

Un enfoque de Recomendación multi-dominio basado en Embeddings y redes semánticas

Ignacio Gatti, Silvia Schiaffino, and Andrés Díaz Pace

ISISTAN (CONICET-UNCPBA), Campus Universitario, Tandil
{ignacio.gatti,silvia.schiaffino,andres.diazpace}@isistan.unicen.edu.com

Abstract. Actualmente, la generación de información en la web crece exponencialmente. Los Sistemas de Recomendación se valen de ésta para ofrecer a los usuarios potenciales ítems que pueden ser valiosos para ellos. Sin embargo, la mayoría de estos sistemas se mueven dentro de un sólo dominio (por ej., películas), desaprovechando información de otros contextos que pueden ser valiosa para la construcción del perfil del usuario, especialmente en escenarios de *cold-start*. El objetivo de este trabajo es aprovechar la información de un dominio puntual (películas) para generar recomendaciones más precisas en uno nuevo (libros). Básicamente, se propone un enfoque híbrido que, por un lado, utiliza un grafo semántico que relaciona artistas; y, por el otro, utiliza la técnica de *word embeddings* para encontrar relaciones de similitud entre las sinopsis de ambos ítems. Los resultados de una evaluación inicial muestra que el enfoque constituye una alternativa interesante para problema del *cold-start*.

Keywords: Recomendación multidominio, *word embedding*, redes semánticas

1 Introducción

En esta época donde el concepto de producción se da en su máxima expresión, desde la generación de datos e información hasta el desarrollo de nuevos productos; donde los usuarios son asfixiados por el peso de ésta, los Sistemas de Recomendación (SR) proponen una alternativa para que ellos puedan “filtrar” de forma automatizada de qué cosas quieren ser informadas. Básicamente, un SR es una pieza de software cuyo propósito es sugerirle al usuario potenciales ítems que consumirá en el futuro [14].

En sintonía con lo anterior, la popularidad de las redes sociales y los comercios electrónicos han permitido la recolección de infinidad de datos sobre el comportamiento de los usuarios a la hora de recorrer dichos lugares, dejando rastros de aquellos productos que les llamó la atención, de aquellos que compraron y los que no; de las películas que les gustaron. Estas huellas sirven de insumo para la construcción de un perfil personalizado. Asimismo, estas redes han permitido el desarrollo de diversas comunidades en la Web donde se deposita información sobre distintos tópicos, desde científicos a entretenimiento. Así, de manera análoga a los perfiles de usuario, es posible generar una descripción compleja de los ítems a recomendar.

Sin embargo, la construcción de dichos perfiles, presenta un desafío en el área de los SR. En la mayoría de los SR desarrollados, se observa que los perfiles generados se restringen al dominio en el que el SR se mueve. Básicamente, al trabajar sobre un único dominio, se presupone que el interés de un usuario sobre algunos de sus elementos está formado únicamente en aspectos propios de éste. Esta suposición coloca varias restricciones para hacer frente a dos problemas clásicos en SR: el arranque en frío (conocido como *cold-start*) y la escasez de ratings. El primero, es un caso particular del segundo, que se da cuando se agrega un nuevo usuario o ítem al sistema, para el cual no se posee ratings que permitan generar recomendaciones personalizadas. En el segundo, los ratings recolectados son tan pocos que repercute negativamente en la calidad de las recomendaciones. En el seno de estas problemáticas, se colocan los sistemas de recomendación multi-dominio. Básicamente, el objetivo que se proponen es poder extraer conocimiento de un dominio origen (D_s) para poder aplicarlo en un dominio objetivo (D_t), de manera tal de poder mejorar las recomendaciones en este último [2].

El presente trabajo propone un enfoque novedoso para atacar la problemática presentada, materializada en los dominios de películas y libros (Fig. 1). En líneas generales, la cuestión radicó en la definición de un espacio de representación donde los ítems de ambos dominios resultasen comparables. Para ellos, se extrajo la información textual que estuviera avalada por la comunidad de ambos dominios (Wikipedia¹, IMDB², GoodReads³). En la literatura, la forma clásica para la representación de documentos es el enfoque TF-IDF, donde cada ítem es mapeado a un vector de palabras de acuerdo a la frecuencia de aparición de cada una de ellas en éste [5]. Sin embargo, aquí se utilizó el concepto de *word embedding* para su representación. Coloquialmente, éstos son "descripciones" de palabras en forma de características o categorías compartidas a través éstas [9]. Finalmente, para vincular a los directores con los escritores se definió una red de influencia que permite calcular la distancia que existe entre cada par. Aprovechando la información semántica que ofrece Wikipedia sobre los artistas, es posible construir una red a partir de las etiquetas "*Influenciado por*" y "*Influencia a*". Si bien los resultados son preliminares, se puede avizorar que el enfoque desarrollado es una alternativa para atacar el problema del *cold-start*. El resto del trabajo se organiza de la siguiente manera. En la sección 2 se analizan los trabajos relacionados que sustentan el enfoque. La sección 3 presenta el enfoque desarrollado. En la 4 muestran los resultados, y se discuten pros y contras del enfoque. En la sección 5 se presentan las conclusiones y se esbozan líneas de trabajo futuro.

¹ <https://www.wikipedia.org/>

² <https://www.imdb.com/>

³ <https://www.goodreads.com/>



Fig. 1: Esquema: SR multi-dominio

2 Trabajos Relacionados

A la hora de analizar SR multi-dominio, Cantador[2] propone una clasificación tomando como criterio la forma en que “explotan el conocimiento”, definiendo dos ramas: *agregación de conocimiento* y *transferencia de conocimiento*. El enfoque presentado en este trabajo se encuentra en la segunda rama, con lo cual a continuación se analizarán algunos de los trabajos en el área.

En este tipo de técnicas, el foco está puesto en tomar al D_s como soporte para extraer conocimiento en forma de patrones para aplicarlo en las recomendaciones a realizar sobre el D_t . En esta dirección, el trabajo de [4] propone un framework basado en grafos semánticos (extraído de *DBpedia Ontology*⁴), para tomar conocimiento de diversos dominios y aplicarlos en la recomendación de ítems en D_t . Puntualmente, ellos aprovechan dicho grafo para encontrar relaciones entre un lugar de interés (POI) que visita el usuario y la música más apropiada para escuchar allí. Para esto, recorren la estructura buscando músicos relacionados con el POI. Por otro lado, en [12] se propone un enfoque de *Transferencia por Factorización Integrativa*, donde buscan atacar el problema de la escasez de datos en D_s (puntualmente películas), tomando una matriz auxiliar del mismo dominio pero de otro sistema, que posee una probabilidad de distribución de ratings para algunos ítems. Básicamente, dicha probabilidad define una restricción a la hora de predecir un ítem en D_t . Según ellos, lo interesante de este enfoque es que dicha distribución de probabilidad puede obtenerse de manera implícita, sin tener que pedirle a los usuarios que voten determinados ítems. Otro trabajo interesante en esta rama es [6], donde la transferencia de conocimiento se realiza a través de la definición de un codebook. Éste se define a partir de patrones extraídos en D_s , generalmente aplicando alguna técnica de factorización sobre la matriz del sistema (Usuarios x Ítems), para luego llevar a D_t , donde se reconstruye la matriz del sistema del dominio objetivo aplicando el codebook.

Los *word embeddings* irrumpieron en el mundo del Procesamiento del Lenguaje Natural (NLP), provocando una revolución en tareas como el reconocimiento de

⁴ <https://wiki.dbpedia.org/services-resources/ontology>

entidades o análisis sintáctico. Desde entonces, el área de Recuperación de la Información (RI) lo ha adoptado como una nueva estrategia para la recuperación de documentos, teniendo resultados alentadores. Vale la pena resaltar el trabajo [7], donde propone definir cada documento y la consulta como un promedio de los vectores de *word embeddings* de sus palabras. De esta forma los lleva a un espacio vectorial donde puede aplicar similitud de coseno para encontrar documentos similares. En la misma dirección, se puede ubicar [11], que utiliza una implementación de *word embeddings*, llamada *Word2Vec*, en el modelado de sus ítems. Básicamente, éste se centra en la recomendaciones de puntos de interés en una ciudad para un visitante, representando cada lugar a partir de los “check-ins” que hace el usuario y mapeando estos a un espacio vectorial utilizando los *word embeddings*.

3 Enfoque propuesto

El presente enfoque propone una estrategia para poder extraer conocimiento de D_s -en este caso películas-, para aplicarlo en D_t -libros-. Puntualmente, el objetivo del trabajo es poder establecer un método que utilice todo el conocimiento que se tiene de D_s -matriz de ratings, información de los ítems aportada por la comunidad-, para poder ofrecerle recomendaciones más precisas y apropiadas a los usuarios de D_s en el D_t -donde solo se posee información de los ítems-.

Desde el punto de vista de los SR tradicionales, Ricci[14] marca dos enfoques para atacar el problema de la recomendación: *filtrado colaborativo* y *basado en contenido*. El primero, solo usa la matriz ratings para calcular las recomendaciones, buscando usuarios similares (es decir que hayan votado un mismo subconjunto de ítems de forma análoga) para recomendar aquellos ítems que un usuario aún no ha consumido y que otro con gustos parecido sí experimentó positivamente. Mientras que el segundo, define un espacio de representación al que son mapeados ítems y usuarios con el fin de encontrar aquellos ítems que son similares para un usuario determinado. Aquí se optó por este último, dado que es conocido su potencialidad para tratar el *cold-start*, como lo muestra [8] en su trabajo de análisis comparativo de ambos enfoques.

Como se dijo anteriormente, uno de los problemas centrales que se presentaron fue la representación de los ítems. Porque si bien ambos pueden entrar en las categorías de “entretenimiento” o “arte”, la naturaleza de ellos es distinta: las películas son objetos audiovisuales, mientras que los libros objetos escritos. Por eso, de la diversas fuentes de datos que existen de ambos (por ej., sinopsis, imágenes, guión), sólo se escogieron aquellas que estuviesen en formato texto, es decir, se tomaron las sinopsis, los géneros, los nombres de los directores y escritores. De esta forma, se logró una homogeneización de los datos para luego poder procesarlos y representar a todos los ítems en un mismo espacio.

De esta selección de datos, un subconjunto de ellos poseen una correlación directa -las sinopsis y los géneros-, mientras que entre directores y escritores existen una serie de relaciones implícitas. Por el lado de los primeros, ambos pueden ser visto como documentos que expresan en palabras el contenido de la obra, ya sea

a través de una reseña o a través de un conjunto de palabras claves (como lo son los géneros); con lo cual, sobre ellos se puede aplicar alguna técnica conocida de RI para modelarlos. Por el otro lado, directores y escritores son casos particulares, porque entre ellos se tejen relaciones de tipo social e intelectual; de manera tal que el conocimiento existente allí se desprende de esa red, siendo necesario algún análisis de grafo.

Por eso, en los siguientes apartados se detallan las dos estrategias definidas para mapear ambos dominios a un mismo espacio: red de influencia y *embedding*.

3.1 Red de influencia

Hoy en día, Wikipedia es una de las mayores fuentes de conocimiento existente en Internet, con una comunidad de más de 307.000 usuarios que participan activamente. Más allá de la revolución que ésta significa por su edición de manera colaborativa, lo que la hace aún más potente es la ontología (*DBpedia Ontology*) que se ha ido desarrollando tras ella. Ésta ha sido generada manualmente (siguiendo el estándar RDF⁵), a partir de los infobox de cada entrada de Wikipedia, cubriendo actualmente 685 clases relacionadas de forma jerárquica, con 2.795 propiedades. Asimismo, ésta puede ser consultada programáticamente, utilizando el lenguaje SPARQL⁶, ofreciendo la posibilidad de hacer diversos recorridos sobre el grafo subyacente.

En este trabajo, se observó que una de las propiedades que poseían los artistas registrados en Wikipedia eran las etiquetas: “*Influido por*” o “*Influye a*”. Con lo cual, resultaba factible poder extraer de dicha ontología un “grafo de influencias” (Fig. 2), donde se materializaba las relaciones intelectuales entre directores-directores, escritores-escritores y directores-escritores. Cuanto más cerca estuvieran los artistas en el grafo, mayor sería su afinidad en sus obras, pues “todo artista siempre es una amalgama entre la tradición que mamó y su pulsión por romper con ella” [10]. A partir de esta premisa, se definió la métrica de influencia 1 para poder ir del espacio de directores al de escritores.

$$Influencia(director, escritor) = 1/camino_mas_corto(director, escritor) \quad (1)$$

De esta forma, es posible determinar cuánta influencia de cada escritor tiene una película a partir de su director (lo mismo vale para el caso director y libro), permitiendo representar una película en el espacio de los libros y viceversa. Así, es factible establecer un espacio de representación común donde películas y libros se vuelven comparables.

3.2 Word Embedding

Con la irrupción de las redes neuronales, en NLP ha surgido un cambio de visión en la forma de representación del lenguaje, inclinándose por modelos vectoriales

⁵ <https://www.w3.org/TR/rdf-schema/>

⁶ <http://dbpedia.org/sparql>

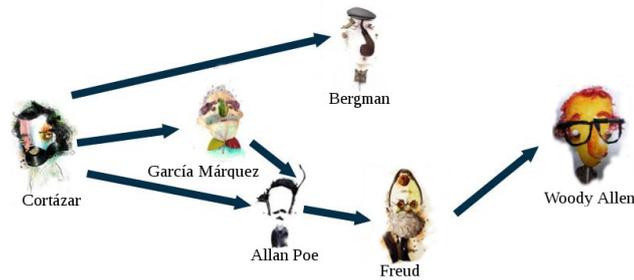


Fig. 2: Red de influencias

que permiten almacenar relaciones internas entre las palabras. Básicamente, estas redes son entrenadas con grandes datasets, obteniendo así modelos que aprenden, por ejemplo, relaciones gramaticales o probabilidad de coocurrencia entre palabras. Entre ellos, *word embeddings* es uno de los modelos que posee mayor popularidad por su buen desempeño en tareas como análisis sintáctico[15], análisis de sentimiento[16] o reconocimiento de entidades[3] .

En palabras de Mikolov[9], los *word embeddings* son vectores n-dimensionales en un espacio vectorial continuo donde cada dimensión representa conceptos o categorías compartidas a través de palabras. Pero se trata de una "categorización interna", que no se traduce a aquellas definidas por el intelecto humano. Por ejemplo, aunque "rey" y "reina" no son sinónimos, hiperónimos ni comparten ninguna relación gramatical, es sabido que están relacionados y, de hecho, sus representaciones están cerca en un espacio generado por los *embeddings* (Fig.3). Si bien los detalles de los tipos de redes neuronales que se utilizan para obtener

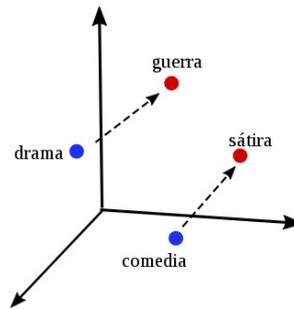


Fig. 3: Word embeddings: espacio vectorial

estos modelos exceden a este trabajo, resulta importante destacar los dos enfoques: *predictivo* y *por conteo*. El primero, parte de la idea que las palabras se ven afectadas inherentemente por el contexto en el que se utilizan. De esta forma, [9] propone generar un modelo que pueda predecir la aparición de una palabra

objetivo en un contexto determinado en función de las características que tiene la *palabra de objetivo*. Si se minimiza el error en este proceso para cada palabra en el vocabulario, se obtendrá efectivamente buenas representaciones para las palabras, porque la predicción de una palabra en este escenario implicará que se hayan modelado correctamente las características de las palabras individuales. Este enfoque también se llama *Skip-Gram* y es el que se utiliza para generar el modelo *Word2Vec*⁷. Por otro lado, [13] propone el otro enfoque, obteniendo el modelo *GLoVe*⁸. En líneas generales, éste consta de dos pasos: primero calcula la probabilidad de co-ocurrencia entre las palabras a lo largo del corpus, determinando que una palabra se relaciona con otra palabra si tienen una alta probabilidad conjunta. Luego, para crear los *embeddings* se realiza un paso de "aprendizaje" donde éstos se refinan para que sean representaciones efectivas de las probabilidades conjuntas, por ejemplo, usando regresión logística. En otras palabras, la idea es crear una matriz que pueda "explicar" la matriz original minimizando la pérdida de reconstrucción, que es la cantidad de información que se pierde cuando se reconstruye la matriz.

En este trabajo se utilizaron dos de los corpus pre-entrenados más utilizados por su robustez: *GLoVe* y *Word2Vec*. Ambos, fueron entrenados con un vocabulario de un par de millones de palabras, generando un vector de 300 dimensiones para cada una de ellas.

A partir de allí, se propuso representar a cada película y libro en el espacio dimensional definido por los *word-embeddings*. Para ello, siguiendo la estrategia definida en [7], se obtuvo el vector promedio de la suma de vectores de *embeddings* de cada palabra de la sinopsis y los géneros.

$$mean_embedding(document) = 1/n * \sum_{d_i \in document} word_embedding(d_i) \quad (2)$$

3.3 Recomendación multi-dominio

Para realizar la recomendación de ítems en D_t -en este caso libros-, se siguió el enfoque clásico basado en contenido definido en [1]. Básicamente, una vez mapeados los ítems de ambos dominios a un mismo espacio vectorial, sólo queda llevar al usuario a dicho espacio, de forma tal de buscar los libros que sean más similares a sus preferencias.

Para ello, se definió el vector de un usuario como el promedio de la suma de los vectores que éste consumió en D_s , ponderado por el rating que le dio.

$$vector_u = 1/n * \sum_{i \in R_u} mean_embedding(i) * r_{ui} \quad (3)$$

Ya en el espacio vectorial, resulta sencillo encontrar vectores similares utilizando similitud del coseno.

$$sim_{cos}(vector_u, embedding_i) = \frac{vector_u \cdot embedding_i}{|vector_u| * |embedding_i|} \quad (4)$$

⁷ <https://github.com/mnihaltz/word2vec-GoogleNews-vectors>

⁸ <https://nlp.stanford.edu/projects/glove/>

4 Evaluación

En esta sección se presentan los resultados de la evaluación del enfoque.

4.1 Dataset

La evaluación del enfoque se realizó sobre el recorte de un dataset de Amazon⁹. Puntualmente, de aquí se extrajeron, por un lado, 6354 títulos de libros y 326 películas, que luego fueron enriquecidos con datos extraídos de la comunidad de Goodreads y IMDB respectivamente. Por el otro, se obtuvieron dos matrices de ratings, una de películas con 67.193 entradas y otra de libros con 223.695.

4.2 Resultados

Para evaluar la propuesta de este trabajo, se tomó como indicador el error en la predicción -tanto el Error Absoluto (MAE) como el Error Cuadrático Medio (MSE)-. Tal como destaca Cantador en [2], en el área de RS multi-dominio, uno de los objetivos es poder lograr predicciones más precisas en D_t utilizando sólo el conocimiento existente en D_s . Por eso, suele utilizarse como dataset de entrenamiento todo lo referente a D_s , dejando para la prueba los ratings de D_t .

A partir del enfoque propuesto, se implementaron una serie de variantes que exploran las posibilidades ofrecidas. Por un lado, se definió toda una rama que mapea los ítems al dominio de las películas (*To movie space*) y el otro que los lleva al de libros (*To book space*). En ambos caso, la variación se encuentra en la forma que se recorre el grafo: de los escritores a los directores (determinando cuánta influencia de cada director tiene un libro) o viceversa (determinando cuánta influencia de cada escritor tiene una película) . A su vez, se posee un grupo de implementaciones cuyo espacio vectorial fue calculado con *GLoVe* y otro con *Word2Vec (W2B)*. La combinación de ambos genera una hibridización, que es manejada con la importancia que se la da a cada aspecto (*Red de Influencias, word embeddings*).

Asimismo, se definieron dos enfoques que se tomaron como base para realizar posteriormente un análisis comparativo. El primero, se trata de uno basado en popularidad, que suele ser uno de los más utilizados en la industria para resolver el problema del *cold-start*[5]. El segundo, es uno que calcula el espacio vectorial utilizando TF-IDF, por ser la alternativa clásica que propone en [1].

El estudio realizado consistió en predecir, para aquellos usuarios que poseían ratings en ambos dominios, los ratings en D_t . Con dichas predicciones se calcularon, los dos indicadores clásicos de error: MAE y MSE. Como se puede observar en la tabla, todas las variaciones del enfoque propuesto mejoran el MAE, mientras que el MSE sólo es mejorado por aquellas variaciones que dan la misma importancia a los *Embeddings* y a la *Red de Influencias*.

Si bien se tratan de resultados preliminares (Tabla 1), en todos los casos se

⁹ <http://jmcauley.ucsd.edu/data/amazon/links.html>

mejora la estimación obtenida con el enfoque basado en popularidad. Sin embargo, al analizar el comportamiento respecto de aquel que usa TFI-DF es posible observar que para un subconjunto de predicciones, los embeddings producen estimaciones que suelen distar mucho del rating definido por el usuario, de allí que el MSE sea un poco mayor. Mientras que si se pone el foco en el MAE, es posible afirmar que las predicciones son más precisas.

Enfoque	Peso Embedding	Peso Red Influencia	MAE	MSE
Popularity baseline			0,8574	1,0773
To movie space	0,5	0,5	0,6837	0,8464
To book space	0,5	0,5	0,7717	0,8854
To movie space Blend (W2B)	0,6	0,4	0,6467	0,9095
To movie space Blend (GLoVe)	0,6	0,4	0,6446	0,9285
To book space Blend (W2B)	0,6	0,4	0,6427	0,8748
To book space Blend (GLoVe)	0,6	0,4	0,6401	0,8998
To movie space Blend (W2B)	0,5	0,5	0,6517	0,8799
To movie space Blend (GLoVe)	0,5	0,5	0,6455	0,8868
To book space Blend (W2B)	0,5	0,5	0,6505	0,8455
To book space Blend (GLoVe)	0,5	0,5	0,6438	0,8587

Table 1: Comparación de métodos

5 Conclusiones

En este trabajo se presentó un enfoque novedoso en el área de SR multi-dominio que utiliza, por un lado, las técnicas de *word embeddings* para representar en un mismo espacio n-dimensional a libros y películas; y, por el otro, usa *DBpedia Ontology* para extraer información del grafo de influencias que subyace entre los artistas. El objetivo fue definir una estrategia que lograra homogeneizar los ítems de D_s y D_t , de forma tal que sean comparables. Allí, es donde juega un rol fundamental la red de influencias, pues ofrece la posibilidad de ver las películas en relación con la influencia que tiene director respecto de los escritores (lo mismo se puede aplicar para el caso de libro con directores).

Por el lado de las técnicas de *word embedding*, éstas se presentan como una alternativa a la clásica TFI-DF, ofreciendo la posibilidad de detectar relaciones implícitas entre palabras que pueden ser provechosas para encontrar similitudes entre las sinopsis de los libros y las películas.

Si bien se trata de una prueba de concepto, los resultados obtenidos mostraron que el enfoque propuesto mejora dos enfoques clásicos que se utilizan ante esta problemática. Por eso, queda como trabajo futuro seguir explorando el enfoque desde otros indicadores como precisión. También se trabajará sobre otras formas de modelar documentos utilizando *word embeddings*.

Bibliografía

1. Balabanović, M., Shoham, Y.: Fab: content-based, collaborative recommendation. *Communications of the ACM* 40(3), 66–72 (1997)
2. Cantador, I., Fernández-Tobías, I., Berkovsky, S., Cremonesi, P.: *Recommender Systems Handbook*, chap. Cross-Domain Recommender Systems. Springer US (2015)
3. Das, A., Ganguly, D., Garain, U.: Named entity recognition with word embeddings and wikipedia categories for a low-resource language. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)* 16(3), 18 (2017)
4. Fernández-Tobías, I., Cantador, I., Kaminskas, M., Ricci, F.: A generic semantic-based framework for cross-domain recommendation. In: *Proceedings of the 2nd International Workshop on Information Heterogeneity and Fusion in Recommender Systems*. pp. 25–32. ACM (2011)
5. de Gemmis, M., Lops, P., Musto, C., Narducci, F., Semeraro, G.: *Recommender Systems Handbook*, chap. Semantic-Aware Content-Based Recommender Systems. Springer US (2015)
6. Li, B., Yang, Q., Xue, X.: Can movies and books collaborate? cross-domain collaborative filtering for sparsity reduction. In: *IJCAI*. vol. 9, pp. 2052–2057 (2009)
7. Lizarralde, I., Rodríguez, J.M., Mateos, C., Zunino, A.: Word embeddings for improving rest services discoverability. In: *Computer Conference (CLEI), 2017 XLIII Latin American*. pp. 1–8. IEEE (2017)
8. Magnini, B., Strapparava, C.: Improving user modelling with content-based techniques. In: *International Conference on User Modeling*. pp. 74–83. Springer (2001)
9. Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S., Dean, J.: Distributed representations of words and phrases and their compositionality. In: *Advances in neural information processing systems*. pp. 3111–3119 (2013)
10. Nietzsche, F.W.: *El nacimiento de la tragedia o Helenismo y Pesimismo*. Espasa-Calpe (1943)
11. Ozsoy, M.G.: From word embeddings to item recommendation. arXiv preprint arXiv:1601.01356 (2016)
12. Pan, W., Xiang, E.W., Yang, Q.: Transfer learning in collaborative filtering with uncertain ratings. In: *AAAI*. vol. 12, pp. 662–668 (2012)
13. Pennington, J., Socher, R., Manning, C.: Glove: Global vectors for word representation. In: *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*. pp. 1532–1543 (2014)
14. Ricci, F., Rokach, L., Shapira, B.: *Recommender Systems Handbook*, chap. Introduction. Springer US (2015)
15. Socher, R., Bauer, J., Manning, C.D., et al.: Parsing with compositional vector grammars. In: *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. vol. 1, pp. 455–465 (2013)
16. Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C.D., Ng, A., Potts, C.: Recursive deep models for semantic compositionality over a sentiment treebank. In: *Proceedings of the 2013 conference on empirical methods in natural language processing*. pp. 1631–1642 (2013)