International Journal of Scientific Research in Science, Engineering and Technology
Print ISSN: 2395-1990 | Online ISSN : 2394-4099 (www.ijsrset.com)
doi : https://doi.org/10.32628/IJSRSET122935

Enhancing E-Commerce Applications with Machine Learning Recommendation Systems

Rafey Ahmed Farooqi, Surabhi Kesarwani, Mohd Shakeeb, Nitin Sharma, Ishita Bhatnagar
Department of Computer Science and Engineering, Babu Banarasi Das Institute of Technology and Management, Lucknow, India

Article Info
Volume 9, Issue 3
Page Number : 85-90
Publication Issue :
May-June-2022
Article History
Accepted : 01 May 2022
Published: 13 May 2022

ABSTRACT

In today's times everything has moved to a digital platform. Even commerce has moved to a digital mode with people now preferring to buy things online rather than going to a physical store. Recommendation Systems are used in such platforms to help users. Recommendation System is one of the most popular application of Machine Learning with various techniques and algorithms to implement it. We have researched these algorithms and have presented an analysis by taking various factors into consideration.

Keywords— E-Commerce Websites, Database Management, Classifiers, Machine Learning, Recommendation Systems, Content-Based, Collaborative filtering, Hybrid.

I. INTRODUCTION

E-commerce is fast gaining ground as an accepted and used business model. More and more business houses are implementing web sites providing functionality for performing commercial transactions over the web. It is reasonable to say that the process of shopping on the web is becoming commonplace. The objective of this paper is to enhance an e-commerce application where any product (such as books, CDs, computers, mobile phones, electronic items, and home appliances) can be bought from the comfort of home through the Internet using recommendation systems which are implemented using Machine Learning algorithms

An e-commerce is a virtual store on the Internet where customers can browse the catalogue and select products of interest. The selected items may be collected in a shopping cart. At checkout time, the items in the shopping cart will be presented as an order. At that time, more information will be needed to complete the transaction. Usually, the customer will be asked to fill or select a billing address, a shipping address, a shipping option, and payment information such as a credit card number. An email notification is sent to the customer as soon as the order is placed

Generally e-commerce applications have thousands of products available in their catalogue. Amazon has reportedly over a million products available for customers to buy. Unlike a physical store which is limited by the storage area available, online stores have no such limitation and can have innumerable products from various vendors.

However, this can lead to navigation issues as users can have trouble finding the exact product that they are looking for. It is imperative that users find the product if it is available and as quickly as possible as users do not wait and sift through thousands of products to find what they want. Unlike a real store , online applications do not have an employee who can guide the user to find the products that they need and can possibly buy.

Recommendation Systems can be used to solve this issue. They can help the users find the products that they require and even recommend products that they might like. Recommendation systems increase the chances of the user adding the products to cart and actually purchasing them.

Currently, the majority of online services use recommendation systems to enhance user experience. Amazon uses them to recommend products. Even entertainment platforms like Netflix and Spotify use them to curate personalized lists for each of their users, recommending movies and songs that they might like based on their activity on their application. Social Media platforms also use them to recommend other users who are similar to a user.

In order to use a recommendation system in an e-commerce platform a lot of data would be required. Specific formatting of said data is done with various pre-processing techniques as only useful data is kept in a data frame. E-commerce platforms therefore, store a lot of user data which helps in making their platforms perform better overall.

Specific data on the products is also kept to find the right product for each user. Data is required on each item to properly categorize and classify them. Moreover, the user's opinion or rating on items is also required for such systems to perform adequately. Without user-ratings all systems fail as there is no data of what the user likes or dislikes. Therefore data collection is also a major requirement in such systems.

II. RELATED WORK

Now that we have understood what recommendation systems are and how are they beneficial, it is time to understand how to implement them.

Implementing a recommendation system can be a little tricky as there are various things to factor in. Understanding what the user likes from their activity. Each user would be different in their own way.

Implementing this would require a lot of calculation. Machine Learning can be used to create a recommendation system that suites the purpose. Machine Learning is a technology which allows computers to "learn" from past experience like human beings do. Machine Learning algorithms use past results to predict new values. There are a multitude of Machine Learning algorithms that can be utilized to create suitable recommendation systems. Let's take a look at the various methods that can be used to create Recommendation Systems

III. METHODS

There are broadly three methods that are widely used to create recommendation systems. All of them are Machine Learning and data analysis based.

I. Collaborative-Filtering

As the name suggests Collaborative Filtering algorithms predict a user's preferences with the help of other similar users. They filter data for information by recognizing patterns of similar behaviour among various users. The algorithm works on the assumption that if user X and Y share an opinion on product A then it is likely that they will have the same opinion on product B.

Methodology – Collaborative filtering can be reduced into two basic steps:

• Look for other users who share similar opinion with the current user.
• Use the information to calculate a prediction for current user.
K – Nearest Neighbour algorithm is generally used for this.

Types – The method can be further classified into two types based on the approach taken.
• Memory Based – It is also called as Neighbourhood collaborative filtering. It makes use of rating data of users and finds similarities between them. It its also further classified into item-based and user-based filtering. User based uses users to predict while item-based finds similarities between items to predict.
• Model Based - Model based systems make use of various machine learning and data analysis models to predict user's preferences. Models require feature extraction from user's data to use as input. Model based approach uses decision trees, Bayesian networks, Clustering models, rule – based approaches and many more.

Collaborative filtering has a lot of advantages. Mainly it is simple to implement. It captures subtleties in data and does not need a firm understanding of it either. They provide high coverage.

There are some disadvantages to using Collaborative filtering as well. The major disadvantage is a problem called the 'Cold-start' problem. The method is not useful for recommending new items or new users. This is because the model works by finding similarities among users and without ant interaction it cannot find similarities.

II. Content-based Filtering

Content based systems predicts user's preferences based on their profile, i.e. find similarities between items that the user has liked previously and predict their preferences on unseen items based on the history. They rely on rating provided by user on items to recommend. The algorithm works on the assumption that if user X likes a product A then it it is likely that they will like product B which is similar to product A.

Requirements: In order for content-based methods to work there are a couple of prerequisites. First is an item data source which is essentially a data-source of all the items present. There should also be ample information about all items to find similarities among them. However, for the information to be useful it all should be in a similar format. Tags or keywords, that accurately describe an item, can be used for this purpose. These tags or keywords can then be further vectorized to be able to be mapped or plotted on a graph. Similar items are then found by measuring cosine distance on the graph.

Secondly, a user data source is also required which stores ratings provided by user on items. These rating help in determining positive feedback on items and finding items that are similar to it. The more user-feedback is available the better the model will work in recommending items.

The one major advantage of using Content-based filtering is that it works really well even if insufficient data is available. It leverages the ratings provided by the user. So even if only one rating is available it will be available to predict much better than collaborating-filtering based systems.

There are some flaws in the model as well. The main one is that the recommendation given are too similar to items user has already consumed. There is a lack of depth in the variety of items. If a user has not rated a specific category of items then that category will never be recommended. There is no diversity in the recommendations. Quite like collaborative-filtering, content-based filtering also fails if there is no data

available. It relies heavily on feedback provided by user. So if there is no feedback, the model would not be able to predict anything.

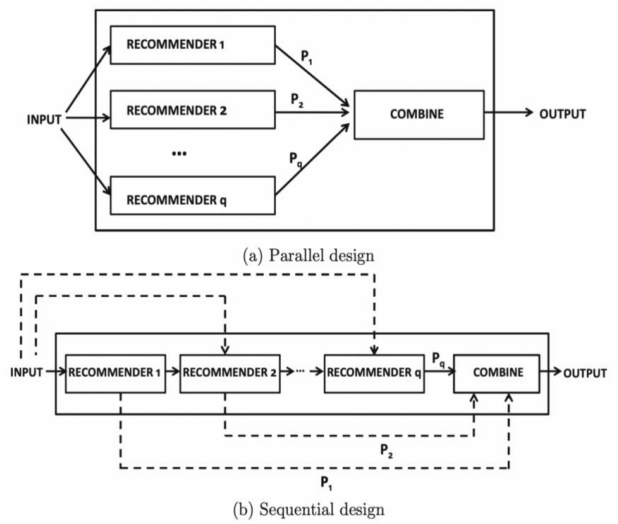
III. Hybrid Systems

There are many methods of implementing a recommendation systems. There are pros and cons of using each system. When used alone these methods can be restrictive and shallow. To counteract the flaws a hybrid system is designed which uses different methods and data based on availability and need. They generally perform better than the methods which are used alone and give out accurate predictions.

Hybrid recommendation systems are broadly classified into two types based on their designs. They are parallel and sequential or series system.

• Parallel – The parallel system provides input to various recommendation systems where each then provide an output. The systems collects the output and combines them to generate a single output .

• Sequential - The sequential systems provide input to different recommendation systems in a serial order.



There are various advantages of using hybrid systems. They combine various models to counteract the weaknesses in each model and therefore perform better. They give better and more personalized predictions. This is the reason that they are the most widely used systems.

However, hybrid systems are computationally expensive and require a lot of resources to run well.

IV. CONDUCTING THE REVIEW

Now we will conduct a review of the methods mentioned above and provide a comparison between them.

In order to compare them we will test each method on a dataset. The data-set used will be from a set of products and user-data of ratings on the products.

In order to use the data we will first perform some preprocessing on it to make it usable for our purpose.

Feature extraction done by choosing keywords from relevant columns and joining them in 'tags' column. Vectorization of the column is done using bag-of-words algorithm and finally each item is assigned a vector and plotted on a graph.

For testing purposes we will randomly convert 20 percent of the user's rated items from the list into unrated products.

Accuracy will be decided on the basis of the products recommended. Out of all products recommended the ones present in the converted list will be considered as accurate recommendations.

Precision will be calculated by the following formula

$$\textbf{Precision} = (\Sigma C / \Sigma \textbf{Total}) * 100$$

where c is the correct recommendations and total signifies the total recommendations made by the system.

Table 1: Precision of different methods

METHOD	PRECISION
Collaborative	37.38
Content-based	37.32
Hybrid	42.06

As it is clear from the result , Hybrid systems outperform collaborative and content-based systems. When the data-set is small content-based systems perform better than collaborative but as the data becomes large collaborative perform better and better. However, hybrid systems require more computational resources than their counterparts so it all comes down to requirements. For e-commerce applications hybrid systems would work best as e-commerce applications generally have a detailed data-base of both products and users.

V. CONCLUSIONS AND FUTURE WORK

All the methods mentioned in the paper are tested and compared with each other. The research done shows the strength and weakness of each method. The research done clearly infers that Hybrid Systems work best among all the methods in terms of precision. Therefore, for an e-commerce platform a hybrid system would perform well.

With this paper we have achieved promising results from all the methods. However, there is still room for improvement among all these systems as in the real world a more precise model would be required to use. Therefore, it would be worthwhile to examine these and more methods in depth to improve our prediction precision.

VI. ACKNOWLEDGEMENT

We are highly grateful to our college Babu Banarasi Das Institute of Technology & Management for providing us robust environment to dive deep into this project and also thankful to our management, mentors and faculties for their guidance and support.

VII. REFERENCES

[1]. Hernandez, Sergio, et al. "Analysis of users' behavior in structured e-commerce websites." IEEE Access 5 (2017): 11941-11958.

[2]. Parallel and Sequential Recommendation System Architecture. Image provided by C.C. Aggarwal, Recommender Systems: The Textbook

[3]. C.C Aggarwal "Recommender Systems: The Textbook" ISBN: 978-3-319-29659-3

[4]. Sharma, Lalita, and Anju Gera. "A survey of recommendation system: Research challenges." International Journal of Engineering Trends and Technology (IJETT) 4.5 (2013): 1989- 1992.

[5]. Isinkaye, Folasade Olubusola, Yetunde O. Folajimi, and Bolande Adefowoke Ojokoh. "Recommendation systems: Principles, methods and evaluation." Egyptian informatics journal 16.3 (2015): 261-273.

[6]. Su, Xiaoyuan, and Taghi M. Khoshgoftaar. "A survey of collaborative filtering techniques." Advances in artificial intelligence 2009 (2009).

[7]. Van Meteren, Robin, and Maarten Van Someren. "Using content-based filtering for recommendation." Proceedings of the machine learning in the new information age: MLnet/ECML2000 workshop. Vol. 30. 2000.

[8]. www.towardsdatascience.com

Cite this article as :

Rafey Ahmed Farooqi, Surabhi Kesarwani, Mohd Shakeeb, Nitin Sharma, Ishita Bhatnagar, " Enhancing E-Commerce Applications with Machine Learning Recommendation Systems, International Journal of Scientific Research in Science, Engineering and Technology(IJSRSET), Print ISSN : 2395-1990, Online ISSN : 2394-4099, Volume 9, Issue 3, pp.85-90, May-June-2022. Available at doi : https://doi.org/10.32628/IJSRSET122935
Journal URL : https://ijsrset.com/IJSRSET122935