2024, 21(2 Special Issue): 0592-0599
https://dx.doi.org/10.21123/bsj.2024.9756
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

Hybrid CNN-based Recommendation System

Muhammad Alrashidi[1], Roliana Ibrahim[1], Ali Selamat[1,2]***

[1]Faculty of Computing, Universiti Teknologi Malaysia, Johor Baru, Johor 80000 Malaysia.
[2]Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, Kuala Lumpur, WP 50088 Malaysia.
*Corresponding Author.
PARS2023: Postgraduate Annual Research Seminars 2023.

Received 29/09/2023, Revised 10/02/2024, Accepted 12/02/2024, Published 25/02/2024

Abstract

Recommendation systems are now being used to address the problem of excess information in several sectors such as entertainment, social networking, and e-commerce. Although conventional methods to recommendation systems have achieved significant success in providing item suggestions, they still face many challenges, including the cold start problem and data sparsity. Numerous recommendation models have been created in order to address these difficulties. Nevertheless, including user or item-specific information has the potential to enhance the performance of recommendations. The ConvFM model is a novel convolutional neural network architecture that combines the capabilities of deep learning for feature extraction with the effectiveness of factorization machines for recommendation tasks. The present work introduces a novel hybrid deep factorization machine (FM) model, referred to as ConvFM. The ConvFM model use a combination of feature extraction and convolutional neural networks (CNNs) to extract features from both individuals and things, namely movies. Following this, the proposed model employs a methodology known as factorization machines, which use the FM algorithm. The focus of the CNN is on the extraction of features, which has resulted in a notable improvement in performance. In order to enhance the accuracy of predictions and address the challenges posed by sparsity, the proposed model incorporates both the extracted attributes and explicit interactions between items and users. This paper presents the experimental procedures and outcomes conducted on the Movie Lens dataset. In this discussion, we engage in an analysis of our research outcomes followed by provide recommendations for further action.

Keywords: CNN, deep learning, Recommendation systems, Social networks, Social recommendation.

Introduction

The majority of recommendation systems use collaborative filtering. There are two major difficulties in the field of collaborative filtering i.e., data sparsity and cold start[1]. Consumers experience data sparsity when they are provided with an excessively large number of products, but even the most active users only score a small portion of those products. It could be challenging to deduce user preferences. In collaborative filtering models, the interaction between a user and an object is frequently represented by a rating matrix, whereas

2024, 21(2 Special Issue): 0592-0599
https://dx.doi.org/10.21123/bsj.2024.9756
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

factorization machines employ a feature vector and the rating as a class label[2].

Another issue is a cold start, which is an issue since there is minimal potential for cold-start items to be revealed to users. It's difficult to figure out what a new user's preferences are. Therefore, side information is often employed in recommender systems to address problems such as cold-starting and data sparsity problems. Even though Deep Learning approaches outperform other model-based recommendation approaches such as Latent factor models and Representation learning models, however, research into how to effectively integrate various side information into Deep Learning techniques has not yet realized its complete potential[3]. Deep learning models can therefore train high-order and non-linear latent representations by utilizing a variety of activation functions, including ReLu, sigmoid, and other activation functions.

Big data, cloud computing, and other innovations have tremendously improved our productivity and quality of life. There has also been a significant increase in the amount of information available on the network. Recommendation systems aim to provide useful recommendations to each user based on their preferences and behaviors. However, with the rapid increase of network data in recent years and taking the characteristics of big data into consideration, the traditional recommendation systems are not completely suitable for the current network environment. In social networks, the user's evaluation record for a certain item is incomplete, and there are a lot of null values in the evaluation-related index of the user and the item. Therefore, the traditional recommendation system suffers from sparse data, which reduces the accuracy of the recommendations.

Deep neural networks are effective at extracting usable representations and underlying explanatory variables from input data. In general, real-world apps provide a variety of meaningful information about products and users. Using this data allows us to improve our knowledge of things and users, resulting in a more accurate recommender. As a result, using deep learning methods to representation learning in recommendation algorithms is an obvious option. deep learning methods for representation learning have the benefit of allowing recommendation models to incorporate diverse content information such as text, photos, audio, and even video[4]. The CNN (Convolutional Neural Network) is a feedforward neural network designed for computer vision. There has been research on the use of CNN to extract supplementary data in order to construct predictions, albeit it is not widely used in the field of recommendation systems. Traditional recommendation systems, such as[5], lack several features that CNN-based recommendation algorithms provide, such as tolerance of faults, parallel processing, and the capacity to self-learn. They can tackle issues such as complicated environmental data, ambiguous historical data, and ambiguous reasoning principles. They allow for more flaws and defects in the data samples. CNNs also are rapid in execution, have strong adaptability and high resolution. CNNs combine feature extraction into a multilayer perceptron framework by adapting the structure and decreasing weights, obviating the time-consuming procedure of extracting features before recognition. The capacity of CNN to generalize is superior to that of other approaches.

Many existing related studies have used Matrix Factorization to create recommendation systems. By the nature of Matrix Factorization, only the user data and item data is projected onto a common latent space. Researchers utilized auto encoders[6-8] in combination with matrix factorization. While Matrix Factorization provides many advantages, it may be limited in the larger scope of features that may be present within a dataset. The Factorization Machine(FM) proposed by[9] takes advantage of the flexibility in feature engineering and the excellent predictive ability of factorization to produce a list of features which are ranked. This will allow for a more holistic approach when creating a prediction. Factorization Machine is an extension of this system that allows for many more features to be considered while also minimizing dimensionality which allows for more accurate and reliable recommendations. When combined with a CNN model framework, the accuracy of the predictions can be effectively improved.

2024, 21(2 Special Issue): 0592-0599
https://dx.doi.org/10.21123/bsj.2024.9756
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

In this study, CNN has been utilized to model user-item ratings from the data, while also employing the factorization machine (FM) method to generate recommendations. The proposed technique ConvFM is just a posed models baseline model to investigate the variation in accuracy due to the additional implementation of CNN in the ConvFM model. This study demonstrates the aforementioned statement while comparing the baseline model to the proposed model.

This study proposes a ConvFM model that first uses feature extraction in tandem with CNN to obtain features from both users and items. It can enhance the recommendation performance by strengthening prediction accuracy and tackling sparsity difficulties.

The proposed model integrates the extracted features as well as explicit item-user interaction that combines the capability of factorization machines for recommendation with deep learning for feature learning.

Materials and Methods

The purpose of this study is to introduce a hybrid deep factorization machine (FM) called ConvFM. ConvFM first uses feature extraction in tandem with CNN to obtain features from both users and items. The major goal of CNN in the proposed model is to create an item latent vector that will be employed in the recommender system's rating prediction. CNN has the ability to extract complicated and important feature representation. Subsequently, the model uses a technique called factorization machines, which utilizes the FM algorithm.

A suggested model ConvFM is divided into four parts, as seen in Fig. 1, including a Data processing phase, a phase for Extracting Features by CNN, a phase for Recommender system, and a phase for Evaluation.

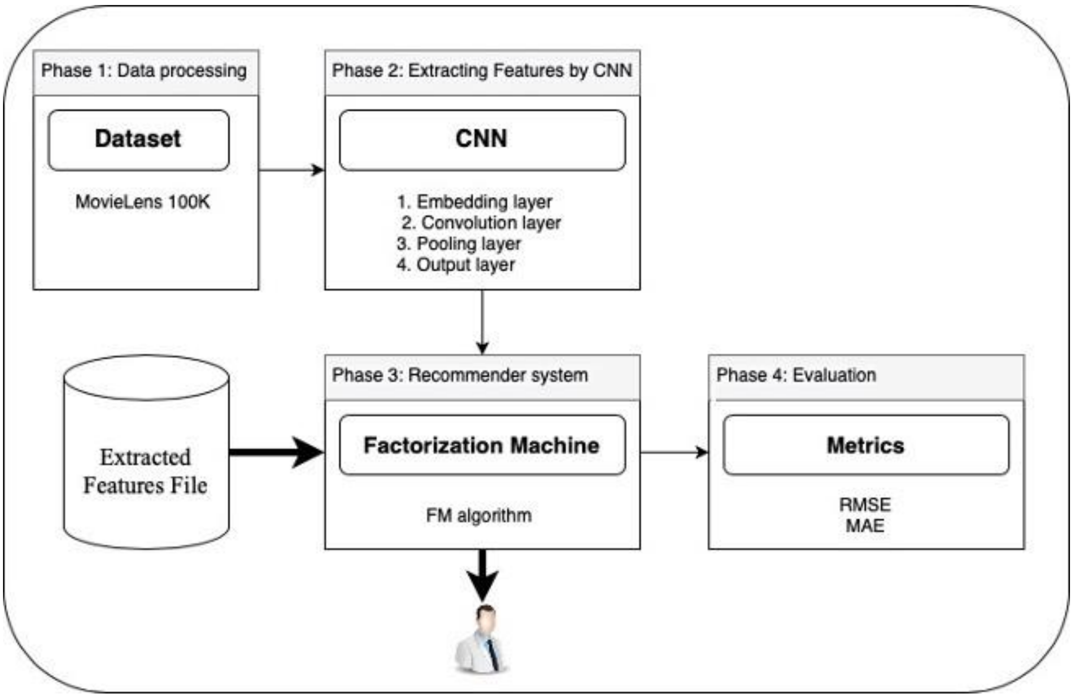

Figure 1. ConvFM Overview.

Experiments

The major goal of CNN in the proposed model is to create an item latent vector that will be employed in the recommender system's rating prediction. CNN has the ability to extract complicated and important feature representation. The proposed ConvFM model applies latent features derived from both the

2024, 21(2 Special Issue): 0592-0599
https://dx.doi.org/10.21123/bsj.2024.9756
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

user and the item by CNN to the FM latent factor training process.

Dataset

The MovieLens 100K Dataset from the MovieLens website served as the study's dataset. The dataset has a sample size of 100,000 ratings on a scale of 1 to 5. There is a dataset of 943 users and a total of 1682 movies. Every user in the dataset was selected with the criteria that they must have rated at least 20 movies and also must have additional demographic information such as age, gender, occupation, and , zip.

Baseline algorithms

In order to validate the performance of our proposed technique, ConvFM is compared with the following state-of-the-art baseline models:

- Hybrid recommendation system: (Gallitelli, 2019) is proposed multiple deep learning architectures for recommendation system (DeepWideNet, CBF-A, AMAR and DeepWideNet + NLP).
- SVD: A well-known SVD algorithm popularized during the Netflix Prize by (Koren, Bell and Volinsky, 2009).
- SVD++: The SVD++ algorithm proposed by (Koren, 2008) is an expansion of the SVD algorithm.
- NMF: Based on non-negative matrix factorization, (Xin Luo et al., 2014) proposes a collaborative filtering technique.

Experimental setup

ConvFM was implemented on Google Colab using TensorFlow framework and LibRecommender 0.6.10 with Python 3.7.12. The initial steps of data preparation involve reading the data and further converting it into TensorFlowow dataframe so it may be used as input in the models. Subsequently, the following steps are applied for the proposed model.

Extracting Features by CNN

The major goal of CNN in the proposed model is to create an item latent vector that will be employed in the recommender system's rating prediction. CNN has the ability to extract complicated and important feature representation. CNN is used to extract features from the dataset. The neural network uses the obtained feature information to classify the data. In the feature extraction process of the CNN architecture, convolution layer piles and sets of pooling layers are used. The convolution layer, as its name suggests, uses the convolution method to modify the data. Convolution layers can be described as a sequence of digital filters. The pooling layer combines the pixels of adjacent pixels into a single pixel. The dimension of the data is subsequently reduced by the pooling layer.

The convolution and pooling layers' methods are essentially in a two-dimensional plane, as CNN's primary focus is the data. Here a GlobalMaxPooling Layer which is used due to its ability to enforce correlations between feature maps and categories, making it more natural to the convolution structure. Furthermore, there is an embedding layer added to convert each word into a non-binary fixed length vector with real number values. The fixed length of these word vectors have reduced dimensions and the words themselves are better represented. Additionally, dense layers are added for matrix vector multiplication. The model described above is demonstrated with the following flow chart of Extracting Features by CNN in Fig. 2.

2024, 21(2 Special Issue): 0592-0599
https://dx.doi.org/10.21123/bsj.2024.9756
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

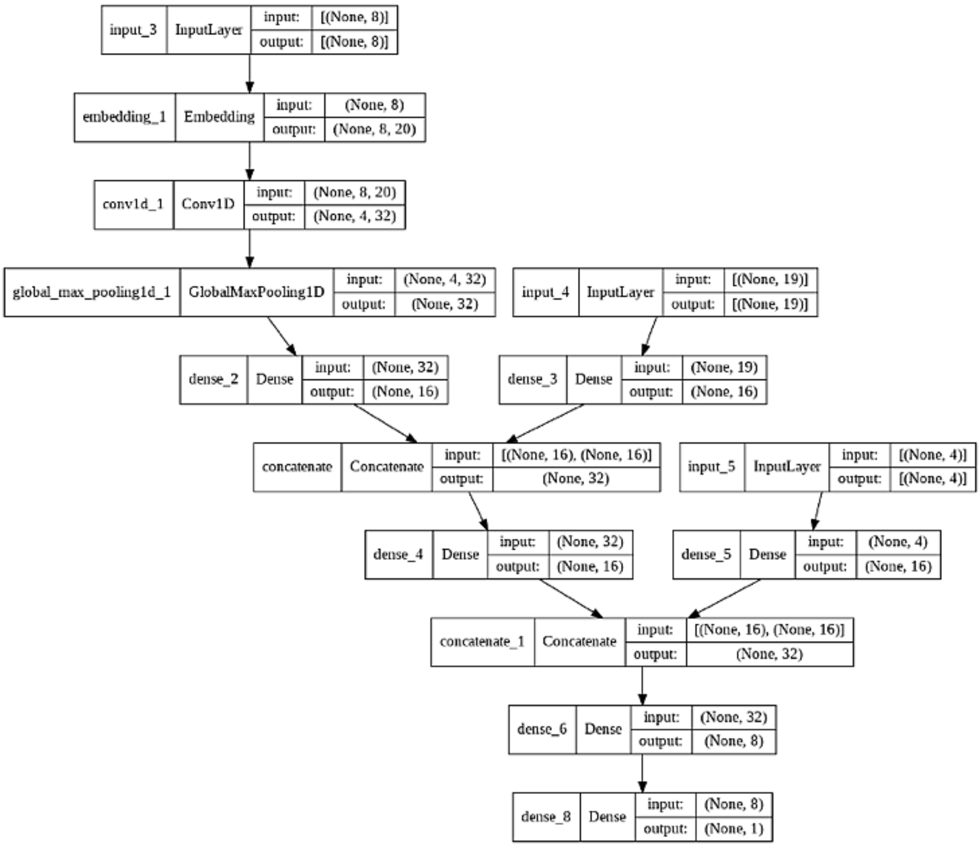

Figure 2. The flow chart of Extracting Features by CNN.

Evaluation method

Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) were utilized as assessment metrics due to their close connection to the objective function of a traditional rating prediction model.

Results and Discussion

In this work, a variety of deep neural network topologies have been proposed in an effort to implement various strategies and address the recommendation problem. All of the provided architectures were created using the Keras deep learning framework with a Tensorflow backend and assessed using well-known datasets like MovieLens and the two most widely used assessment metrics for regression issues, MAE and RMSE.

Table 1. MAE and RMSE values for other existing techniques using MovieLens Datasets

Method	RMSE	MAE
ConvFM	0.708	0.504
DeepWideNet	0.908	0.684
CBF-A	0.969	0.719
AMAR	0.975	0.714
DeepWideNet + NLP	0.889	0.677
SVD	0.934	0.737
SVD++	0.920	0.722
NMF	0.963	0.758

2024, 21(2 Special Issue): 0592-0599
https://dx.doi.org/10.21123/bsj.2024.9756
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

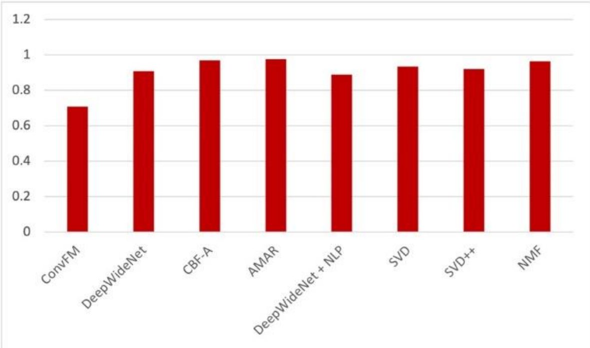

Figure 3. Results of an RMSE comparison.

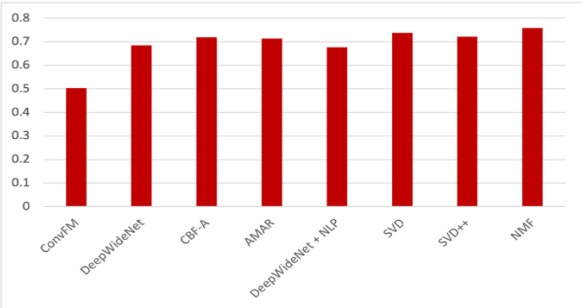

Figure 4. Results of an MAE comparison.

The findings of comparing ConvFM with the aforementioned approaches are listed in Table 1 and illustrated in Fig. 3 and Fig.4. The suggested model acquires a lower root mean square error (RMSE) and mean absolute error (MAE) than current techniques. The Movie Lens 100k Dataset was used for analysis, and the results were compared with those of others. The previous maximum output was achieved by DeepWideNet + NLP technique, and our approach even outperformed that with a good margin. Our Hybrid CNN-based FM Model (ConvFM) out performed all the previous methods (like DeepWideNet, CBF-A, AMAR, DeepWide Net + NLP, SVD, SVD++ and NMF) in both mean square error (MSE) and root mean square error (RMSE). The previous best for RMSE and MSE were 0.889 and 0.677, respectively, and the results have been improved by 0.181 and 0.172, respectively.

Conclusion

This study proposes an innovative convolutional neural network (CNN) structure that combines the feature learning capabilities of deep learning with the recommendation abilities of factorization machines. ConvFM is a model that combines deep learning techniques with factorization machines, resulting in a hybrid architecture. The use of the Factorization Machine technique, sourced from the LibRecommender Library, is employed in conjunction with Convolutional Neural Networks (CNN) to merge the user's explicit input with the latent properties derived from the metadata of the objects. This approach not only enhances the accuracy of predictions but also addresses challenges related to data sparsity. The CNN utilizes a process of extracting features from data that may be learned, and then applies these features in a factorization machine (FM) to calculate the scores for all categories. The experimental findings demonstrate that our proposed methodology exhibits superior performance compared to the frequency modulation (FM) approach that exclusively relies on historical data scoring. The findings indicate that the proposed ConvFM model has superior performance compared to the baseline models in terms of accuracy. The ConvFM model effectively addresses the limitations seen in traditional methodologies and significantly improves the overall performance of the recommendation system. The ConvFM model proposed in this study effectively addresses the cold-start issue and significantly improves the efficiency and accuracy of recommendation processing. This assertion is supported by the experimental results obtained from the analysis of the MovieLens dataset. Moreover, next investigations will be conducted on diverse datasets exhibiting different levels of density in order to showcase the capability of the proposed model to effectively manage sparsity in data and provide accurate recommendations.

Acknowledgment

This work was supported/funded by the Ministry of Higher Education under the Fundamental Research Grant Scheme (FRGS/1/2018/ICT04/UTM/01/1). The authors sincerely thank Universiti Teknologi Malaysia under Research University Grant Vot-20H04 and Malaysia Research University Network Vot 4L876, for the completion of the research.

Authors' Declaration

- Conflicts of Interest: We have no conflicts of interest to disclose.

Authors' Contribution Statement

M. A., R. I. and A. S. contributed to the design and implementation of the research, to the analysis of the results and to the writing of the manuscript.

References

1. Et al. A-B. A Study on the Accuracy of Prediction in Recommendation System Based on Similarity Measures. Baghdad Sci. J. . 2019 Mar 17;16(1(Suppl.)):0263. https://doi.org/10.21123/bsj.2019.16.1(Suppl.).0263
2. Loni B, Shi Y, Larson M, Hanjalic A. Cross-Domain Collaborative Filtering with Factorization Machines. SpringerLink . 2014; p. 656–61. https://doi.org/10.1007/978-3-319-06028-6_72.
3. Sun Z, Guo Q, Yang J, Fang H, Guo G, Zhang J, et al. Research commentary on recommendations with side information: A survey and research directions. Electron Commer Res Appl. .2019 Sep;37:100879. https://doi.org/10.1016/j.elerap.2019.100879.
4. Zhang S, Yao L, Sun A, Tay Y. Deep Learning Based Recommender System. ACM Comput Surv. . 2020 Jan 31;52(1):1–38. https://doi.org/10.1145/3285029.
5. Lin P-C, Arbaiy N. User-Oriented Preference Toward a Recommender System. Baghdad Sci J . 2021 Mar 30;18(1(Suppl.)):0746. https://doi.org/10.21123/bsj.2021.18.1(Suppl.).0746
6. Tahmasebi H, Ravanmehr R, Mohamadrezaei R. Social movie recommender system based on deep autoencoder network using Twitter data. Neural Comput Appl. 2021 Mar 16;33(5):1607–23. https://doi.org/10.1007/s00521-020-05085-1.
7. Pan Y, He F, Yu H. Learning social representations with deep autoencoder for recommender system. springer. 2020 Jul 7;23(4):2259–79. https://doi.org/10.1007/s11280-020-00793-z.
8. C C N, Mohan A. A social recommender system using deep architecture and network embedding. Appl Intell . 2019 May 18;49(5):1937–53. https://doi.org/10.1007/s10489-018-1359-z.
9. Rendle S. Factorization Machines. In: 2010 IEEE International Conference on Data Mining. IEEE. 2010; p. 995–1000. https://doi.org/10.1109/ICDM.2010.127.

2024, 21(2 Special Issue): 0592-0599
https://dx.doi.org/10.21123/bsj.2024.9756
P-ISSN: 2078-8665 - E-ISSN: 2411-7986

Baghdad Science Journal

نظام التوصية الهجين القائم على CNN

محمد الرشيدي[1]، روليانا ابراهيم[1]، علي سلامات[1،2]

[1]كلية الحاسبات، الجامعة التكنولوجية الماليزية، جوهور بارو، جوهور، 80000، ماليزيا.
[2]المعهد الماليزي الياباني الدولي للتكنولوجيا، الجامعة التكنولوجية الماليزية، كوالالمبور، WP 50088، ماليزيا.

الخلاصة

يتم الآن استخدام أنظمة التوصية لمعالجة مشكلة المعلومات الزائدة في عدة قطاعات مثل الترفيه والشبكات الاجتماعية والتجارة الإلكترونية. على الرغم من أن الطرق التقليدية لأنظمة التوصية قد حققت نجاحًا كبيرًا في تقديم اقتراحات العناصر، إلا أنها لا تزال تواجه العديد من التحديات، بما في ذلك مشكلة البداية الباردة وتناثر البيانات. وقد تم إنشاء العديد من نماذج التوصيات لمعالجة هذه الصعوبات. ومع ذلك، فإن تضمين معلومات خاصة بالمستخدم أو العنصر لديه القدرة على تحسين أداء التوصيات. نموذج ConvFM عبارة عن بنية شبكة عصبية تلافيفية جديدة تجمع بين إمكانيات التعلم العميق لاستخراج الميزات وفعالية آلات التحليل لمهام التوصية. يقدم العمل الحالي نموذجًا جديدًا لآلة التحليل العميق الهجين (FM)، يشار إليه باسم ConvFM. يستخدم نموذج ConvFM مزيجًا من استخراج الميزات والشبكات العصبية التلافيفية (CNNs) لاستخراج الميزات من كل من الأفراد والأشياء، أي الأفلام. بعد ذلك، يستخدم النموذج المقترح منهجية تعرف بآلات التحليل، والتي تستخدم خوارزمية FM. ينصب تركيز CNN على استخراج الميزات، مما أدى إلى تحسن ملحوظ في الأداء. من أجل تعزيز دقة التنبؤات ومواجهة التحديات التي يفرضها التناثر، يتضمن النموذج المقترح كلاً من السمات المستخرجة والتفاعلات الواضحة بين العناصر والمستخدمين. تعرض هذه الورقة الإجراءات والنتائج التجريبية التي أجريت على مجموعة بيانات Movie Lens. في هذه المناقشة، ننخرط في تحليل نتائج بحثنا متبوعًا بتقديم توصيات لاتخاذ مزيد من الإجراءات.

الكلمات المفتاحية: CNN، التعلم العميق، أنظمة التوصية، الشبكات الاجتماعية، التوصية الاجتماعية.