

Aprendizaje profundo para la extracción de aspectos: tarea esencial en la creación y uso de las ontologías

Dionis López Ramos^{1,2} and Leticia Arco García²

¹ Departamento de Informática, Universidad de Oriente, Santiago de Cuba, Cuba.
Avenida de Las Américas s/n, Santiago de Cuba, Cuba

dionis@uo.edu.cu

² Departamento Ciencia de la Computación, Universidad Central “Marta Abreu” de Las Villas. Carretera a Camajuaní km 5 1/2 Santa Clara. Villa Clara. Cuba CP 54830. leticiaa@uclv.edu.cu

Abstract. On the internet, a large amount of information is organized and related through the technologies and components that define the Semantic Web. Other data sources such as blogs, web pages and digital documents contain important information, but it is not well organized or structured. For many institutions and people, it is important to extract and organize this information automatically by using knowledge sources such as ontologies. The efficient use of these sources of knowledge is conditioned by an effective process of information extraction. This process requires performing several natural language processing tasks. One of the most important tasks is the aspect extraction or the extraction of entity characteristics, events or objects in the processed information. An efficient aspect extraction allows an effective use of ontologies. In recent years, researchers have applied models based on Deep Learning algorithms to extracting aspects with good results. This paper is focus on the study of the main results related to the application of Deep Learning techniques in the aspect extraction task. The analysis presented here will facilitate the selection of Deep Learning methods for aspect extraction, as well as the textual representation models to be used that allow obtaining good results, and thus contribute satisfactorily to the use of ontologies.

Keywords: Aspect Extraction, Deep Learning, Ontology, Semantic Web

Resumen. En internet gran cantidad de información es organizada y relacionada a través de las tecnologías y componentes que definen a la Web Semántica. Otras fuentes de datos como los blogs, páginas web y documentos digitales contienen información importante, pero no organizada o estructurada. Para muchas instituciones y personas es importante extraer y organizar esta información de manera automática empleando fuentes de conocimiento como las ontologías. El uso eficiente de estas fuentes de conocimiento está condicionado por un eficaz proceso de extracción de la información. Este proceso requiere realizar varias tareas del procesamiento del lenguaje natural. Una de las tareas más importantes

es la extracción de aspectos o características de entidades, eventos u objetos presentes en la información procesada. Una extracción de aspectos eficiente permite un empleo efectivo de las ontologías. En los últimos años, los investigadores han aplicado los modelos basados en algoritmos del Aprendizaje Profundo a la extracción de aspectos con buenos resultados. En este trabajo se realizó un estudio de las investigaciones relacionadas con estos algoritmos y la extracción de aspectos. El análisis que aquí se presenta facilitará la selección de los métodos de Aprendizaje Profundo para la extracción de aspectos, así como las formas de representación textual a utilizar que permitan obtener buenos resultados, y así contribuir satisfactoriamente al uso de las ontologías.

Palabras Claves: Extracción de Aspectos, Aprendizaje Profundo, Ontologías, Web Semántica

1 Introducción

Las técnicas y componentes de la Web Semántica han permitido la creación de contenidos y recursos como Wikitology [14], DBpedia [6], YAGO [56], Freebase [10] y WordNet [34]. La extracción de conocimiento e información con estos recursos permite la creación de servicios y el análisis efectivo de los datos [45]. Otros datos son publicados en fuentes como blogs, páginas web y documentos digitales de forma no estructurada. Para el procesamiento y recuperación de la información no estructurada se realizan varias tareas del procesamiento del lenguaje natural (PLN). Una de las tareas importantes es la extracción de aspectos porque permite obtener mayor información de los documentos y oraciones de textos digitales. Esta tarea se encarga de extraer aspectos o características de las entidades, eventos u objetos escritos en los documentos. Por ejemplo, en la oración “La comida de ese restaurante es deliciosa y barata” se debe extraer el aspecto “comida” de la entidad “restaurante”. En este ejemplo el aspecto aparece como una palabra simple, pero puede contener frases compuestas que hacen esta tarea aún más compleja [30].

Desde inicios del 2000 han aumentado las investigaciones dirigidas a abordar la extracción de aspectos. Varias propuestas [1,49] para la extracción automática de aspectos han empleado ontologías como: WordNet [34], SentiWordNet [8], WordNet-Affect [54], o ConceptNet [53]. En el estado del arte realizado en [25] se reporta, desde el 2010 hasta la fecha, la existencia de varios trabajos que emplean ontologías como estrategias de búsquedas de información o combinan el empleo del análisis semántico con ontologías. En varios trabajos se realiza la extracción de aspectos como un paso previo al uso de las ontologías [5,52]. Éstas se usan luego para determinar el dominio o el significado correcto de los aspectos según su contexto de ocurrencia. En otras publicaciones, se realiza la extracción de aspectos a partir del uso directo de las ontologías [26,38,55].

Con el objetivo de acelerar el proceso de desarrollo de las ontologías los investigadores en ocasiones tratan de reusar los recursos no ontológicos, como son los esquemas de clasificación, tesauros, lexicones y folcsonomías que han alcanzado

un grado de consenso en la comunidad científica [55]. Mediante la reutilización de recursos no ontológicos se puede reducir el tiempo y los costos asociados al desarrollo de ontologías. Además, se pueden desarrollar ontologías de mejor calidad mediante la reutilización de recursos estandarizados o consensuados por la comunidad de expertos en diversos dominios. Existen varias metodologías que proponen guías para el desarrollo de ontologías: METHONDOLOGY, On-To-Knowledge, DILIGENT, NeOn, entre otras [42].

En NeOn [55] se propone una guía para tener en cuenta recursos no-ontológicos durante la creación de una ontología. La guía plantea un escenario que muestra el caso donde los desarrolladores de ontologías necesitan analizar un recurso no ontológico y decidir, según los requerimientos de la ontología a crear, cuales recursos (i.e., tesauros, glosarios, bases de datos, etc.) deben ser seleccionados para crear la red ontológica. Así, el proceso consta de tres pasos: (1) ingeniería inversa de los recursos no-ontológicos, (2) transformación de los recursos no-ontológicos, e (3) ingeniería avanzada de la ontología.

El primer paso tiene como objetivo analizar un recurso no ontológico para identificar los componentes o información presente y crear la representación de los datos en los diferentes niveles de abstracción (de diseño, requerimientos y conceptual). El segundo tiene como propósito generar un modelo conceptual a partir de estos recursos [55]. Por tanto, la aplicación de técnicas para el PLN y en especial la extracción de aspectos constituyen herramientas de gran utilidad para llevar a cabo los pasos 1 y 2.

Siguiendo ideas muy similares a las planteadas por la metodología NeOn, varios trabajos proponen estrategias para la construcción de ontologías desde conjuntos de documentos. En [60] se muestran los avances más recientes para la construcción de taxonomías desde textos, se definen subareas relevantes, recursos para la evaluación y retos para investigaciones futuras. En [26] se propone una estrategia semántica para extraer taxonomías de dominios desde textos que transita por cuatro etapas. La primera extrae términos de un conjunto de documentos. En la segunda etapa, de estos términos se toman los más relevantes para un dominio específico usando un filtro. En la tercera etapa se desambiguan los términos y son generados los conceptos. En la última etapa se establecen relaciones entre conceptos usando una técnica para asociarlos a una categoría, teniendo en cuenta su coocurrencia en el texto. Una vez generada la nueva taxonomía, ésta puede convertirse automáticamente en una ontología. En [16] se usa una estrategia empleando la Arquitectura General para la Ingeniería de Texto (*General Architecture for Text Engineering*; GATE ¹) para poblar ontologías desde documentos.

El proceso de extracción de términos o conceptos de las propuestas anteriores es un campo de acción para los métodos de extracción de aspectos. En [26] se propone un algoritmo semi-supervisado difuso para minar una ontología de productos a partir de los conceptos que los definen. En este artículo los conceptos están asociados a los aspectos explícitos o implícitos en el texto para cada producto. Para la extracción de aspectos se emplea el método Asignación La-

¹ <https://gate.ac.uk/>

tente de Dirichlet (*Latent Dirichlet Allocation*; LDA) [9]. Cada t3pico o aspecto seleccionado por el algoritmo representa un concepto. Para estimar el grado de inclusi3n entre conceptos se propone un modelo de lenguaje probabilístico. A partir del valor probabilístico de la pertenencia de un concepto con respecto a otro se define una funci3n de relaci3n taxon3mica difusa. Estos trabajos demuestran la utilidad de la extracci3n de aspectos para la creaci3n de ontologías a partir de recursos no ontol3gicos.

Una subtarea de gran relevancia para el análisis de sentimiento y el PLN es la extracci3n de aspectos porque permite obtener mayor informaci3n de los documentos y oraciones de textos digitales. Esta tarea se encarga de extraer aspectos o característicās de las entidades, eventos u objetos escritos en los documentos [30]. La extracci3n de aspectos y las ontologías tienen gran importancia para el análisis de sentimientos en opiniones sobre productos, por ejemplo, en el caso de productos electrodomésticos. Al procesar informaci3n sobre éstos pudiéramos encontrar la siguiente oraci3n: “La batería es muy buena aun cuando se usa el flash y el LCD”. En este ejemplo se pueden extraer como opiniones candidatas “batería-buena”, “flash-buena” y “lcd-buena”. Para una efectiva extracci3n de las relaciones se pueden agrupar las característicās de los productos con una ontología de productos como se muestra en la Fig. 1. Esto permite agrupar aspectos de productos con significados similares.

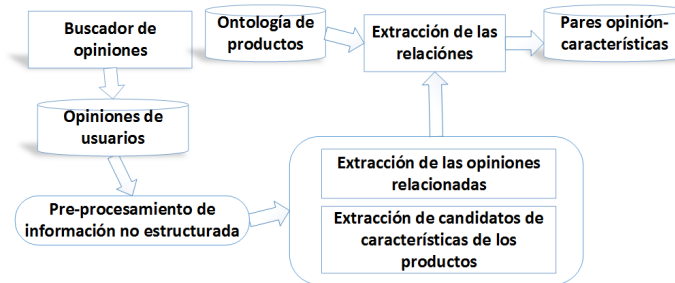


Fig. 1. Ejemplo de arquitectura para la extracci3n de aspectos en un sistema para el análisis de sentimientos [52].

Un reto muy importante para la Web Semántica es relacionar la informaci3n no estructurada o que no ha sido creada con relaciones semánticas en su contenido. Las técnicas que logran la extracci3n de aspectos permiten a investigadores y desarrolladores tener herramientas para lograr relacionar el contenido. Esto se logra al poder enlazar las entidades con los aspectos o característicās asociadas a ellos y que están presentes en el texto digital. Seleccionar una estrategia correcta para la extracci3n de aspectos es determinante para lograr sistemas que permitan construir recursos para la Web Semántica.

Por otro lado, el concepto de Aprendizaje Profundo agrupa a un conjunto de algoritmos que han tenido mucho éxito al resolver problemas en varios domin-

ios del conocimiento humano, entre ellos, el procesamiento de imágenes [50] y el procesamiento del lenguaje natural [7]. Estos algoritmos usan redes neuronales y permiten el aprendizaje automático de las características del problema. Definen varias capas de abstracción y logran que los sistemas aprendan las más complejas representaciones del conocimiento. Estos métodos adquieren la habilidad de aprender las características que son importantes y usan grandes volúmenes de información para el aprendizaje [13]. Las diferentes estrategias o algoritmos del Aprendizaje Profundo han sido aplicadas con éxito en la tarea de extracción de aspectos [17,29,36]. Por esta razón, en esta investigación se hace un análisis de los trabajos que emplean estos algoritmos para la extracción de aspectos con el objetivo de determinar las estrategias de mejor desempeño, así como las formas de representación del conocimiento efectivas.

2 Metodología empleada en la investigación

Esta investigación se basa en los criterios expuestos en [62] y se orientó a los artículos relacionados a la extracción de aspectos empleando algoritmos del Aprendizaje Profundo. Los artículos analizados comprenden un período de tiempo entre enero 2011 hasta febrero 2018. Si los estudios han sido publicados en más de una fuente o memorias de conferencias, se eligió el trabajo más completo. Se tuvieron en cuenta las publicaciones de varias fuentes de búsqueda de investigación científica: *ACM Digital Library*, *IEEE Explorer*, *ScienceDirect*, *Scopus*, *Springer Link* y *Google Scholar*. No se tomaron en cuenta los trabajos que no tienen referencia de la revista, conferencia o memoria de evento, ni aquellos que son un resumen o publicación parcial de otro artículo. Los términos principales de búsqueda fueron: extracción de aspectos aprendizaje profundo (*aspect extraction deep learning*), nivel de aspecto aprendizaje profundo (*aspect-level deep learning*), y nivel de características aprendizaje profundo (*feature level deep learning*). En el trabajo se analizaron las formas de representación de la información, los algoritmos o métodos más empleados y los resultados obtenidos. El análisis comprende 32 artículos de investigación y 26 artículos de revisión.

3 Análisis de los resultados y discusión

Para contribuir a la investigación es útil realizar una evaluación, identificación e interpretación de las investigaciones más relevantes sobre el tema que se aborda hasta la fecha. Una búsqueda sobre revisiones de la literatura o estados del arte reveló la existencia de varios trabajos de este tipo.

En [58] se hace un análisis de varias estrategias del Aprendizaje Profundo para el análisis de sentimiento y sólo se hace alusión a tres artículos que abordan el objeto de investigación de este trabajo. En [31] el objetivo fue estudiar los algoritmos que realizan la extracción de aspectos pero sólo se menciona una propuesta que emplea algoritmos del Aprendizaje Profundo. En [35] se hace un estudio del estado del arte de métodos que han dado una propuesta para la tarea de extracción de aspectos pero no se menciona ninguna propuesta que emplee

métodos del Aprendizaje Profundo. La revisión realizada en [66] menciona solo dos propuestas que emplean algoritmos del Aprendizaje Profundo. En el trabajo publicado en [3] se presenta una revisión sistemática de la literatura orientada a las propuestas existentes para el análisis de sentimiento y se reporta la existencia de tres propuestas que realizan la extracción de aspectos empleando algoritmos del Aprendizaje Profundo. Ramya y un colectivo de autores [43] analizan varias propuestas que realizan la extracción de aspectos o características y solamente referencian una investigación sobre el tema de esta investigación. El estudio presentado en [44] es un estado del arte de las investigaciones que enfrentan la tarea de la extracción de aspectos y no se menciona ninguna propuesta que use algoritmos del Aprendizaje Profundo. En [21] solo se menciona la investigación publicada en [41] y se afirma que ésta es la primera que realizó la extracción de aspectos con algoritmos del Aprendizaje Profundo. El resto de los estados del arte analizados, por solo citar los más recientes [2,4,15,25,32,46,48] no mencionan investigaciones sobre la extracción de aspectos o el empleo de técnicas del Aprendizaje Profundo para esta tarea.

Los 26 estados del arte o revisiones sistemáticas de la literatura que se consultaron en esta investigación tienen una escasa referencia a investigaciones que enfrentan la tarea de la extracción de aspectos y usan algoritmos del Aprendizaje Profundo. Durante el período de enero de 2011 a febrero de 2018 se encontraron 36 artículos que muestran resultados al enfrentar la tarea de extracción de aspectos empleando diferentes algoritmos del Aprendizaje Profundo. De la información obtenida se realiza un análisis orientado a estudiar, por su importancia, los algoritmos del Aprendizaje Profundo y sus resultados en la extracción de aspectos, los modelos para la representación de la información y otros datos de interés para los investigadores. Por otro lado, las ontologías son un valioso recurso para la Web Semántica. Por tanto, esta investigación contribuye a tener una visión general del empleo de técnicas que pueden mejorar el uso de ontologías en el PLN y su enriquecimiento a partir de la extracción de aspectos.

3.1 Algoritmos del Aprendizaje Profundo empleados en la extracción de aspectos

Los algoritmos del Aprendizaje Profundo se pueden usar en la solución de problemas a través de enfoques supervisados, no supervisados o pueden usar propuestas híbridas combinando la salida de un método no supervisado con otro supervisado [13]. Algunos ejemplos de algoritmos del Aprendizaje Profundo propuestos en [13] y [28] son: las Redes Neuronales Convolucionales (*Convolutional Neural Networks*; CNN) [27], las Redes Neuronales Recurrentes (*Recurrent Neural Network*; RNN) [61], la propuesta nombrada Memoria a Corto Plazo (*Long Short Term Memory*; LSTM) [22], *Autoencoders*[11], Redes de Creencia Profunda (*Deep Belief Network*; DBN), Máquinas de Boltzman Restringidas (*Restricted Boltzmann Machines*; RBM) [51], Máquinas Profundas de Boltzman (*Deep Boltzman Machine*; DBM) [47].

En el análisis realizado en esta investigación se encontraron más trabajos con modelos supervisados que no supervisados o híbridos. De las propuestas que

siguen un enfoque supervisado, 12 corresponden a variantes que usan CNN. En [18,41,19] se emplea una variante de CNN mediante el uso de una secuencia de redes convolucionales donde la salida de una red es la entrada de la otra. Este algoritmo es conocido como Redes Convolucionales Apiladas o en Cascada (*Convolutional Stacked Network*; CSN). La selección de este tipo de algoritmo por parte de los investigadores se justifica por la variedad de problemas del PLN resuelto aplicando las CNN y los buenos resultados obtenidos. Esta técnica permite a partir de la representación de las palabras aplicar una operación de convolución o de selección de características importantes en cada capa de la red.

Otros enfoques del Aprendizaje Profundo muy empleados de forma supervisada son las LSTM y las RNN. De los trabajos analizados, 13 usan algoritmos de uno de estos dos tipos. El empleo de estos algoritmos por parte de los autores se justifica porque estas técnicas están especializadas para procesar secuencias de valores. Estos métodos procesan una oración desde el inicio hasta el final, analizando una palabra a la vez. Además, estos métodos se auxilian de las relaciones de dependencias y los árboles sintácticos para extraer a nivel de palabras las relaciones semánticas y sintácticas. De esta forma, logran capturar las representaciones del conocimiento más abstractas y de más alto nivel en diferentes capas. Por otro lado, las RNN son capaces de modelar secuencias de tamaño arbitrario por la aplicación de unidades recurrentes a lo largo de las secuencias de tokens. Las RNN tienen como desventajas el desvanecimiento o explosión del gradiente. Esto provoca que las RNN no sean suficientes para modelar dependencias de gran tamaño. Este problema ha motivado que varias propuestas usen las LSTM para la extracción de aspectos [57,67].

El resto de los trabajos revisados emplean los algoritmos del Aprendizaje Profundo de forma no supervisada, como se muestra, por ejemplo, el uso de Autoencoders en la investigación publicada en [17]. Esta técnica implementa una red neuronal que copia los datos de la capa de entrada en la capa de salida. Internamente, tiene una capa oculta que a través de dos funciones f y r se encargan de codificar y decodificar los datos de entrada. Por lo general, estas funciones se definen de forma que la copia sea aproximada y de esta manera el modelo es forzado a priorizar aquellos aspectos que sean propiedades útiles de los datos.

En [60] se prueba un RBM para la extracción de aspectos. Esta red neuronal es un modelo basado en energía con una distribución de probabilidad conjunta especificada por una función de energía. El RBM representa un grafo no dirigido probabilístico que contiene una capa de variables observables y una capa de variables latentes. Su representación modela un grafo bipartito entre la capa de entrada y de variables latentes donde el cálculo de algunas variables puede ser exhaustivo y costoso. En [60] las unidades de la capa oculta representan aspectos, sentimientos previamente seleccionados y palabras de rechazo, mientras que la capa de entrada está asociada a las palabras de las oraciones de entrenamiento.

En [20,63] se usa una Red de Memoria Profunda (*Deep Memory Network*; DMN) que es entrenada a partir de un conjunto de aspectos predefinidos. En varios trabajos [12,20,23,63] se reporta el empleo de un mecanismo conocido

como Mecanismo de Atención (*Attention Mechanism*) que promedia los pesos que pueden ser relevantes en otros puntos de una red neuronal. Este mecanismo permite incluir características lingüísticas o sintácticas al proceso de aprendizaje de la red neuronal que implementa el algoritmo del Aprendizaje Profundo. Varias propuestas [24,41,65] agregan reglas lingüísticas al empleo de algoritmos del Aprendizaje Profundo. Debido a que la extracción de aspectos es una tarea de gran importancia en el análisis de sentimientos basado en aspectos, aparecen muchas propuestas que realizan la extracción de aspectos y la clasificación de la polaridad de los aspectos (positiva, negativa y neutra) de forma paralela [37,58,64], lo que dificulta determinar cuáles técnicas fueron propiamente utilizadas para la extracción.

De todos los trabajos analizados, el propuesto en [41], empleando un CNN con siete capas consecutivas o apiladas, obtuvo los mejores resultados alcanzando un 86 % para la Medida-F al evaluar su propuesta en un conjunto de datos de opiniones de restaurantes [40]. Es importante señalar que la calidad de la extracción de los aspectos no solo depende de las técnicas de Aprendizaje Profundo que se apliquen, en los resultados también repercute la forma de representación textual. Por ello, que en la próxima sección abordaremos sobre las principales formas de representación de la información y su repercusión en la extracción de aspectos.

3.2 Formas de representación de la información

La forma de representación de la información está asociada a la posible organización de la información del texto no estructurado. Los textos generalmente se conforman por párrafos, oraciones y palabras. Una correcta organización de la información no estructurada es necesaria para lograr un efectivo entrenamiento de los algoritmos del Aprendizaje Profundo y consecuentemente una eficaz extracción de aspectos.

El Aprendizaje Profundo se basa esencialmente en el trabajo con redes neuronales. Uno de los retos más importantes al utilizar redes neuronales es lograr una forma de representación correcta para los datos de entrada a la red. Los conjuntos de entrenamiento de estas redes están formados por documentos u oraciones que contienen palabras. El *Word Embedding*[33] se definió para lograr un mejor entrenamiento de las redes neuronales que se utilizan en el procesamiento del lenguaje natural. *Word Embedding* es el nombre de un conjunto de lenguajes de modelado y técnicas de aprendizaje dónde las palabras o frases del vocabulario son vinculadas a vectores de números reales. Conceptualmente transforma un espacio con una dimensión por cada palabra a un espacio vectorial continuo con menos dimensiones [27]. La herramienta *word2vec*² fue creada para la creación de estos vectores a partir de un conjunto de palabras.

De las propuestas analizadas, 31 emplean como forma de representación el *Word Embedding*. Este necesita de grandes volúmenes de información para la creación de los vectores asociados a las palabras. En [33] se propuso e hizo

² <https://code.google.com/archive/p/word2vec/>

público un conjunto de vectores pre-entrenados, a partir de un conjunto de datos de noticias procedentes del sitio Google News ³ (100 mil millones de palabras). El modelo contiene un vector de dimensión 300 para 3 millones de palabras o frases. Este conjunto de vectores fue empleado por 13 de los 31 trabajos que emplean *Word Embedding*. El resto usa la herramienta word2vec para construir los vectores asociados al conjunto de palabras. En varios trabajos [12,59,19] se usan conjuntos de datos, como el propuesto en Yelp ⁴ o Amazon ⁵, para el entrenamiento del vector asociado a las palabras.

El concepto de vector de valores reales asociados a palabras puede ser extendido a oraciones o párrafos, a partir del conjunto de datos de entrenamiento. En [20] se realiza el entrenamiento del vector de valores reales en función de las oraciones presentes en el conjunto de datos y en el trabajo publicado en [39] se emplean los párrafos. Estas formas de representación de la información repercutieron negativamente en la calidad de la extracción de aspectos.

En [17] se emplea una bolsa de palabras (*bag-of-words*) y se obtiene un vector binario que codifica la presencia/ausencia de unigramas y bigramas. En [60] se tiene un vector de palabras y se calcula la frecuencia de los términos (*Term Frequency*; TF) para los sustantivos en el conjunto de datos de entrenamiento y se calcula la frecuencia inversa del documento (*Inverse Document Frequency*; IDF) en un conjunto de datos de n-gramas de Google ⁶. Las propuestas que utilizan estas formas de representación no necesitan grandes conjuntos de datos para su entrenamiento y no tiene una representación vectorial de grandes dimensiones. Sin embargo, pierden la riqueza semántica que posee el *Word Embedding* y, por tanto, no superan los resultados de las propuestas que lo emplean.

4 Conclusiones y recomendaciones

Las ontologías son usadas en el proceso de extracción de la información no estructurada. En este proceso es necesario realizar varias tareas del PLN, especialmente la subtarea de extracción de aspectos. Esta subtarea se usa en combinación con ontologías para obtener mayor información de la fuente de datos no estructurados. Se han desarrollado varias propuestas que intentan extraer aspectos, dentro de éstas, las técnicas del Aprendizaje Profundo han ganado gran popularidad por los buenos resultados obtenidos. Sin embargo, los 26 estados del arte o revisiones sistemáticas de la literatura que se consultaron en esta investigación tienen una escasa referencia a investigaciones que enfrentan la tarea de la extracción de aspectos y usan algoritmos del Aprendizaje Profundo. Es por ello que en esta investigación se realizó un estudio de las principales publicaciones sobre el tema y se logró suplir así las carencias de los artículos de revisión consultados. El análisis de los 36 artículos sobre el uso del Aprendizaje Profundo

³ <https://news.google.com>

⁴ <https://www.yelp.com/datasetchallenge/>

⁵ <http://www.cs.jhu.edu/~mdredze/datasets/sentiment/>

⁶ <http://books.google.com/ngrams/datasets>

para la extracción de aspectos publicados desde enero de 2011 hasta febrero de 2018 arrojó las siguientes conclusiones:

- CNN es la técnica del Aprendizaje Profundo más empleada en la extracción de aspectos. La selección de este tipo de algoritmo por parte de los investigadores se justifica por la variedad de problemas del PLN resueltos aplicando las CNN y los buenos resultados obtenidos al usar este algoritmo.
- *Word Embedding* con vectores asociados a palabras es la forma de representación más empleada para la extracción de aspectos aplicando Aprendizaje Profundo. La forma más empleada de entrenar estos vectores parte de un conjunto de datos de Google Noticias. No obstante, algunos investigadores emplean exitosamente otros conjuntos de datos para el entrenamiento inicial del *Word Embedding*. Esta forma de representación es muy útil para datos de entradas de las redes neuronales de los algoritmos del Aprendizaje Profundo, debido a que representan un vector de números reales. Los grandes conjuntos de datos con los que se crean los vectores permiten cubrir gran cantidad de ejemplos y representar la relación semántica entre palabras, información útil en el proceso de extracción de aspectos.

La extracción automática de aspectos en datos no estructurados utilizando las CNN con la representación Word Embedding es una solución prometedora para un uso efectivo posterior de las ontologías. El trabajo futuro debe estar orientado a la revisión de los resultados de la extracción de aspectos en múltiples dominios y su impacto en los resultados al emplear ontologías.

References

1. Agarwal, B., Mittal, N., Bansal, P., Garg, S.: Sentiment analysis using common-sense and context informations. *Computational Intelligence and Neuroscience* 30 (2015)
2. Ain, Q.T., Ali, M., Riaz, A., Noureen, A., Kamran, M., Hayat, B., Rehman, A.: Sentiment analysis using deep learning techniques: a review. *International Journal of Advanced Computer Science and Applications* (6), 424–433 (2017)
3. Al-Moslmi, T., Omar, N., Abdullah, S., Albared, M.: Approaches to cross-domain sentiment analysis: a systematic literature review. *IEEE Access* pp. 16173–16192 (2017)
4. Alasmari, S.F., Dahab, M.: Sentiment detection, recognition and aspect identification. *International Journal of Computer Applications* pp. 31–37 (2017)
5. Alfonso, M., Sardinha, R.: Ontology based aspect level opinion mining. *International Journal of Engineering Sciences and Research* 5, 797–804 (2016)
6. Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R., Ives, Z.: Dbpedia: a nucleus for a web of open data. *The Semantic Web. Lecture Notes in Computer Science* 4825, 722–735 (2007)
7. Azzam, A., Tazi, N., Hossny, A.: Text-based question routing for question answering communities via deep learning. In: *Symposium on Applied Computing, Marrakech, Marruecos*. pp. 1674–1678 (2017)

8. Baccianella, S., Esuli, A., Sebastiani, F.: SentiWordNet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In: Language Resources and Evaluation Conference (LREC), Valletta, Malta. pp. 2200–2204 (2010)
9. Blei, D.M., Ng, A., Jordan, M.I.: Latent dirichlet allocation. *Journal of machine Learning research* (1), 993–1022 (2003)
10. Bollacker, K., Evans, C., Paritosh, P., Sturge, T., Taylor, J.: Freebase: a collaboratively created graph database for structuring human knowledge. In: ACM SIGMOD international conference on Management of data, Vancouver, Canada. pp. 1247–1250 (2008)
11. Bourlard, H., Kamp, Y.: Auto-association by multilayer perceptrons and singular value decomposition. *Biological Cybernetics* pp. 291–294 (1988)
12. Cheng, J., Zhao, S., Zhang, J., King, I., Zhang, X., Wang, H.: Aspect-level sentiment classification with heat (hierarchical attention) network. In: ACM on Conference on Information and Knowledge Management, Singapore, Singapore. pp. 97–106 (2017)
13. Deng, L., Yu, D.: Deep learning in medical image analysis. *Foundations and Trends in Signal Processing* (3-4), 197–387 (2014)
14. Finin, T., Syed, Z.: Creating and exploiting a web of semantic data. In: International Conference on Agents and Artificial Intelligence, Valencia, España. pp. 7–18 (2010)
15. Ganeshbhai, S.Y., Shah, B.K.: Feature based opinion mining: a survey. In: IEEE International Advance Computing Conference (IACC), Bangalore, India. pp. 919–923 (2015)
16. Ganino, G., Lembo, D., Scafoglieri, F.: Ontology Population from Raw Text Corpus for Open-Source Intelligence. *Current Trends in Web Engineering* pp. 173–186 (2018)
17. Glorot, X., Bordes, A., Bengio, Y.: Domain adaptation for large-scale sentiment classification: a deep learning approach. In: 28th international conference on machine learning (ICML-11), Bellevue, Washington, USA. pp. 513–520 (2011)
18. Gu, X., Gu, Y., Wu, H.: Cascaded convolutional neural networks for aspect-based opinion summary. *Neural Processing Letters* pp. 581–594 (2017)
19. H. Ye, Z. Yan, Z.L.y.W.C.: Dependency-tree based convolutional neural networks for aspect term extraction. In: Pacific-Asia Conference on Knowledge Discovery and Data Mining, Jeju, South Korea. pp. 350–362 (2017)
20. He, R., Lee, W.S., Ng, H.T., Dahlmeier, D.: An unsupervised neural attention model for aspect extraction. In: 55th Annual Meeting of the Association for Computational Linguistics, Vancouver, Canada. pp. 388–397 (2017)
21. Hemmatian, F., Sohrabi, M.K.: A survey on classification techniques for opinion mining and sentiment analysis. *Artificial Intelligence Review* pp. 1–51 (2017)
22. Hochreiter, S., Schmidhuber, J.: Long short-term memory. *Neural Computation* (8), 1735–1780 (1997)
23. Huang, M., Wang, Y., Zhu, X., Zhao, L.: Attention-based lstm for aspect-level sentiment classification. In: Empirical Methods in Natural Language Processing, Austin, Texas, USA. pp. 606–615 (2016)
24. Joty, S., Liu, P., Meng, H.M.: Fine-grained opinion mining with recurrent neural networks and word embeddings. In: Conference on Empirical Methods in Natural Language Processing, Lisboa, Portugal. pp. 1433–1443 (2015)
25. Kumar, P.K., Nandagopalan, S.: Insights to problems, research trend and progress in techniques of sentiment analysis. *International Journal of Electrical and Computer Engineering (IJECE)* 7(5), 2818–2822 (2017)

26. Lau, R.Y., Li, C., Liao, S.S.: Social analytics: learning fuzzy product ontologies for aspect-oriented sentiment analysis. *Decision Support Systems* 65, 80–94 (2014)
27. LeCun, Y.: Generalization and network design strategies. *Connectionism in perspective* pp. 143–155 (1989)
28. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. *Nature* (7553), 436–444 (2015)
29. Liao, J., Wang, S., Li, D., Li, X.: FREERL: Fusion relation embedded representation learning framework for aspect extraction. *Knowledge-Based Systems* pp. 9–17 (2017)
30. Liu, B.: *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge University Press (2015)
31. Luo, C., Sun, S., Chen, J.: A review of natural language processing techniques for opinion mining systems. *Information Fusion* pp. 10–25 (2015)
32. Madhoushi, Z., Hamdan, A.R., Zainudin, S.: Sentiment analysis techniques in recent works. In: *Science and Information Conference (SAI), 2015, London, United Kingdom*. pp. 288–291. IEEE (2015)
33. Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S., Dean, J.: Distributed representations of words and phrases and their compositionality. *Advances in Neural Information Processing Systems* p. 3111–3119 (2013)
34. Miller, G.A.: WordNet: a lexical database for english. *Communications of the ACM* 38(11), 39–41 (1995)
35. More, P., Ghotkar, A.: A study of different approaches to aspect-based opinion mining. *International Journal of Computer Applications* (6)
36. Nguyen, T.H., Shirai, K.: PhraseRNN: phrase recursive neural network for aspect-based sentiment analysis. In: *Conference on Empirical Methods in Natural Language Processing, Lisboa, Portugal*. pp. 2509–2514 (2015)
37. Nguyen-Hoang, B.D., Ha, Q.V., Nghiem, M.Q.: Aspect-based sentiment analysis using word embedding restricted boltzmann machines. In: *International Conference on Computational Social Networks*. pp. 285–297 (2016)
38. Peñalver-Martínez, I., Garcia-Sanchez, F., Valencia-Garcia, R., Rodríguez-García, Miguel Á., M.V., Fraga, A., Sánchez-Cervantes, J.L.: Feature-based opinion mining through ontologies. *Expert Systems with Applications* 41(13), 5995–6008 (2014)
39. Pham, D.H., Le, A.C.: Learning multiple layers of knowledge representation for aspect based sentiment analysis. *Data and Knowledge Engineering* (2017)
40. Pontiki, M., Galanis, D., Pavlopoulos, J., Papageorgiou, H., Androutsopoulos, I., Manandhar, S.: Semeval-2014 task 4: Aspect based sentiment analysis. In: *8th international workshop on semantic evaluation (SemEval 2014)*. p. 27–35 (2014)
41. Poria, S., Cambria, E., Gelbukh, A.: Aspect extraction for opinion mining with a deep convolutional neural network. *Knowledge-Based Systems* pp. 42–49 (2016)
42. Poveda, M.: *Metodología NeOn Aplicada a la representación del contexto*. Maestría de investigación en inteligencia artificial, Universidad Politécnica de Madrid, Madrid, España (2010)
43. Ramya, R.S., Venugopal, K.R., Iyengar, S.S., Patnaik, L.M.: Feature extraction and duplicate detection for text Mining: a survey. *Global Journal of Computer Science and Technology*. (5) (2017)
44. Rana, T.A., Cheah, Y.N.: Aspect extraction in sentiment analysis: comparative analysis and survey. *Artificial Intelligence Review* (4) (2016)
45. Ristoski, P., Paulheim, H.: Semantic web in data mining and knowledge discovery: a comprehensive survey. *Web semantics: science, services and agents on the World Wide Web* 36(11), 1–22 (2016)
46. Rojas-Barahona, L.M.: Deep learning for sentiment analysis. *Language and Linguistics Compass* (12), 701–719 (2016)

47. Salakhutdinov, R., Hinton, G.: Deep boltzmann machines. *Artificial Intelligence and Statistics* pp. 448–455 (2009)
48. Schouten, K., Frasinca, F.: Survey on aspect-level sentiment analysis. *IEEE Transactions on Knowledge and Data Engineering* (3), 813–830 (2016)
49. Shein, K.P.P.: Ontology based combined approach for sentiment classification. In: 3rd International Conference on Communications and information technology, Florencia, Italia. pp. 112–115 (2009)
50. Shen, D., Wu, G., Suk, H.I.: Deep learning in medical image analysis. *Annual Review of Biomedical Engineering* pp. 221–248 (2017)
51. Smolensky, P.: Information processing in dynamical systems: Foundations of harmony theory. Ph.D. thesis, Colorado University at Boulder, Department of Computer Science (1986)
52. Somprasertsri, G., Lalitrojwong, P.: Mining feature-opinion in online customer reviews for opinion summarization. *Journal of Universal Computer Science* 16(6), 938–955 (2010)
53. Speer, R., Havasi, C.: ConceptNet 5: A large semantic network for relational knowledge. In: *The People’s Web Meets NLP*. pp. 161–176. Springer (2013)
54. Strapparava, C., Valitutti, A.: WordNet Affect: an affective extension of wordnet. In: 4th International Conference on Language Resources. p. 1083–1086 (2004)
55. Suárez-Figueroa, M.C., Gómez-Pérez, A., Fernández-López, M.: The NeOn methodology for ontology engineering. *Ontology engineering in a networked world* pp. 9–34 (2012)
56. Suchanek, F.M., Kasneci, G., Weikum, G.: Yago: a core of semantic knowledge. In: 16th International Conference on World Wide Web, Banff, Alberta, Canada. pp. 697–706 (2007)
57. Sun, C., Wang, X., Liu, Y., Wang, B., Wang, X.: Predicting polarities of tweets by composing word embeddings with long short-term memory. In: *Association for Computational Linguistics (ACL)*, Beijing, China. pp. 1343–1353 (2015)
58. Tang, D., Qin, B., Liu, T.: Deep learning for sentiment analysis: successful approaches and future challenges. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* (6), 292–303 (2015)
59. Toh, Z., Jian, S.: NLANGP: Supervised Machine Learning System for Aspect Category Classification and Opinion Target Extraction. In: 9th International Workshop on Semantic Evaluation, Denver, Colorado, USA. pp. 496–501 (2015)
60. Wang, C., He, X., Zhou, A.: A Short Survey on Taxonomy Learning from Text Corpora: Issues, Resources and Recent Advances. In: 2017 Conference on Empirical Methods in Natural Language Processing. pp. 1190–1203. ACM (2017)
61. Williams, D., Hinton, G.: Learning representations by back-propagating errors. *Nature* (6088), 533–538 (1986)
62. Wohlin, C., Prikladniki, R.: Systematic literature reviews in software engineering. *Information and Software Technology* (6), 919–920 (2013)
63. Xiong, S., Zhang, Y., Ji, D., Lou, Y.: Distance metric learning for aspect phrase grouping. In: *Conference on Computational Linguistics (Coling)*, Osaka, Japon. p. 2492–2502 (2016)
64. Xu, L., Lin, J., Wang, L., Yin, C., Wang, J.: Deep convolutional neural network based approach for aspect-based sentiment analysis. *Artificial Intelligence and Statistics* pp. 199–204 (2017)
65. Ying, D., Yu, J., Jiang, J.: Recurrent neural networks with auxiliary labels for cross-domain opinion target extraction. In: 31st AAAI Conference on Artificial Intelligence: San Francisco, USA. pp. 3436–3442 (2017)

66. Yousif, A., Niu, Z., Tarus, J.K., Ahmad, A.: A survey on sentiment analysis of scientific citations. *Artificial Intelligence Review* pp. 1–34 (2017)
67. Yuan, J., Zhao, Y., Qin, B., Liu, T.: Local contexts are effective for neural aspect extraction. In: *IEEE International Advance Computing Conference (IACC)*, Bangalore, India. pp. 244–255 (2017)