G.H Raisoni College of Engineering & Management, Pune

(An Autonomous Institute Affiliated to SPPU)

Department: Computer Engineering

Subject: **MPTRW**

Topic: ONLINE GROCERY RECOMMENDATION SYSTEM.

Submitted To:

Prof. Gayatri Bedre

Submitted by:

1.Sitaram Mane. (TCOB98)

2.Sumit Biradar. (TCOB101)

3.Swapnil Rankhamb. (TCOB105)

4.Tejas Rathod. (TCOB107)

DECLARATION

We hereby declare that the project report entitled "**Online Grocery Recommendation System**" submitted by us to **G.H.RAISONI COLLEGE OF ENGINERRING AND MANAGEMNETY** in partial fulfillment for the award of degree of B.Tech in computer science and engineering is a record of bonafide project work carried out by us under the guidance of **Mrs. Gayatri Bedre.**

I further declare that the work reported in this project has not been submitted and will not be submitted, either in part or full, for the award of any other degree in this institute or any other institute or University.

Sitaram Mane	(TCOB98)
Sumit Biradar	(TCOB101)
Swapnil Rankhamb	(TCOB105)
Tejas Rathod	(TCOB107)

ACKNOWLEDGEMENT

The satisfaction that accompanies that the successful completion of any task would be incomplete without the mention of people whose ceaseless cooperation made it possible, whose constant guidance and encouragement crown all efforts with success.

We are grateful to our project guide **Mrs. Gayatri Bedre** for the guidance, inspiration and constructive suggestions that helpful us in the preparation of this project.

We also thank our colleagues who have helped in successful completion of the project.

Sitaram Mane (TCOB98)
Sumit Biradar (TCOB101)
Swapnil Rankhamb (TCOB105)
Tejas Rathod (TCOB107)

ABSTRACT

Now days there is a lot of trend of online shopping. Almost all the products which are used in day to day life are available online especially grocery. So there are many projects going on online grocery shopping systems. And most of them have recommendation systems in them. Recommendations are used for making the work of the customer more easy and fast. This reduces their valuable time and also the efforts. For this the recommendations given to the customer should be exact and should be fast. And most importantly they should none irritate the customer. These recommendations are mostly given based on their necessity can be predicted from the people have interest same as that customer. In our project of online grocery recommendation system, we are going to develop a recommendation system which will recommend the customer products of his interest and similarity. One more additional feature in our project is that the customer is recommended a special basket based on his profile and also purchase history to some extent.

Contents

Chapter 1: Introduction

1.1 Definition:

Collaborative filtering is a technique that can filter out items that a user might like on the basis of reactions by similar users. It works by searching a large group of people and finding a smaller set of users with tastes similar to a particular user.

1.2 Purpose:

The main objective of this e-commerce websites is to find out which products the customers might like to purchase based on his/her previous purchase history. Recommendation systems also allows preparing more relevant personalized offers.

1.3 project main idea:

The main idea of the project is to make a recommendation application to recommend grocery product for customers so the market could save its money spent advertising and customers save their time for searching for their desired products.

1.4 problem Statement:

Companies spent money on advertising and sometimes high portion of this money doesn't make an effect because these advertising go to the wrong customers that don't prefer these products or don't match their needs, so we want to make an application that would recommend products to the most desirable Customers that have high probability to buy these items.

1.5 Proposed solution:

Our solution is to make a recommendation application that recommend the grocery items to the customers according to their needs, so company can recommend the best products for their customer that have high probability to buy these products.

1.6 Objective:

1) The purpose of a recommender system is to suggest relative items to users.
2) Reduce transaction costs of finding and selecting items in an online shopping environment.
3) Main objective of these recommendation system is to growth in the business.

Chapter 2: Literature survey

Overfitting and underfitting come into picture when we create our statistical models. The models might be too biased to the training data and might not perform well on the test data set. This is called overfitting. Likewise, the models might not take into consideration all the variance present in the population and perform poorly on a test data set. This is called underfitting. A perfect balance needs to be achieved between these two, which leads to the concept of Bias-Variance trade off.

2.1 Domain of Study:

Python was the major technology used for the implementation of machine learning concepts the reason being that there are numerous inbuilt methods in the form of packaged libraries present in python. For implementation of this project different type of machine learning algorithm are used. Following are prominent libraries/tools we used in our project.

1]SCIPY SciPy: It is a free and open-source Python library used for scientific computing and technical computing. SciPy contains modules for optimization, linear algebra, integration, interpolation, special functions, FFT, signal and image processing, ODE solvers and other tasks common in science and engineering.

2]SCIKIT LEARN: Scikit-learn provides a range of supervised and unsupervised learning algorithms via a consistent interface in Python. It is licensed under a permissive simplified BSD license and is distributed under many Linux distributions, encouraging academic and commercial use.

3] NUMPY NumPy: It is a general-purpose array-processing package.It provides a highperformance multidimensional array object and tools for working with these arrays. It is the fundamental package for scientific computing with Python. Besides its obvious scientific uses, NumPy can also be used as an efficient multi-dimensional container of generic data.

4]JUPITER NOTEBOOK: The Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations, and narrative text. It includes data cleaning and transformation, numerical simulation, statistical modeling, data visualization, machine learning, and much more.

2.2 Literature Survey

Sr No	Author	Title	Source	Publish Date	Findings
1	Shraddha Gupta, Ankit Maithani	A Literature Review on Recommendation Systems	IRJET	Sep-2020	Recommendation system using collaborative filtering, Implement Collaborative based methods separately and then merge their predictions.
2	Sharon J. Moses, L.D. Dhinesh Babu	Buyagain Grocery Recommender Algorithm for Online Shopping of Grocery and Gourmet Foods	International Journal of Web Services	July-2019	Based on collaborative filtering where the slope one method was employed for predicting user behaviour based on other users behaviours
3	Yi-Jing Wu and WeiGuang Teng	An Enhanced Recommendation Scheme For Online Grocery Shopping	International Symposium on Consumer Electronics. IEEE.	Jan-2011	Common Recommendation Techniques, Proposed Scheme For Online Grocery Shopping,

4	Lamiyah Khattar, Dr.Geetika Munjal	Analysis of Online Grocery Recommendation Systems	2021 IEEE 11th International Conference on Cloud Computing, Data Science & Engineering	Jun-2011	Clustering of similar user data and similar item. Implementation of Collaborative filtering model steps.
5	Lee. Yunkyoung	RECOMMENDATION SYSTEM USING COLLABORATIVE FILTERING	Master's Projects. 439	Dec-2015	We Studied the process of data filtering in the collaborative system.
6	Sharon J. Moses, L.D.Dhinesh Babu	Buyagain Grocery Recommender Algorithm for Online Shopping of Grocery and Gourmet Foods	International Journal of Web Services Research	Sep-2018	We Studied algorithm for Online Grocery Recommendation System

Chapter 3: System Analysis

3.1 Introduction

Recommender systems that recommend items through consumer collaborations and are the most widely used and proven method of providing recommendations. There are two types: user-to-user collaborative filtering based on user-to-user similarity and item-to-item collaborative filtering based on item-to-item similarity.

3.2 Feasibility Study

In this Grocery Recommendation System model, we use a variety of algorithms such as regression and classification. Here we use Linear Regression, Logistic Regression and Lesso Regression Technique. continuous volatility. Used continuous output prediction and constant volume. Lesso Regression is usually a small square adjusted to reduce the total coefficient. Logistic regression helps us predict event opportunities by entering data. to get into work. Edition is the process of dividing a particular set of data into categories. And division algorithm divides/Separates data into different categories. For example in our database we have product name and it's purchase history so by using the Collaborative algorithm, it can split item based data. data such as Product id, product name, Purchase history. This will help us to predict or suggest the similar type of product to the another customer.

3.3 Operational Feasibility

The system working is quite easy to use and learn due to its simple but attractive interface. User requires no special training for operating the system. Technical performance include issues such as determining whether the system can provide the right information for the purchased product details, and whether the system can be organized so that it always delivers this information at the right place whenever customer going for shopping and on time using internet services.

3.4 Technical Feasibility

The technical requirement for the system is economic and it does not use any other additional Hardware and software. Technical evaluation must also access whether the existing systems can be upgraded to use the new technology and whether the organization has the expertise to use it.Install all upgrades framework into the .py package supported windows based application. This application depends on Jupyter Notebook,Google Colab,and Vs-code. Enter Grocery details report to excel sheet and save Excel file with .csv extension as gorcey.csv

Chapter 4: Methodology/Planning

4.1 Planning

A grocery recommendation system is a tool that is designed to provide suggestion to customers for what he would want to buy next.

1) Data Collection: First and the most important thing that you need is data. Data and customers and products on sale.
2) After collecting the data, it has to be cleaned and filtered in order to get a better model and better predictions.
3) Check for the rows and columns. Remove all the data that you don't need. Remove the fields that are incomplete. As this will result in wrong predictions.
4) There are different recommendation systems out there in the market and we use collaborative filtering Machine Learning algorithms.
5) When a user enters a website or installs the app for the first, the agency has no information about the user of what he will like or what he is trying to buy. There we have to apply a popularity-based strategy. Show the user all the trending products available. This will narrow down the choices of the user. The website will also get to know the location of the user, from which website he came from etc.
6) Recommendations depend on products which you like on thesebehaviour.
7) The collaborative method uses a nearest-neighbour algorithm to identify products.
8) Testing of data: Train/Test is a method to measure the accuracy of your model. Once your machine learning model is built (with your training data), you need unseen data to test your model. Test a machine learning program after it has been trained on an initial training dataset.
9) Output Analysis: Process of analysing the correct prediction and correct output.it is the process of study of all the output in the recommendation system.

4.2 Methodology:

Single machine learning classifier approach that has been used in all previous researches was also tested in this research. The whole data set collected in this research has been split into training (90 percent) and testing (10 percent) subsets and Artificial Neural Network, Support Vector Machine and Random Forest classifiers models werbuilt.

To recommend the product to customer following algorithms and Machine Learning libraries are used:

1)Seaborn Library

Seaborn is a Python data visualization library based on matplotlib. Seaborn is a library mostly used for statistical plotting in Python. It is built on top of Matplotlib and provides beautiful default styles and colour palettes to make statistical plots more attractive seaborn library aims to make a more attractive visualization of the central part of understanding and exploring data.

Seaborn offers the following functionalities:

1.Automatic estimation and plotting of linear regression plots.

2.Visualizing univariate and bivariate distribution.

3.Dataset oriented API to determine the relationship between variables.

2)Linear Regression

Linear Regression is a machine learning algorithm based on super\vised learning. It performs a regression task. Regression models a target prediction value based on independent variables. It is mostly used for finding out the relationship between variables and forecasting. Linear Regression is a supervised machine learning algorithm where the predicted output is continuous and has a constant slope Linear regression is a quiet and simple statistical regression method used for predictive analysis and shows the relationship between the continuous variables. Linear regression shows the linear relationship between the independent variable (X-axis) and the dependent variable (Y-axis), consequently called linear regression. If there is a single input variable (x), such linear regression is called simple linear regression.

4.3 The Dataset

To experiment with recommendation algorithms, you'll need data that contains a set of items and a set of users who have reacted to some of the items.

The reaction can be explicit (rating on a scale of 1 to 5, likes or dislikes) or implicit (viewing an item, adding it to a wish list, the time spent on an article).

While working with such data, you'll mostly see it in the form of a matrix consisting of the reactions given by a set of users to some items from a set of items. Each row would contain the ratings given by a user, and each column would contain the ratings received by an item. A matrix with five users and five items could look like this:

	i_1	i_2	i_3	i_4	i_5
u_1	5		4	1	
u_2		3		3	
u_3		2	4	4	1
u_4	4	4	5		
u_5	2	4		5	2

The matrix shows five users who have rated some of the items on a scale of 1 to 5. For example, the first user has given a rating 4 to the third item.

In most cases, the cells in the matrix are empty, as users only rate a few items. It's highly unlikely for every user to rate or react to every item available. A matrix with mostly empty cells is called sparse, and the opposite to that (a mostly filled matrix) is called dense.

Chapter 5: System Module

The system comprises of 2 major modules with their sub-modules as follows:

5.1 Admin:

- **Login:** Admin can login in his personal account using id and password.
- **Add Grocery:** Admin can add Groceries.
- **View User:** Admin can view all information about the user.

5.2 User:

- **Login:** User can login his account using id and password.
- **View Products:** User can view the products.
- **View Product on search:** User can view the product on basis of the searches.
- **View Recommends:** User will get the collaborative filtering of grocery.

Chapter 6: Requirement Specifications

6.1 Software and Hardware Requirements:

	Sr.No	Resource	Specification	Quantity
Software	1	Python	Version 3.93	1
	2	IDE	Jupyter Notebook	1
	3	Library	NumPy, Pandas, matplotlip, sciPy, scikit learn	1
Hardware	1	Processer	Core i5	1
	2	Hard Disk	160GB	1
	3	Memory	1GB RAM	1

6.2 Functional Requirements:

Grocery Recommendation System involves the following functions- -

Easily track purchase history details.

-To recommend the best item to customer.

Chapter 7: Architecture And Design

7.1 Block Diagram :

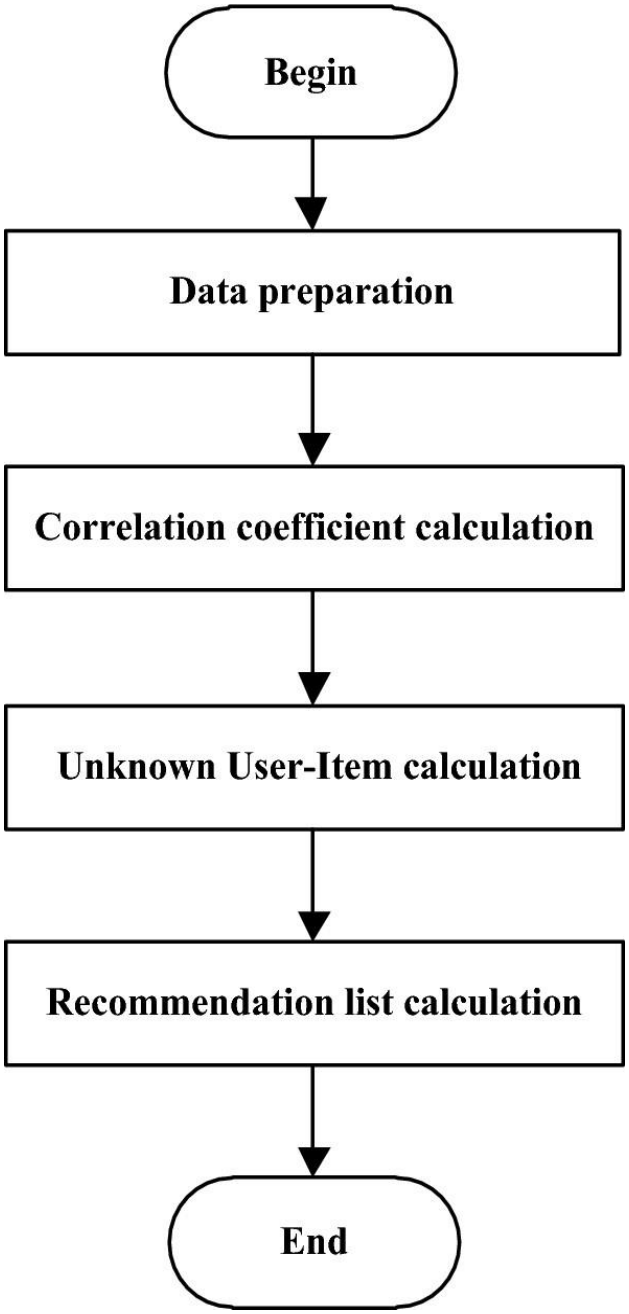


Figure 1: Block Diagram

7.2 Use Case Diagram:

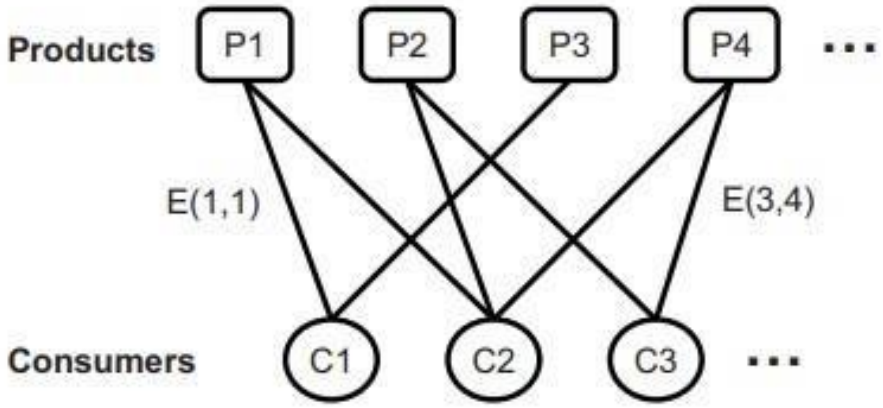


Figure 2: User Case Diagram (1)

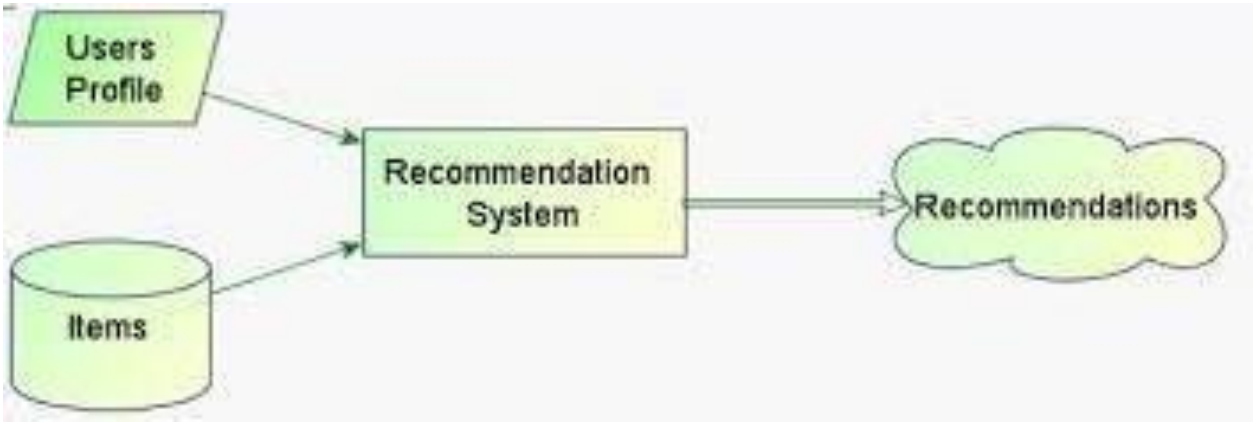


Figure 3: User Case Diagram (2)

7.3 Data Flow Diagram (DFD):

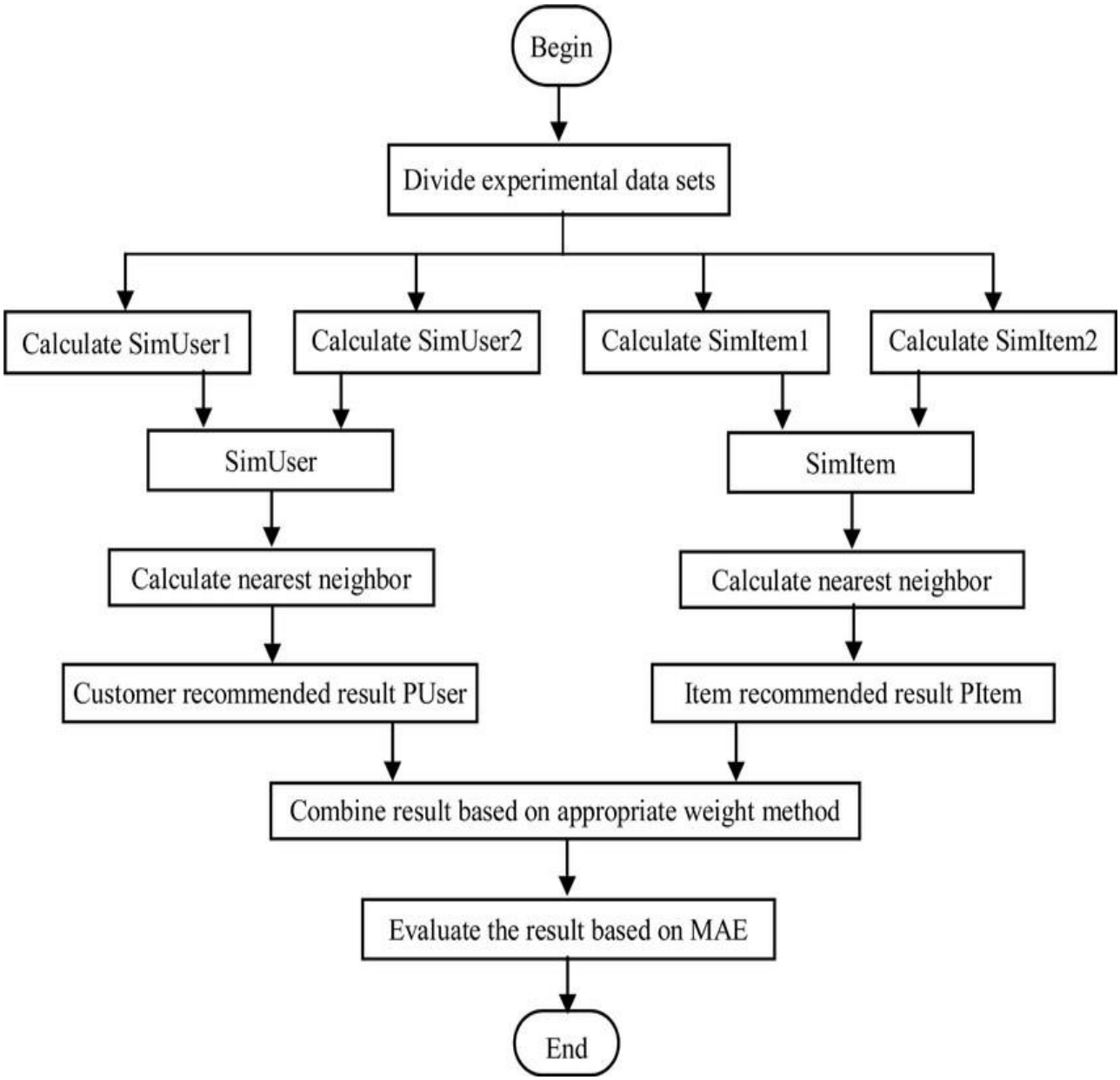


Figure 4: Data Flow Diagram

7.4 Activity Diagram:

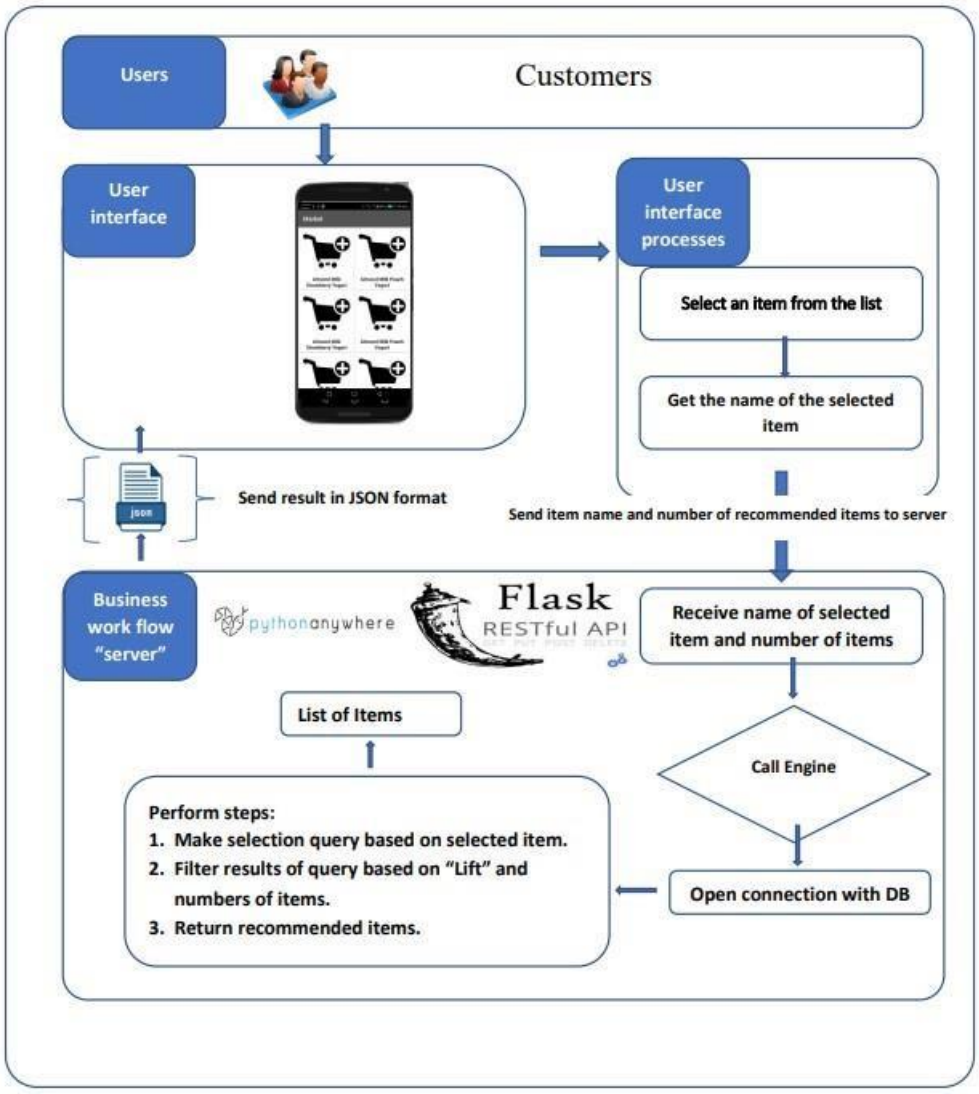


Figure 4: Activity Diagram

Chapter 8: Expected Outcomes

1) Input = Purchase history dataset with multiple parameters to train the model.
2) With the help of the machine learning algorithm the model predict the accurate result in the form of graphs. Machine learning algorithm are applied on each and every column to predict the result. The task of recognizing emotion, nonetheless, is an innately trouble, some issue, more so than numerous other PC vision undertakings.
3) The proposed prediction model has been evaluated on the test subset and model achieved overall accuracy of 87.38 percent. This proves that combination of multiple machine learning classifiers strengthens the classification performance overall.
4) Expected Output = Predict the next food on the bases of previous purchase history dataset and using the similarity of data.

ID	**Input**	**Expected Output**	**Actual output**	**Pass/Fail**
ID_1	Purchase history dataset	Dataset executed and imported well in the Jupyter notebook	Successfully imported the dataset	Pass
ID_2	Dataset	Dataset will be easily analyse and visualize for understand.	Data get analyse and classified using collaborative filtering	Pass
ID_3	Dataset Cleaning	Clean dataset and remove empty data	Clean the data	Pass
ID_4	Dataset	Splitting data	Data get split into multiple groups	Pass
ID_5	Data	Similar type of Dataset	Similar type of data get separated in groups	Pass

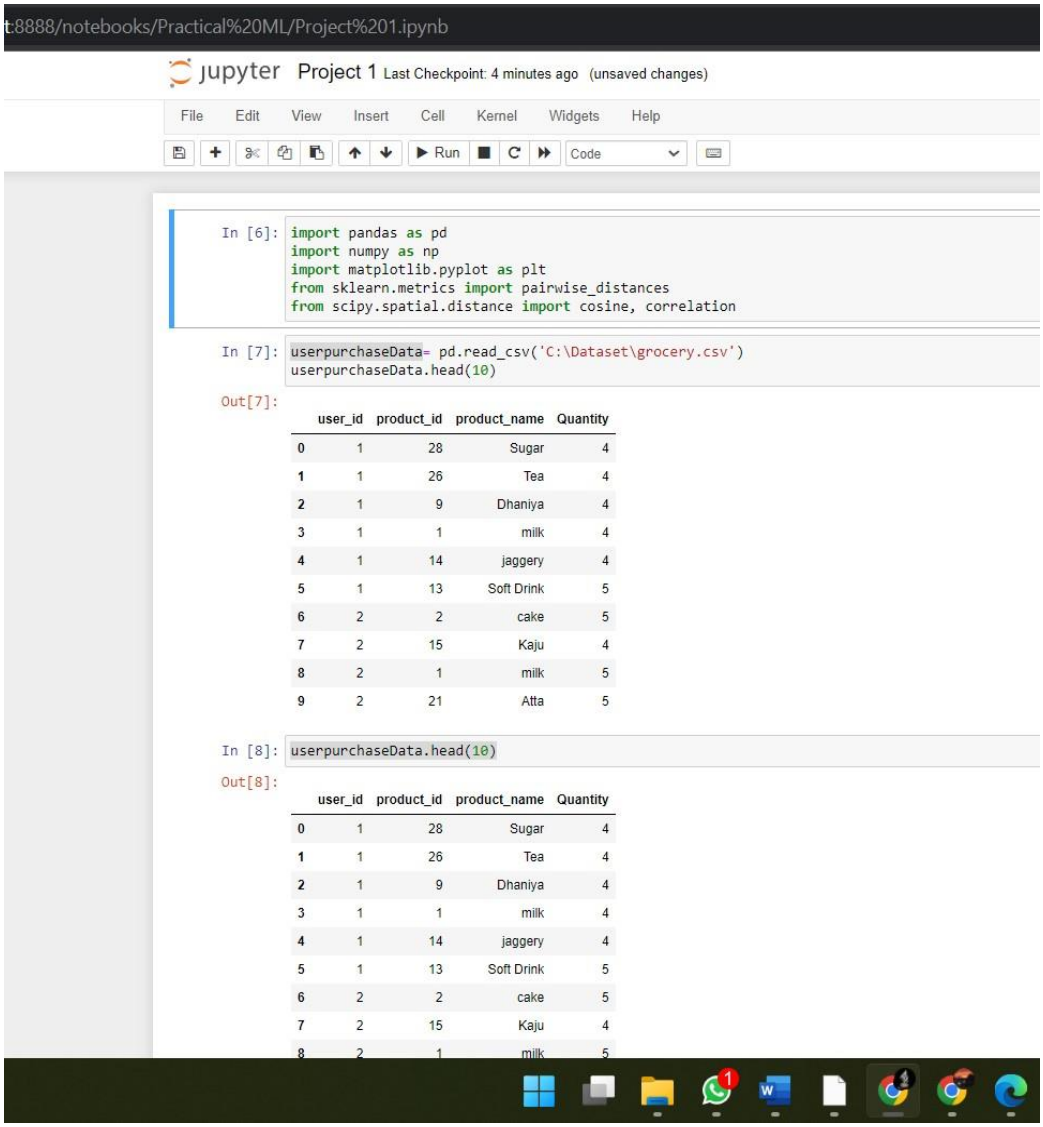
ID_6	Similar Data	Data split on the basis of User purchase product similarity	Data get slitted in the similar group of purchase history	Pass
ID_7	Splitting data	Training of data with machine learning algorithm	Machine learning Model get Train with Machine learning algorithm such as linear regression.	Pass
ID_8	Training Data	Testing data	Test whether the machine learning model satisfies the machine learning algorithm or not	Pass
ID_9	Recommendation Data	Recommend the correct product to user based on collaborative filtering	Recommended the right product to customer	Pass
ID_10	Training Model	After selected the one product recommend another product based on similarity user to each other	Recommend the similar product to similar customers	Pass

Chapter 9: Implementation

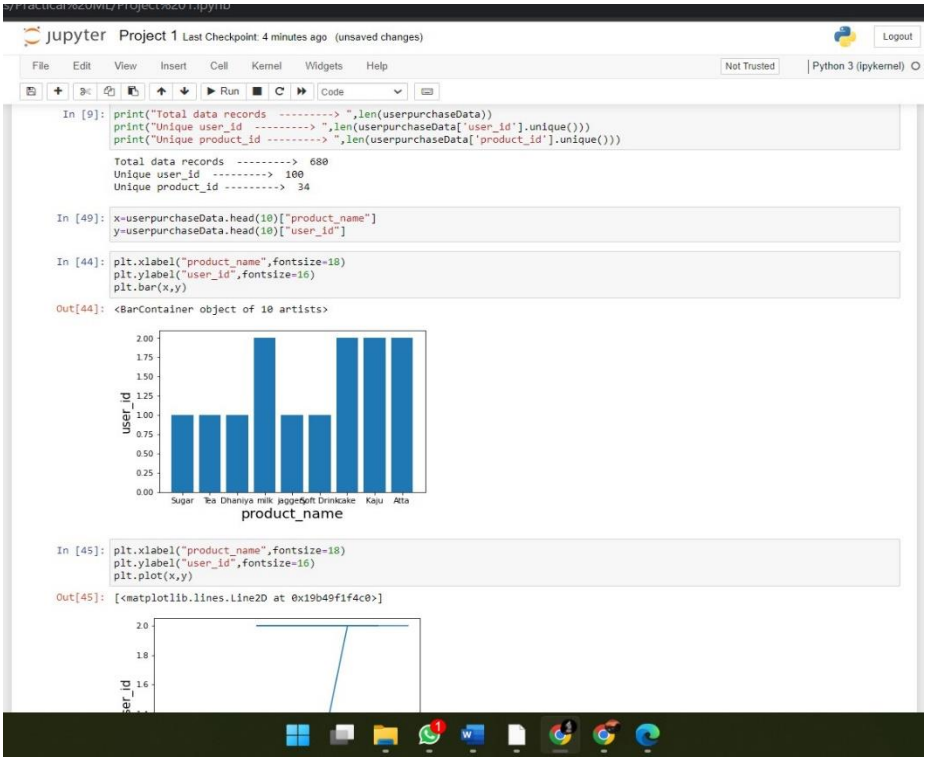
9.1 Implementation Steps:

We implemented our code in the Jupyter notebook using the grocery purchase history dataset of the 600 users records

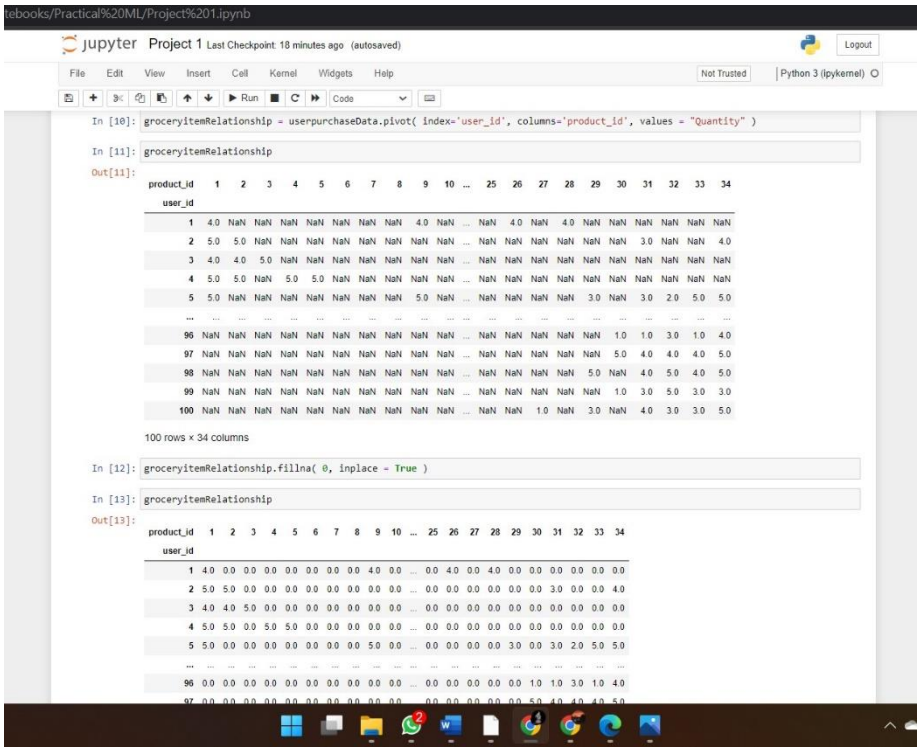
Step 1: Import important libraries and read the csv file.



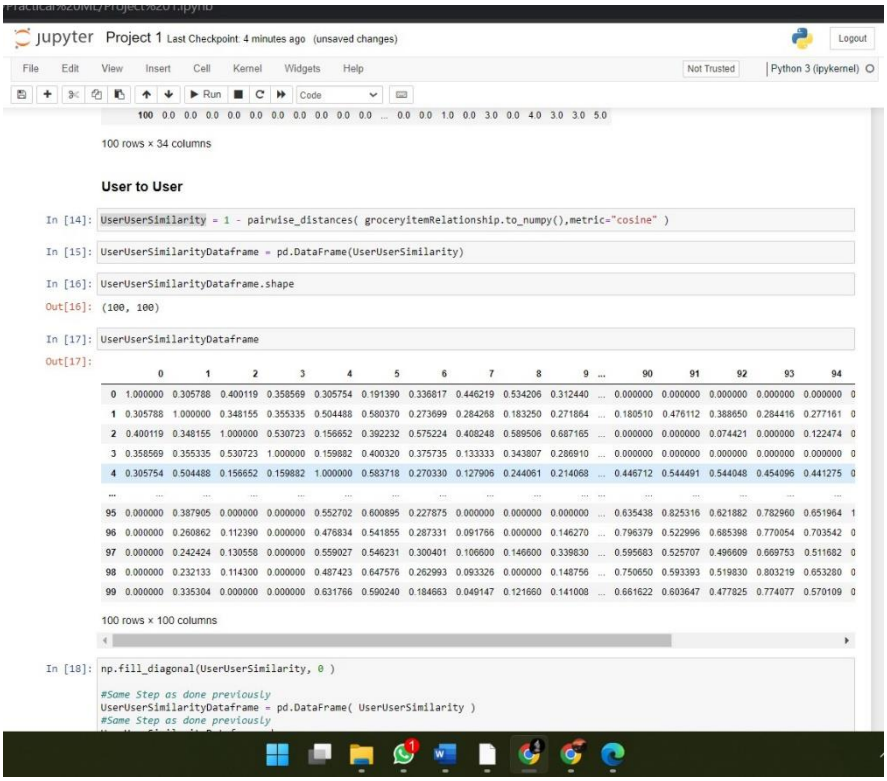
Step 2: In this step we clean the dataset and visualize the dataset with the help of libraries, such as matplotlib and SciPy.



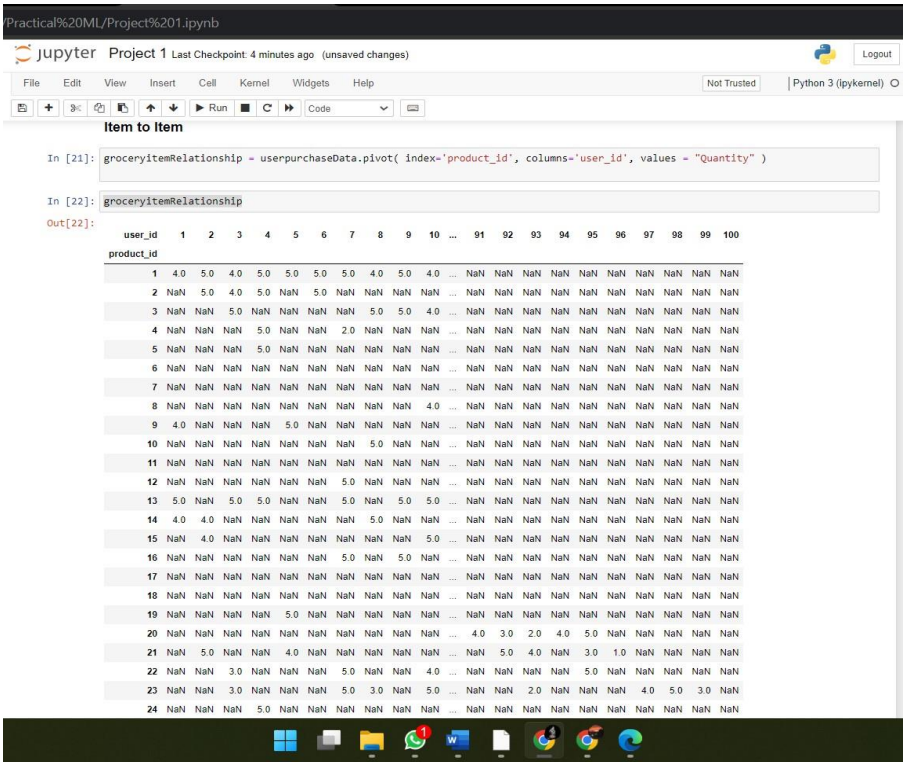
Step 3: In this step we find out the relation between users.



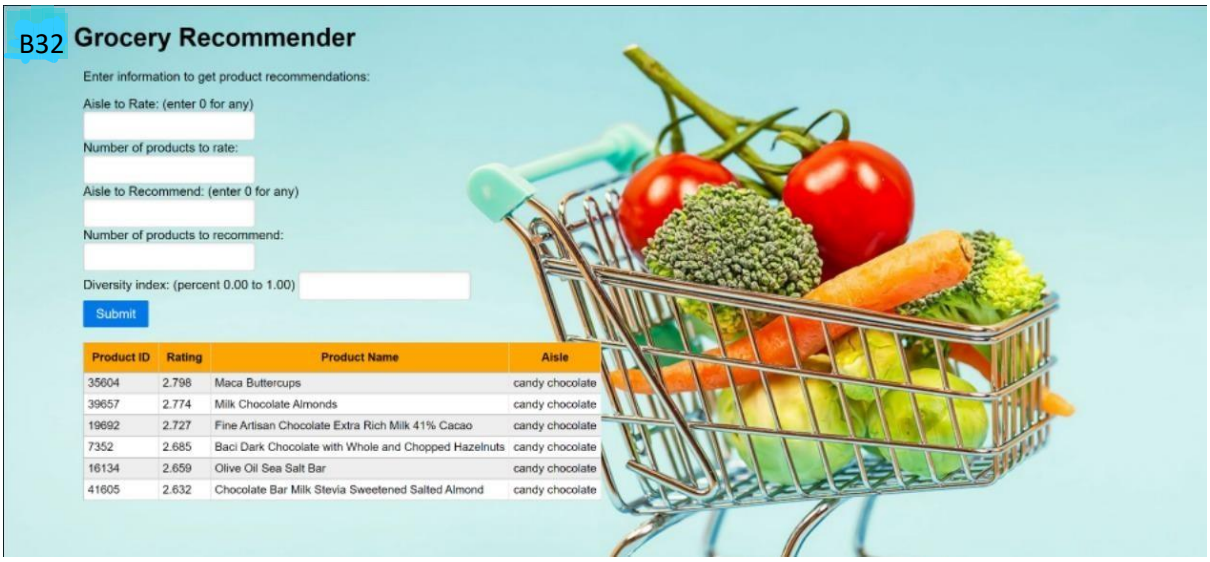
Steps 4: In this step we find out the user to user relationship.



Step 5: In this step we find out the item to item relationship.



9.2 Model Output:



Chapter 10: Conclusion

10.1 Conclusion:

A recommendation system is a subclass of information filtering systems that predict the items the user may be interested in based on the user past behaviour.

Collaborative filtering is one such recommendation technique that filters items of user interest based on user/item similarity. Due to ease of use and domain-free, it is being used and explored at a large scale by researchers. Data cleaning is one of the processes that increases prediction performance, yet insufficient for the cases of complex data sets as the one in this research. Applying single machine algorithm on the data set accuracy was less than 50 percent. Therefore, the ensemble of multiple machine learning algorithms has been proposed and this combination of ML methods gains accuracy of 92.38 percent. In this project, we have implemented item-based collaborative filtering to recommend product to users using cosine similarity. Other similarity metrics such as the Pearson correlation coefficient and Jaccard similarity could also be explored. This is still an open area of research with the motive to provide the user with the most relevant items.by using these we recommend the similar product to every customer.

10.2 Future Scope:

Recommendation system is a powerful tool which will be used by many websites to make the selection process easier for the user. Recommendations are used for making the work of the customer easier and faster. This reduces their valuable time and also the efforts. For this the recommendations given to the customer by this system is exact and fast.

Chapter 11: Reference

[1] Shraddha Gupta, Ankit Maithani, “A Literature Review on Recommendation Systems”, International Research Journal of Engineering and Technology (IRJET), e-ISSN: 2395-0056 Volume: 07 Issue: 09 | Sep 2020

[2] Sharon J. Moses, L.D. Dhinesh Babu, “Buyagain Grocery Recommender Algorithm for Online Shopping of Grocery and Gourmet Foods”, International Journal of Web Services Research Volume:15; Issue 3; July-September 2019

[3] Yi-Jing Wu and Wei-Guang Teng, “AN ENHANCED RECOMMENDATION SCHEME FOR ONLINE GROCERY SHOPPING”, 2011 IEEE 15th International Symposium on Consumer Electronics.

[4] Lamiyah Khattar, Dr. Geetika Munjal, “Analysis of Online Grocery Recommendation Systems” 2021 IEEE 11th International Conference on Cloud Computing, Data Science & Engineering;16-June-2021.

[5] Lee, Yunkyoung, "RECOMMENDATION SYSTEM USING COLLABORATIVE FILTERING" (2015). Master's Projects. 439; December 2015.

[6] Sharon J.Moses,L.D.Dhinesh Babu, “ Buyagain Grocery Recommender Algorithm for Online Shopping of Grocery and Gourmet Foods”, International Journal of Web Services Research; Volume 15; Issue 3; July-sep-2018.