International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 09 | Sep 2020 www.irjet.net p-ISSN: 2395-0072

Improved System for Performance Evolution in Recommendation Model

Robin Sihag[1], Ruchi Rautela[2]

[1]*Student, Vivekanand Education Society's Institute of Technology, Mumbai*
[2]*Assistant Professor, Dept. of MCA, Vivekanand Education Society's Institute of Technology, Mumbai*

Abstract - *In digital era recommendation became an integral part of business. Association rule mining augments the performance of recommendation system. But it still faces the challenge of data sparsity, cold start as well as recommendations to new user. Data sparsity arises due to limited reviews and ratings available to build a recommendation system. In this paper we propose deep learning to diminish the impact of these challenges. The Recurrent Neural Network extract features and contextual data from the recommendations. Results present the influence of RNN with contextual data on the recommendation models. This paper represents influence of the deep learning with contextual data on the recommendation models.*

Key Words: Deep Learning, Association Rule Mining(ARM), Recommendation System, Data Sparsity.

1. INTRODUCTION

Recommendation system are subclass of information filtering system that make predictions based on the user's interest of things. Recommendation system need large amount of data from users and training time. As the data size increases with time the sparsity of data increases. To reduce sparsity of data and drawing out same patterns repetitively from the same datasets only limited and essential data patterns must be stored in memory. Hence proposed system mainly focuses on the four components Data Preprocessing, Adding metadata and tags, Association rule mining and ranking, Deep learning and prediction.

Association rules is widely used for making recommendations. It uses the conditional probability of two mutually exclusive and independent events. Artificial neural network is used in various tasks such as prediction, classification, pattern recognition tasks

1.1 Association Rule Minning(ARM)

As the name implies, association rule mining is where association rules are simple If-then statements that help discover relationships between independent transactions within a dataset or itemset. Most machine learning algorithms tend to be mathematical because they operate on numeric data sets. However, association rule mining is suitable for non numeric categorical data and requires less than a simple count. The purpose of association rule mining is to observe frequently occurring patterns, correlations or associations from datasets of different types of data resources.

1.2 Deep Learning(DL)

Deep learning is an artificial intelligence feature that processes data used for decision making and mimics the work of the human brain when creating patterns. Deep learning is a subset of machine learning in artificial intelligence (AI), with networks that can learn unsupervised from unstructured or unlabelled data. Also called deep neural learning or deep neural network. Deep learning systems require large amounts of data to produce accurate results. Therefore, the information is provided as a large dataset. When processing data, artificial neural networks can categorize the data with the answers received from a series of binary true or false questions that involve very complex mathematical calculations.

1.3 Recurrent Neural Network(RNN)

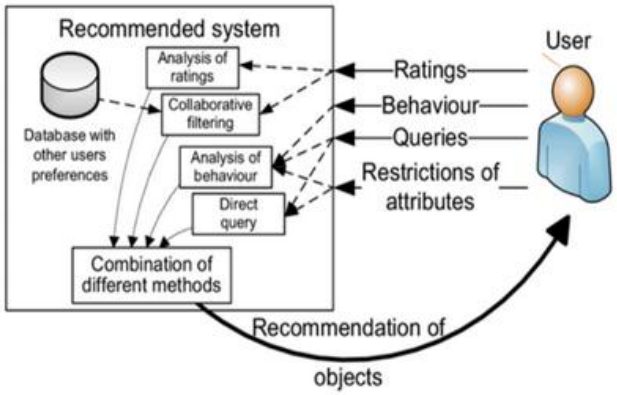
The Recurrent Neural Network remembers the past, and its decisions are influenced by learning from the past. Basic feedforward networks also "remember" things, but remember what they learned during training. For example, an image classifier learns what a "1" looks like during training, and uses that knowledge to classify things in production. RNNs learn similarly during training, but also remember what they learned from previous inputs when generating outputs. It is part of the network. An RNN can take one or more input vectors and produce one or more output vectors, and the output is "hidden" representing context based on a priori, as well as weights applied to inputs like a regular NN Is also affected by the state vector. Input Output. Therefore, the same input can produce different outputs depending on previous inputs in the series.

2. LITERATURE SURVEY

The recommender system solves the problem of information overload of users by searching a large amount

of dynamically generated data, and provides users with personalized contents and services. The primary purpose of the online service portal recommendation system is to rank and list items that are not visible to the user or that the user has not searched for, rated or displayed on the online shopping portal [2]. In an e-commerce setting, a recommendation system increases revenue because it is an effective means of selling more products. From an e commerce perspective, a recommendation system can be viewed as a tool that helps users find records of knowledge related to their interests and preferences. In recent years, various approaches have been developed to build recommendation systems that can use either collaborative, content-based, or hybrid filtering.

Figure -1: Existing system of recommendation model [10]



Recommendation systems are built on filtering of transactions or item sets in the relational database. Mainly two types of filtering techniques are used, that are, Collaborative filtering and Content-based filtering.

· A collaborative approach recommends the product by considering other users with similar preferences, and uses opinions to recommend items to active users. The collaborative approach has issues such as cold start, sparsity, and scalability issues.

· Content-based filtering techniques typically make recommendations based on user information and ignore posts from other users. Some of the issues associated with content-based filtering techniques are limited content analytics and over specialization, resulting in sparse data.

· The hybrid approach combines a content-based approach with a collaborative approach to mitigate the problems faced separately by the two filtering techniques. Hybrid recommendation systems can also be enhanced with various algorithms and knowledge-based techniques to further improve the accuracy of the recommendation system. The specific consideration of knowledge-based systems raises the need for knowledge acquisition, a well known bottleneck for many artificial intelligence applications.

3. LIMITATIONS OF EXISTING SYSTEM

3.1 Data Sparsity:

The recommendation system relies on users in the user review matrix to recommend to other users. However, when a user places a purchase in a large item or browses a movie or music list, If users did not rate these items, data sparsity evolved. Data sparsity simply means absence of data of selected features.

3.2 Cold start:

Cold start issues mainly occur when there are new users on the site or when adding new items to the system. First, he does not know how to recommend the item to new users and has not yet rated the item. Second, anyone who can recommend this new item to others, nobody will rate this item

3.3 Diversity:

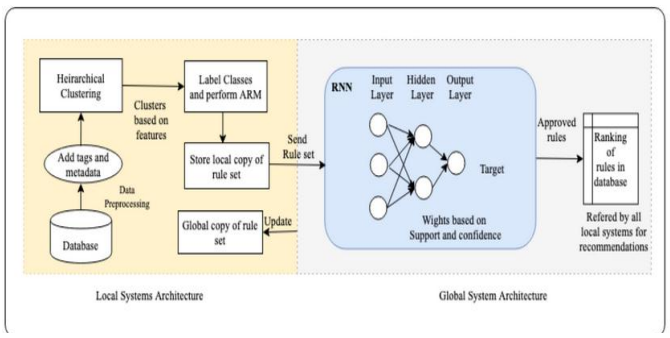
Recommender results should span as much of the item area as possible and not all come from the same cluster. Producing diverse results is not possible with little information.

4. PROPOSED SYSTEM

Proposed system is suitable for categorical data with unknown labels. Therefore, use a clustering algorithm to find the correlation between each transaction. The similarity between these transactions determines the cluster of data points in the plane. To build recommendation model data pre processing and providing metadata are two important steps involved.

Following are the details of building recommendation model with deep learning

Figure 2. Proposed System Architecture



1. Collection of data:
Data is collected from various database sources. This data may contain user information, product information, user ratings and reviews on product.

2. Adding metadata and tagging:
Perform data cleaning and pre-processing on the collected data. Add suitable tags to the data and add metadata to the products or items and users. For example, milk, curd, cheese are dairy product so add tag as "dairy". Adding tags to data helps to reduce cold start problem for new users and helps to deal with data sparsity.

3. Data Pre-processing:
Apply clustering algorithm on the cleaned data. Give labels to the clusters and select features from the data. Apply feature selection algorithm,

4. Association Rule Mining Algorithm:
Apply hybrid filtering on the dataset and rank the items based in the scores. Prepare review datase along with the ranking dataset.

5. Feed item sets and scores to the RNN with the reviews dataset and predict the score of the each item. Split data in train and test datasets.

6. Distribute ranked copies over the local system

5. Conclusion

Current system is studied and limitations were identified .Hybrid approach have better chances in improvising the result than the traditional approach. In traditional approach large amount of data is processed together to draw out patterns. This proposed system is useful to build recommendation models on contextual data and enormous amount of user-item correlated data.

REFERENCES

[1] K. Shah, A. Salunke, S. Dongare and K. Antala, "Recommender systems: An overview of different approaches to recommendations," 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), Coimbatore, 2017, pp. 1-4. "Self-Tuning Approach for Implementing a Multidimensional Recommendation System Using PID", International Conference on Innovations in Power and Advanced Computing Technologies [i-PACT2017], 978-1-5090-5682- 8 /17/$31.00 ©2017 IEEE

[2] F.O. Isinkaye, Y.O. Folajimi, B.A. Ojokoh, "Recommendation systems: Principles, methods and evaluation", Egyptian Informatics Journal, vol. 16, 2015, pp. 261–273.

[3] Z. Chen, W. Liang, X. Gao, Z. Zhou and M. Wu, "Research on the Accurate Recommendation Management System for Employment of College Graduates on Hadoop," 2019 5th International Conference on Big Data and Information Analytics (BigDIA), Kunming, China, 2019, pp. 19-22.

[4] Agrawal, R., Gehrke, J., Gunopulos, D. et al. Automatic Subspace Clustering of High Dimensional Data. Data Min Knowl Disc 11, 5– 33 (2005). https://doi.org/10.1007/s10618-005- 1396-1.

[5] M. Alfarhood and J. Cheng, "DeepHCF: A Deep Learning Based Hybrid Collaborative Filtering Approach for Recommendation Systems," 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, 2018, pp. 89-96.

[6] Schmidhuber, J. (2015). Deep Learning in Neural Networks: An Overview. Neural Networks 61: 85-117.

[7] Deshpande, M., Karypis, G.: Item-based top-N recommendation algorithms. ACM Transaction on Information Systems 22(1), 143177 (2004).

[8] Kenneth Wai-Ting Leung, Wilfred Ng, and Dik Lun Lee, "Personalized Concept-Based Clustering of Search Engine Queries", Knowledge and Data Engineering IEEE

[9] https://www.researchgate.net/figure/Structure-of-a-recommender-system_fig2_220827211

[10] https://www.kaggle.com/gspmoreira/articles sharing-reading-from-cit-deskdrop