EasyChair Preprint

№ 9176

Enhanced Product Recommendation System for e-Commerce Using Machine Learning

R.M Rani, Sourav Singh, K Abinav and Suresh Krishna

October 28, 2022

Enhanced Product Recommendation System for e-commerce Using Machine Learning

[1] R.M.Rani [2] Sourav Kumar Singh [3]Abinav K [4] Suresh Krishna S
[1, 2, 3, 4] Department of Information Technology, SRM Institute of Science and Technology, Ramapuram, Chennai,
E-mail: 1ranir@srmist.edu.in 2ss3617@srmist.edu.in 3ak5886@srmist.edu.in 4sk7280@srmist.edu.in

ABSTRACT

In the modern information technology age, finding users favorite product in large application databases becomes a serious issue for Recommendation developers. Recommendation means providing relevant suggestions to the user as per his/her interest and their needs. In this proposed Model we combined the corrwith() method which computes the Pearson correlation coefficients with our proposed collaborative filtering method which uses SVD++ that will help us in improving accuracy and targeting all sorts of users and recommends products supported ratings, previous purchase history, products on sale and recently viewed products etc. The experiments are performed on Amazon products dataset which consists the meta data of product reviews, feature, ratings, sales rank and similar products etc. The results of recommendations with this approach not only provide recommendations for specific products, but also provide recommendations on categories like groceries and gourmet foods. When a user selects a product, the model suggests related products that other users have also purchased with that product according to ratings, reviews and user's purchase history. Thus, the recommendations are going to be more varied and in line with user interest. This model paper also discusses about various shortcomings in the current recommendation techniques and suggests potential solutions that that could enhance the existing recommendation systems used by e-commerce websites.

Keywords: Recommender system, e-commerce, collaborative filtering, User reviews, User Ratings

1.INTRODUCTION

The majority of firms are working to develop some form of mechanism that is nothing more than a recommendation system in order to improve the customer experience and increase product sales. Therefore, recommender system appears to complete this task. A recommender system is required to boost the market's effectiveness or the selling of the goods. For making the best decisions a recommender system needs lots of information and the data supplied to the recommender system must be of a consistent in nature. The Recommendation system in existing models operates in two steps: first, it analyses the user's item search and basic user interests and second, it looks for a similar set of products that the user could find interesting. This in turn resulted in better product selections.

This proposed model used the customer data to make a recommendation using either popularity based filtering or collaborative filtering or sometime uses the hybrid recommendation techniques using collaborative filtering and corrwith() method which calculates the Pearson correlation coefficients depending on which one is best for the situation.

Collaborative filtering is a system for gathering information from multiple agents, customer point of view and data sources such as ratings, user likes, and so on. On the other hand, content filtering which uses Characteristics Information about the items/products (keywords, categorization) as well as the user's preferences, profile, etc. There is one more recommendation system called hybrid system which is combination of both combination of any two or more than two methods to prevent the issues that arise from using just one type of system.

In our proposed model, it has been to focused on grocery and gourmet products by using hybrid recommendation system which focused on combination of collaborative filtering and corrwith() method which is used to calculate the Pearson correlation coefficients and recommend products based on the product ratings, reviews and recent viewed products.

2.RELATED SURVEY

A product recommendation system is filtering system that helps to predict and show items that a user would like to purchase. In this, the system which is already existed is just like the most cutting-edge methods for recommendations, including collaborative filtering, popularity-based filtering, hybrid approaches, and social network-based approaches. In content-based

filtering there's an absence of novelty and variety. There is more to recommendations than relevance. And also, the disadvantage of collaborative filtering is it's unable to spot synonyms. There is no human interaction when predicting customer purchase behaviour. The Hybrid Recommendation System (HRS) clearly outperforms other modern algorithms in terms of accurate prediction of customer attitude towards recommending a product for a specific user. The problem with this model is that the data and application are defined interoperability.

3.Enhanced Product Recommendation System Architecture

In our proposed system, we combined the corrwith() method which calculates the Pearson correlation coefficients with our proposed collaborative filtering method which uses SVD++ which will help us in improving accuracy and targeting all sorts of users and recommend products supported ratings, previous purchase history, and recently viewed products etc. In Our proposed model the Amazon dataset is chosen and operations are performed on Grocery and Gourmet category where the system recommends the product based on similar purchased history. If any user chooses one product, it will recommend similar products which were bought by other users with good rating and based on customers previous purchase history. Thus, the recommendations are going to be more in line with user interest and helps in higher revenue generation for the e-commerce website. Fig 1. Shows the architecture of the proposed model and the proposed model used Amazon product dataset which contains product reviews and ratings and meta data about all products in form of text and images. It takes different ratings, reviews and more meta data on various products from users in Grocery and gourmets foods category and perform data preprocessing and cleansing after that we have used popularity-based recommendation, collaborative-based filtering approach and corrwith() method together with collaborative-based filtering approach and trained our dataset to recommend user the most effective recommended products list which are of similar interest to the targeted user.

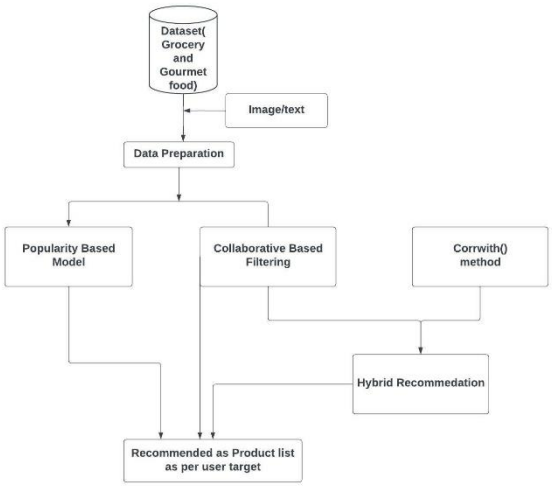

Fig 1: Enhanced Product Recommendation System

A. Popularity-Based Recommender Model

It performs by recommending highly rated products that are viewed and purchased by most of the users. It is not a personalized recommendation. As shown in fig 2, The implementation of Popularity-Based Filtering is very simple we just need to sort our products according to user ratings and show the top-rated products from that list.

- For rating of products, first create a metric score
- Calculate score for every product
- Sort the scores from highest to lowest and recommend the best rated product to the users. we used the IMDB's weighted rating formula to score the products, as follows:

weighted rating (WR) = (Nv / (Nv + Mv)) * M + (Mv / (Nv + Mv)) * MR

- M is the mean rating of a product.
- Nv is the count of votes for any product.
- Mv is the minimum votes required for a product be added in the popular products list.
- MR is the mean rating of products across the whole dataset.

We already have M or Mean Rating for each product and Mv or the Number of Rating. So, we only need to calculate MR (the mean rating across the whole dataset).

Fig 2: Popularity Based Recommendation Model

B. Collaborative filtering Recommender Model

Collaborative filtering uses similarities between users and products/items simultaneously to provide recommendations in order to overcome the drawbacks of popularity-based filtering systems. As shown in fig 2. Collaborative filtering is a system for gathering information from multiple agents, points of view, data sources such as ratings, user likes, and so on. It also considers other user's similar interests and behaviors towards a similar product when offering a product to target customer.

In order to implement Collaborative Filtering Recommender Model, this model used the best possible recommendations from people having similar interest. In other words, it makes predictions about how someone would rate an item based on past item ratings made by similar people. As shown in fig 3. Collaborative filtering recommender model recommends the set of products based on user-item interaction matrix which, represents the unique user preferences for items in a group, for recommender systems and are generally used in for collaborative filtering (CF) techniques

User-Item Interaction Matrix

User / Item	Item 1	Item ...	Item m
User 1	3		1
User ...	2		
User n		5	

Fig 3: User-Item Interaction Matrix

Our Model uses the enhanced Singular Value Decomposition (SVD++). This technique will customize our recommendation according to the same user group based on the above user-item interaction matrix and minimize Root Mean Square Error (RMSE) that is calculated by Kfold Cross Validation and give great recommendations.

Fig 4: Collaborative filtering Recommender Model

C. Hybrid Recommender Model

Each methods have its own strengths and weaknesses. To offer a better recommendation, our Proposed model combines the corrwith() method which is used to calculate the Pearson correlation coefficients together with the collaborative filtering. This concept leads towards the development of another recommendation model, which is Hybrid Recommender Model. As mentioned in fig 4. our proposed model used both the combination of corrwith() method which computes the Pearson to achieve more better and accurate results. In this system the combination of both is implemented to overcome most of the weakness of both the systems which will result in better accuracy and precision. The main process of the system are as follows:

- Step 1: Give UserID and ProductID as input
- Step 2: Create a pivot table which contains userIds as rows and productIds as columns
- Step 3: As Ouput we will get Similar products sorted based upon the expected score or ratings by a particular user.

Fig 5: Hybrid Recommender Model

4.RESULT & DISCUSSION

Our model is evaluated against Mean Square Error (MSE) and Root Mean Square Error (RMSE) to evaluate predictions. RMSE score is used to find the root mean square error between the actual and predicted ratings (R_{ai} and R_{pi}).

$$RMSE = \sqrt{[\Sigma (R_{pi} - R_{ai})^2 / n]}$$

and RMSE Obtained of enhanced SVD or SVD++ was about **0.871**.

MAE score used to find the mean square error between the actual and predicted ratings.

MAE = $(\Sigma^{n}_{i=1} |xi-xt|) / n$

and MAE of enhanced SVD or SVD++ was about **0.644**.

The combination of corrwith() method which computed the Pearson correlation coefficients with our proposed collaborative filtering(SVD++) is used to build a hybrid recommender which takes userId and productId as input and can suggest up to 10 products that were similar to the input productId based on the estimated ratings that was internally calculated for the input userId. This Hybrid System took advantage of both Pearson method and our proposed Collaborative filtering (SVD++) and therefore made reliable predictions.

TABLE 1: Performance metrics of all Algorithms

	Model	RMSE	MAE
1	**SVD**	**0.8746375**	**0.645896**
2	**SVD++**	**0.871489**	**0.644017**
3	**KNNBasic**	**0.979768**	**0.696134**
4	**KNNWithZScore**	**0.934450**	**0.650497**
5	**KNNWithMeans User-User**	**0.917915**	**0.644701**
6	**KNNWithMeans Item-Item**	**0.920770**	**0.647257**

Data shown in table 1 is also described below in figure 6. It is demonstrated that the performance of SVD++ is better than all other models.

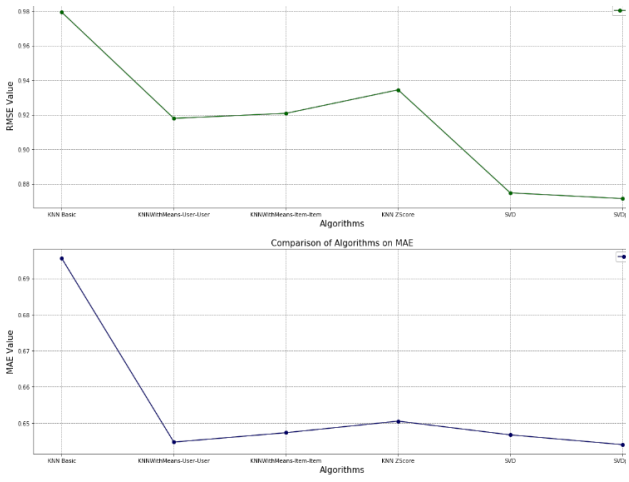

Fig 6: Comparison of Algorithms

Figure 6 demonstrate the accuracy of the proposed and reported methods using RSME and MAE value in percentage It is clear that the proposed SVD++ algorithm performs best in terms of accuracy and has the best RMSE value. with parameters {'n_epochs': 25, 'lr_all': 0.01, 'reg_all': 0.4} and after that SVD is performing second best RMSE value with parameters {'n_epochs': 20, 'lr_all': 0.005, 'reg_all': 0.2}. SVD++ is also giving the best MAE value and KNNWithMeans is giving second best MAE value. SVD++ is having the best RMSE in Matrix Factorization Based Algorithms also.

5.CONCLUSION

The concept of corrwith() method which computed the Pearson correlation coefficients with combination of our proposed collaborative filtering method(SVD++) to build an Hybrid Recommender Model which gives the better results as compared to other existing models. This Hybrid System took advantage of both Pearson method and Collaborative filtering and therefore made reliable predictions. hybrid recommender took userId and productId as input and suggested up to 10 products that were similar to the input productId based on the estimated ratings that was internally calculated for the input userId. This technique can help the users to save time and also improve the site's average order value which in turn will result in higher revenue generation for the e-commerce website.

REFERENCES

[1] Yi Cai, Ho-fung Leung, Qing Li, Huaqing Min, Jie Tang, Juanzi Li,"Typicality-Based Collaborative Filtering Recommendation", IEEE Transactions on Knowledge & Data Engineering, vol.26, no. 3, pp. 766-779, March 2014, doi:10.1109/TKDE.2013.7.

[2] S. Varma, M. Jain, D. Sharma and A. Beniwal, "Refined and diversified query suggestion with latent semantic personalization," 2015 IEEE UP Section Conference on Electrical Computer and Electronics (UPCON), Allahabad, 2015, pp.1-6. doi: 10.1109/UPCON.2015.7456725

[3] C.N. Ziegler, S. M. McNee, J. A. Konstan, and G. Lausen. “Improving recommendation lists through topic diversification”, In Proceedings of the 14th international conference on World Wide Web, WWW '05, , New York, NY, USA, pp. 22-32,2005

[4] Y. Wang, S. C.-F. Chan, and G. Ngai, “Applicability of demographic recommender system to tourist attractions: A case study on trip advisor,” in Proceedings of the 2012 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology-Volume 03, pp 97–101, 2012

[5] J. Wei, J. He, K. Chen, Y. Zhou, and Z. Tang, “Collaborative filtering and deep learning-based recommendation system for cold start items,” Expert Syst. Appl., vol. 69, pp. 29–39, 2017

[6] M. Balabanovi, Y. Shoham, Fab: content-based, collaborative recommendation. Commun.

ACM,40(3):66-72,1997.
[7] H.J Ahn, A new similarity measure for collaborative filtering to alleviate the new user cold-starting problem, Inf. Sci. 178 (1) (2008) 37–51
[8] X. Su, T.M. Khoshgoftaar, A survey of collaborative filtering techniques, Adv. Artif. Intell. 2009 (2009) 1–19.
[9] Karim, J., 2014, Hybrid Systems for Personalized Recommendations, Research Challenges in Information Science (RCIS), 2014 IEEE Eighth International Conference, May 2014
[10] P.-Y. Chen, et al., "The impact of online recommendations and consumer feedback on sales," ICIS 2004 Proceedings, p. 58, 2004
[11] G. Shani and A. Gunawardana, "Evaluating recommendation systems," in Recommender systems handbook, ed: Springer, pp. 257- 297, 2011.
[12] G. Shani and A. Gunawardana, "Evaluating recommendation systems," in Recommender systems handbook, ed: Springer, pp. 257- 297, 2011.