 2013.
- [19] K.-M. Park, H. Park, H.-G. Kim, and H. Ko, "Review summarization based on linguistic knowledge," in *Proceedings of the 17th international conference on Database Systems for Advanced Applications (DASFAA'12)*, 2012, vol. 7240, pp. 105–114.
- [20] K. Ganesan, C. Zhai, and E. Viegas, "Micropinion generation: An Unsupervised Approach to Generating Ultra-Concise Summaries of Opinions," in *Proceedings of the 21st international conference on World Wide Web - WWW '12*, 2012, pp. 869–878.
- [21] S.-W. Lee, Y.-I. Song, J.-T. Lee, K.-S. Han, and H.-C. Rim, "A new generative opinion retrieval model integrating multiple ranking factors," *J. Intell. Inf. Syst.*, vol. 38, no. 2, pp. 487–505, May 2011.
- [22] L. Guo and X. Wan, "Exploiting syntactic and semantic relationships between terms for opinion retrieval," *J. Am. Soc. Inf. Sci. Technol.*, vol. 63, no. 11, pp. 2269–2282, 2012.
- [23] E. Filatova, "Irony and Sarcasm: Corpus Generation and Analysis Using Crowdsourcing," in *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC-2012)*, 2012, pp. 392–398.
- [24] A. Reyes and P. Rosso, "Making objective decisions from subjective data: Detecting irony in customer reviews," *Decis. Support Syst.*, vol. 53, no. 4, pp. 754–760, 2012.
- [25] J. Savoy, "Authorship Attribution Based on Specific Vocabulary," *ACM Trans. Inf. Syst.*, vol. 30, no. 2, pp. 1–30, 2012.
- [26] M. Montesi and T. Navarrete, "Classifying web genres in context: A case study documenting the web genres used by a software engineer," *Inf. Process. Manag.*, vol. 44, no. 4, pp. 1410–1430, 2008.
- [27] A. Mukherjee, B. Liu, and N. Glance, "Spotting fake reviewer groups in consumer reviews," in *Proceedings of the 21st international conference on World Wide Web (WWW '12)*, 2012, pp. 191–200.
- [28] S. Xie, G. Wang, S. Lin, and P. S. Yu, "Review spam detection via temporal pattern discovery," in *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2012, pp. 823–831.
- [29] G. Wang, S. Xie, B. Liu, and P. S. Yu, "Identify Online Store Review Spammers via Social Review Graph," *ACM Trans. Intell. Syst. Technol.*, vol. 3, no. 4, pp. 1–21, 2012.
- [30] K. Gutiérrez-Batista, J. R. Campaña, M.-A. Vila, and M. J. Martín-Bautista, "Fuzzy Analysis of Sentiment Terms for Topic Detection Process in Social Networks," in *International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, 2018, pp. 3–14.
- [31] R. R. Yager, "Families of OWA operators," *Fuzzy Sets Syst.*, vol. 59, no. 2, pp. 125–148, 1993.
- [32] S.-M. Zhou, F. Chiclana, R. I. John, and J. M. Garibaldi, "Type-1 OWA operators for aggregating uncertain information with uncertain weights induced by type-2 linguistic quantifiers," *Fuzzy Sets Syst.*, vol. 159, no. 24, pp. 3281–3296, Dec. 2008.