|  |
| --- |
|  |

|  |
| --- |
| TECHNICAL REPORT  Electrical & Computer Engineering & Computer Science (ECECS) |

|  |  |
| --- | --- |
| SPRING 2023 |  |



Contents

[Project Name 2](#_Toc96341549)

[Executive Summary 2](#_Toc96341550)

[Technical Report 3](#_Toc96341551)

[Highlights of Project 3](file:///C:\Users\ardia\Desktop\TEACHING\SPRING%2022\DSCI6002%20Data%20Science\Projects\Technical%20report%20Template.docx#_Toc96341552)

[Submitted on: 3](file:///C:\Users\ardia\Desktop\TEACHING\SPRING%2022\DSCI6002%20Data%20Science\Projects\Technical%20report%20Template.docx#_Toc96341553)

[Abstract 4](#_Toc96341554)

[Methodology 5](#_Toc96341555)

[Results Section 5](#_Toc96341556)

[Discussion 6](#_Toc96341557)

[Conclusion 6](#_Toc96341558)

[Contributions/References 7](#_Toc96341559)

|  |
| --- |
| Product Recommendation System |

|  |
| --- |
| Executive Summary Amazon is a multinational technology company based in Seattle, Washington, that focuses on e-commerce, cloud computing, digital streaming, artificial intelligence and more. One of its most popular services is the e-commerce website ‘Amazon.com.’ One of the ways Amazon keeps its shoppers engaged is by way of its product recommendation system. The product recommendation system uses content and collaborative based filtering, looking at a shopper’s past purchases and purchases by shoppers who purchased similar items and then recommends items that the shopper may like based off of these two behaviors. Collaborative filtering enables the shopper to see items from other categories, while content-based filtering allows shoppers to see domain specific items.  Unlike content-based filtering, which suggest only similar products or domain specific products, collaborative filtering recommendation systems allows users to find similar products and jump (or switch) to other categories of products based on similarities with other users and products. In this case, we explore 3 different ways to use collaborative based filtering.   1. The first method suggests products that other users like to a specific user based on their rating of a specific product. Essentially, it finds users that rated a product similar to the specific user and returns the top products the other users liked. 2. The second method predicts what products a user may like based on past rating history and the rating history of other users. This method is not specific to the rating of one product like the first method. 3. The third method predicts products similar to a specific product by finding the nearest neighbors to that product. This method is good for users who may want to purchase products as bundles.   We evaluate our methods by:   * Testing the recommendation systems to see what products they return. * Understanding matrix factorization and how it effects collaborative filtering systems. * Selecting and tuning various recommendation algorithms to find the best performing algorithm. * Understanding estimated ratings vs. true ratings.   <https://github.com/UNH-Fall2022-DistributedSystems-G6/UNH-DS-Fall22> |
| person at a table writing in a notebook with people around  **Questions?**  **Team Members: Contact:**  Pavani Billapati [pbill1@unh.newhaven.edu](mailto:pbill1@unh.newhaven.edu)  Likhitha Gupta Thallapally [lthal1@unh.newhaven.edu](mailto:lthal1@unh.newhaven.edu)  Trilok Kumar Pidikiti tpidi1@unh.newhaven.edu  Harianth Kumar Kancharla hkanc1@unh.newhaven.edu |

|  |  |
| --- | --- |
| Amazon's website offers a plethora of different types of products ranging from automotive to sports to groceries to hardware tools to digital products and more. A shopper can purchase almost anything on Amazon.com while taking advantage of its recommendation system.  TECHNICAL REPORT  Recommendation systems are great at recommending similar products to shoppers, but they tend to be too concentrated, which does not let a shopper go outside their comfort zone or see something they may not have known they wanted. For example, watching a video on YouTube will start to flood a user’s browser with several similar videos and not enough diversity in content. As for the case with Amazon, we want to build a recommendation system that recommends a diverse selection of items or content to users as needed. |  |
| Data have grown up uncontrollably due to which  ABSTRACT  there are large number of products that are listed on e-commerce  website today. In a way through which users can quickly find a  favorite item from big resources, the user requires such technology  by which it can be automated so recommendation systems were  introduced. Recommendation systems are mainly used by  companies like e-commerce to help the user discover items they  have to found out by them and increase the sales of the company.  A recommender engine can recommend user interest products. In  this paper the discussion various different recommendation system,  evaluation techniques and also the challenges and problem in the  system are discussed. Also using Amazon electronics data building  of popularity-based recommender engine and recommender engine  using collaborative filtering that is based on singular value  decomposition is discussed. The goal of the model is to  recommend users 5 top products to the user and performance of  each model is evaluated  Data have grown up uncontrollably due to which  there are large number of products that are listed on e-commerce  website today. In a way through which users can quickly find a  favorite item from big resources, the user requires such technology  by which it can be automated so recommendation systems were  introduced. Recommendation systems are mainly used by  companies like e-commerce to help the user discover items they  have to found out by them and increase the sales of the company.  A recommender engine can recommend user interest products. In  this paper the discussion various different recommendation system,  evaluation techniques and also the challenges and problem in the  system are discussed. Also using Amazon electronics data building  of popularity-based recommender engine and recommender engine  using collaborative filtering that is based on singular value  decomposition is discussed. The goal of the model is to  recommend users 5 top products to the user and performance of  each model is evaluated  Data have grown up uncontrollably due to which there are large number of products that are listed on e-commerce website today. In a way through which users can quickly find a favorite item from big resources, the user requires such technology by which it can be automated so recommendation systems were introduced. Recommendation systems are mainly used by companies like e-commerce to help the user discover items they have to find out and increase the sales of the company. A recommender engine can recommend user interest products. In this report, various recommendation systems, evaluation techniques and also the challenges and problem in the system are discussed. Also using Amazon electronics data building of popularity-based recommender engine and recommender engine using collaborative filtering that is based on singular value decomposition is discussed. The goal of the model is to recommend users 5 top products to the user and the performance is evaluated. |

Introductory Section

Recommendation systems had changed the way of

interaction between user and websites and are increasingly

important today. Recommender system enhance accesses

and take charge to recommend appropriate items to users by

in view of the users raised choices and objective behaviors.

Being an online advertisement or e-commerce websites,

recommender system cannot be avoided today. Every other

company is trying to use the power of recommendation

system. These systems have huge application in different

sectors that are education, economy and researches, like

much other work[1]–[5]. The rate of digital information is

increasing rapidly due to rapid growth of information

technology. The recommendation system has attained great

results solving the problem of data overloaded. There is a

very large number of products that are listed today on e-

commerce websites like Flipkart, Amazon. Recommender

system helps user when they face huge amount of choices[6].

There are almost more than 30 million products present on

Flipkart today. Due to which it has become tough for

customers to choose their desired choice. The recommender

system deals with many data present by filtering the most

Recommendation systems have changed the way of interaction between user and websites and are increasingly important today. Recommender system enhances access and take charge to recommend appropriate items to users by view of the users raised choices and objective behaviors. Being an online advertisement or e-commerce websites, recommender system cannot be avoided today. Every other company is trying to use the power of recommendation system. These systems have huge applications in different sectors that are education, economy and researches, like much other work. The rate of digital information is increasing rapidly due to rapid growth of information technology. The recommendation system has attained great results solving the problem of data overload. There is a very large number of products that are listed today on e-commerce websites like Flipkart, Amazon. Recommender system helps user when they face huge number of choices. There are almost more than 30 million products present on Flipkart today. Due to which it has become tough for customers to choose their desired choice. The recommender system deals with many data present by filtering the most important information based on historical data of a user which takes care of the user’s preference and interest. Recommender system can predict whether a particular user would prefer an item or not based on the user’s profile. It can predict whether a user would prefer a product or not based on the user’s historical data. Recommender systems are beneficial to both service providers and users. The quality and decision-making process has also improved through these kinds of systems.

Recommender systems result in mainly things stated below:

* Benefits users in finding items of their interest.
* Help item providers deliver their items to the right user.
* Identity products that are most relevant to the user.
* Personalized content.
* Help website improves user engagement.

## 

## Methodology

**Data Collection and Preprocessing**:

The first step in building a recommendation system is to collect and preprocess the data. In this case, we will use the Amazon product reviews dataset, which contains customer ratings and reviews for various products sold on Amazon. The dataset is available on Amazon's website and can be downloaded in a CSV format.

Once the dataset is downloaded, we need to preprocess the data to make it suitable for our recommendation system. We will start by removing any irrelevant columns and duplicate entries from the dataset. Next, we will calculate the average rating for each product and store it in a separate column. This will help us identify the most popular products based on their average rating.

**Model Training and Testing**:

With the preprocessed data, we can now train our recommendation system model. There are several algorithms we can use to build a recommendation system, such as collaborative filtering, content-based filtering, and matrix factorization. For this project, we used collaborative filtering algorithm.

**Collaborative filtering:**

Collaborative filtering is a technique that recommends products to a user based on the preferences of other users who have similar tastes. To implement this algorithm, we need to create a user-item matrix that stores the ratings given by each user for each product. We will then split the dataset into training and testing sets and use the training set to train the model. Once the model is trained, we will use the testing set to evaluate its performance.

Collaborative filtering is a popular technique used in recommendation systems to suggest products or services to users based on their preferences and the preferences of similar users. In collaborative filtering, the system identifies similarities among users or items and uses this information to make personalized recommendations.

There are two main types of collaborative filtering: user-based and item-based. In user-based collaborative filtering, the system identifies users who have similar preferences to the target user and recommends products that these similar users have liked. In item-based collaborative filtering, the system identifies items that are similar to the items that the target user has liked and recommends these similar items.

Collaborative filtering is based on the assumption that users who have similar preferences in the past are likely to have similar preferences in the future. However, this assumption may not always hold true, and the performance of collaborative filtering may be affected by the sparsity of the data or the presence of noisy or irrelevant data.

To address these issues, researchers have developed more advanced techniques such as matrix factorization and deep learning-based models, which can better capture the underlying patterns in the data and improve the performance of recommendation systems.

**SVD (singular value decomposition):**

The SVD (singular value decomposition) algorithm is a popular technique for building recommendation systems, including for Amazon's product recommendation system. The SVD algorithm is a matrix factorization technique that decomposes the user-item ratings matrix into three matrices: U, S, and V.

U is a matrix that represents the user features, where each row corresponds to a user and each column corresponds to a latent feature. S is a diagonal matrix that represents the strength of each latent feature, and V is a matrix that represents the item features, where each row corresponds to an item and each column corresponds to a latent feature.

The SVD algorithm learns the latent features from the user-item ratings matrix and uses them to predict the ratings for items that the user has not yet rated. The predicted ratings can then be used to generate personalized recommendations for the user.

To implement the SVD algorithm in Amazon's product recommendation system, we start by preprocessing the user-item ratings matrix to remove any missing values or outliers. We then decompose the matrix using the SVD algorithm to obtain the U, S, and V matrices. We can then use these matrices to predict the ratings for items that the user has not yet rated.

To evaluate the performance of the SVD algorithm, we can use metrics such as RMSE (root mean squared error) and MAE (mean absolute error). These metrics measure the difference between the predicted and actual ratings, with lower values indicating better performance.

Overall, the SVD algorithm is a powerful technique for building recommendation systems and has been successfully applied in various domains, including e-commerce, music, and movies. It can help Amazon's product recommendation system provide personalized and accurate recommendations to its users.

**Results Section**

By using SVD algorithm, we performed data slicing, visualizing selected featured by plotting their Bar graph, plotting the features to analyze the label and its frequency and trained the model. We further created pivot table, diagonal array in svd to make further conclusion.

Text

Description automatically generated

Text

Description automatically generated

## Graphical user interface, text, application Description automatically generated

## Graphical user interface, text, application Description automatically generated

Chart, bar chart

Description automatically generated

## Discussion

We considered various factors like 'user Id', 'product Id', 'ratings', 'timestamp’, and other factors to determine the top ‘n’ product recommendations. We cleaned the data by replacing ‘Nan’ values with 0 to get proper accuracy. While executing the project we omitted the columns which did not contribute much to the prediction accuracy.

## Conclusion

Unlike content-based filtering, which suggest only similar products or domain specific products, collaborative filtering recommendation systems allows users to find similar products and jump (or switch) to other categories of products based on similarities with other users and products.

This method suggests products that other users like to a specific user based on their rating of a specific product. Essentially, it finds users that rated a product similar to the specific user and returns the top products the other users liked. This method was used as an introduction to recommendation systems with a simple, built from scratch, recommender that did not incorporate actual training of models or prediction methods.

## Contributions/References

* <https://www.researchgate.net/publication/3419552_Linden_G_Smith_B_and_York_J_'Amazoncom_recommendations_item-to-item_collaborative_filtering'_Internet_Comput_IEEE_7>
* <https://www.semanticscholar.org/paper/A-Survey-on-Recommendation-System-Das-Sahoo/5e90a13c6b159dbc85772309f5301dc9f308f9d9>
* <https://en.wikipedia.org/wiki/Recommender_system>