Journal of Cases on Information Technology
Volume 26 • Issue 1

Social Recommender System Based on CNN Incorporating Tagging and Contextual Features

Muhammad Alrashidi, Universiti Teknologi Malaysia, Malaysia*
Ali Selamat, Universiti Teknologi Malaysia, Malaysia
https://orcid.org/0000-0001-9746-8459
Roliana Ibrahim, Universiti Teknologi Malaysia, Malaysia
Hamido Fujita, Universiti Teknologi Malaysia, Malaysia

ABSTRACT

The Internet's rapid growth has led to information overload, necessitating recommender systems for personalized suggestions. While content-based and collaborative filtering have been successful, data sparsity remains a challenge. To address this, this article presents a novel social recommender system based on convolutional neural networks (SRSCNN). This approach integrates deep learning and contextual information to overcome data sparsity. The SRSCNN model incorporates user and item factors obtained from a neural network architecture, utilizing features from item titles and tags through a CNN. The authors conducted extensive experiments with the MovieLens 10M dataset, demonstrating that the SRSCNN approach outperforms state-of-the-art baselines. This improvement is evident in both rating prediction and ranking accuracy across recommendation lists of varying lengths.

KEYWORDS

CNN, Deep Learning, Recommender Systems, Social Networks, Social Recommendation

1. INTRODUCTION

The rapid growth of the Internet has not only contributed to making life easier but has also generated an enormous volume of information, leading to the issue of information overload. It has become challenging for individuals to select their desired items that align with their preferences from the vast array of choices available. To address this, recommender systems (RSs) play a crucial role in alleviating information overload and enhancing user experiences. RSs are intelligent tools and methodologies designed to propose items that are likely to capture the interest of a particular user. These recommendations frequently play a role in various decision-making processes, such as helping users decide which products to purchase, which songs to listen to, and which articles to read. The term

DOI: 10.4018/JCIT.335524
*Corresponding Author

"item" commonly refers to the suggestions provided by the system. The system typically presents personalized recommendations in the form of ranked lists. These recommendations are generated by predicting the most suitable items or services for an individual based on their preferences (Ricci et al., 2022).

Three main types of RSs are commonly employed: content-based filtering, collaborative filtering (CF), and hybrid approaches (Da'u & Salim, 2020). These methods rely on the fundamental principle that individuals with similar past preferences are likely to make comparable future choices. CF recommenders specifically utilize data about user-item interactions to formulate their suggestions. Additionally, there exists a subset of RSs that incorporate contextual information related to users and/or items to enhance their recommendations. Among these types, content-based recommenders stand out as a distinct category within RSs. Moreover, certain hybrid RSs amalgamate insights from both user and item perspectives to provide comprehensive recommendations (Batmaz et al., 2019).

Despite the effectiveness of CF methods in various scenarios, the issue of data sparsity remains a significant hurdle (Da'u & Salim, 2020) (Batmaz et al., 2019). This challenge, stemming from an insufficient number of user ratings for each item, often hampers the performance of conventional approaches. Another notable concern is the cold start problem (Shokeen & Rana, 2020a), which significantly impacts both new users and new items. Providing meaningful recommendations without access to relevant historical data proves to be challenging. While novel users can receive suggestions for popular items, this approach falls short of delivering truly personalized recommendations.

Addressing these challenges necessitates the incorporation of supplementary data such as user profiles, product descriptions, and information from social platforms (Sun et al., 2019). Several research projects have explored the collection of additional data, and one approach for tackling data scarcity involves extracting insights from review texts (L. Zheng et al., 2017). In instances where adequate historical user-item interaction data is lacking, leveraging contextual information becomes pivotal in making informed inferences about user preferences. Individuals' preferences can be aggregated through social networks, allowing for the deduction of their tastes. Further data sources, like item tags or categories, can be directly employed to comprehend user preferences. For instance, a user's favorite movie genres or music album categories can shed light on the types of content they prefer (Sun et al., 2019). Contextual factors, such as timing, location, or tags, possess the capacity to influence an individual's perception of a particular item (Chen & Li, 2019). Tag-related data, on the other hand, supplements existing datasets by elucidating item attributes and capturing user preferences through tagging behaviors (J. Zhang et al., 2019b). Consequently, integrating supplementary information into the recommendation process emerges as a viable solution to tackle the challenges posed by data sparsity and the cold start problem in RSs.

While incorporating supplementary data often enhances the accuracy of RSs, it is important to note that this additional data is frequently more sparse compared to the user-item rating matrix (Ahmadian et al., 2022). In contrast, matrix representations connecting users, items, and tagged data possess high dimensions due to the typical abundance of users, products, and tags in real-world scenarios. Consequently, adding this extra information to the recommendation process can make techniques a lot more complicated, which could make them less useful for large-scale systems (Z. Zhang et al., 2017).

Deep learning techniques have demonstrated effectiveness in RSs by leveraging textual attributes like reviews and item descriptions within deep learning models (L. Zheng et al., 2017) (Yin et al., 2017) (Tuan & Phuong, 2017) (Catherine & Cohen, 2017) (Yao et al., 2017). However, challenges persist in terms of the sparsity issue within the user-item rating matrix and handling the additional data, which could potentially render the learned latent factors ineffective. Moreover, the efficacy of extracting valuable features from diverse data sources critically influences the success of recommendation techniques in terms of enhancing prediction accuracy. In light of these considerations, a potential solution to the data sparsity problem in recommendations involves the strategic utilization of additional data during the recommendation process. Directly harnessing supplementary information, such as