

**Module Code & Module Title**

**CU6051NT Artificial Intelligence**

**Assessment Weightage & Type 75% Individual Coursework**

**Year and Semester 2021-22 Autumn**

**Student Name: Roshani karki**

**London Met ID:19031836**

**College ID: NP05CP4A190137@iic.edu.np**

**Assignment Due Date: 2022/01/27**

**Assignment Submission Date: 2022/01/27**

**Title: Product Recommendation System Word Count (Where required):**

*I confirm that I understand my coursework needs to be submitted online via Google Classroom under the relevant module page before the deadline in order for my assignment to be accepted and marked. I am fully aware that late submissions will be treated as non-submission and a mark of zero will be awarded.*

Acknowledgement

I would like to express our special thanks to our lecturer Mr. Prateek kokh Shrestha sir for the valuable guidance when performing this project which helped us patch this project and make it full completion at a time and for the continuous support while teaching, giving feedback while doing this coursework for his patience, motivation, enthusiasm and immense knowledge who gave us the golden opportunity to do this wonderful project which also helped us in doing lot of research. His instruction suggestion has served us as the great contributor towards the project. Because of his presence from the beginning session to end we go through the practice, research, online books regarding the coursework. He engaged us on tutorial, lecture classes from the first day and helped to do research in different topic for completing coursework given by college. I would be forever grateful for providing to complete the project of Artificial Intelligence, which helps us gain many knowledges towards this subject. It is always a pleasure for teachers and colleagues in the Itahari international College for their sincere guidance that we received to complete our project work in Artificial Intelligence.

I would like to thank my friends too for their guidance while I am in trouble in any questions while doing this project. The completion of coursework helps to uphold knowledge and skills toward Artificial Intelligence in future.

Abstract

Because of the rapid expansion of the internet and social media in recent years. Because of its ability to provide improved product, a product recommendation is vital in our social lives. Users can be recommended a set of products based on their interests or the popularity of the products. Despite the fact that a number of product recommendation systems have been presented, the majority of them either cannot recommend a product to existing users efficiently or cannot recommend a product to a new user at all. We offer a product recommendation system in this research that can recommend product to both new and existing users.

Recommender systems, also known as recommendation systems, are information filtering systems that are typically connected with a variety of consumer and commercial applications. These systems operate as a link between numerous content providers, such as social media websites, e-commerce portals, streaming platforms, and app users, by proposing items from the app database that match the user's tastes and previous activity. Such personalized systems play a vital role, especially when the user is unclear of the item to the searched for.

Table of Contents

[1. Introduction 1](#_Toc94080354)

[1.1. Explanation of the topic/AI concept used 1](#_Toc94080355)

[1.1.1. Artificial intelligence 1](#_Toc94080356)

[1.1.2. Machine learning 3](#_Toc94080357)

[1.1.3. AI concept used in recommendation system 6](#_Toc94080358)

[1.2. Explanation/introduction of the chosen problem domain/topic 7](#_Toc94080359)

[2. Background 9](#_Toc94080360)

[2.1. Research work done on the chosen topic/problem domain 9](#_Toc94080361)

[2.1.1. Collaborative Filtering System 11](#_Toc94080362)

[2.1.2. Example of Collaborative Filtering System 12](#_Toc94080363)

[2.1.3. Types of collaborative filtering 13](#_Toc94080364)

[2.1.4. Advantage and Disadvantage of collaborative filtering 15](#_Toc94080365)

[2.1.5. Dataset 16](#_Toc94080366)

[2.2. Analysis of existing work 17](#_Toc94080367)

[3. Solution 23](#_Toc94080368)

[3.1. Explanation of the proposed solution/approach to solving the problem 23](#_Toc94080369)

[3.2. Algorithm 25](#_Toc94080370)

[3.3. Pseudocode of the solution 31](#_Toc94080371)

[3.4. Diagrammatic representations of the solution (Flowchart) 32](#_Toc94080372)

[3.5. Explanation of the development process (with explanation of the used tools and technologies/libraries) 34](#_Toc94080373)

[Tools 34](#_Toc94080374)

[Libraries 35](#_Toc94080375)

[3.6. Achieved results (screenshots of the application/screenshots of the results attained) 40](#_Toc94080376)

[4. Conclusion 42](#_Toc94080377)

[4.1. Analysis of Work Done 42](#_Toc94080378)

[4.2. The Solution to Real-World Problems 43](#_Toc94080379)

[4.3. Further Work 44](#_Toc94080380)

[**Bibliography** 45](#_Toc94080381)

Table of figures

[Figure 1: Artificial Intelligence 2](#_Toc93993722)

[Figure 2: Machine Learning 3](#_Toc93993723)

[Figure 3: Recommendation system 8](#_Toc93993724)

[Figure 4: method Recommendation System 10](#_Toc93993725)

[Figure 5: collaborative filtering system 11](#_Toc93993726)

[Figure 6: Example of collaborative filtering system 12](#_Toc93993727)

[Figure 7: types of collaborative filtering 14](#_Toc93993728)

[Figure 8: Dataset of product recommendation system 16](#_Toc93993729)

[Figure 9: recommendation of Facebook 18](#_Toc93993730)

[Figure 10: recommendation of amazon 19](#_Toc93993731)

[Figure 11: recommendation of Netflix 20](#_Toc93993732)

[Figure 12: recommendation of Spotify 21](#_Toc93993733)

[Figure 13: recommendation of LinkedIn 22](#_Toc93993734)

[Figure 14: Item-Item based algorithm 26](#_Toc93993735)

[Figure 15: Item-Item based algorithm 26](#_Toc93993736)

[Figure 16: formula of cosine similarity 27](#_Toc93993737)

[Figure 17: formula of prediction computation 27](#_Toc93993738)

[Figure 18: Value of s (I1, I2) 28](#_Toc93993739)

[Figure 19: Value of s (I2, I3) 29](#_Toc93993740)

[Figure 20: Value of s (I1, I3) 29](#_Toc93993741)

[Figure 21: rating of User 1 Item 2 30](#_Toc93993742)

[Figure 22: rating of User 2 Item 3 30](#_Toc93993743)

[Figure 23: rating of User 4 Item 1 30](#_Toc93993744)

[Figure 24: flowchart of product recommendation system 33](#_Toc93993745)

[Figure 25: Importing Libraries 36](#_Toc93993746)

[Figure 26: Loading Dataset 36](#_Toc93993747)

[Figure 27: Displaying data 36](#_Toc93993748)

[Figure 28: pivoting in dataset 37](#_Toc93993749)

[Figure 29: Finding Similarity between product 37](#_Toc93993750)

[Figure 30: Creating DataFrame 38](#_Toc93993751)

[Figure 31: Recommendation of product 38](#_Toc93993752)

[Figure 32: Rating multiple products by user 40](#_Toc93993753)

[Figure 33: sum of every row of values 41](#_Toc93993754)

Table of tables

Table 1: Example of item-item based algorithm ........................................................... 26

# 1. Introduction

## 1.1. Explanation of the topic/AI concept used

### Artificial intelligence

Artificial intelligence (AI) is a wide-ranging branch of computer science concerned with building smart machines capable of performing tasks that typically require human intelligence. Artificial Intelligence refers to the intelligence displayed by machines. In today's world, Artificial Intelligence has become highly popular. It is the simulation of human intelligence in computers that have been programmed to learn and mimic human actions. These machines can learn from their mistakes and do human-like tasks. Artificial intelligence (AI) will have a significant impact on our quality of life as it develops. It's only natural that everyone today wants to connect with AI technology in some way, whether as a consumer or as a professional in the field. (Haton, 2006)

The phrase is widely used to refer to a project aimed at creating systems with humanlike cognitive abilities, such as the ability to reason, discern meaning, generalize, and learn from past experiences. Since the invention of the digital computer in the 1940s, it has been proved that computers can be programmed to perform extremely complicated jobs with ease, such as finding proofs for mathematical theorems or playing chess. Despite ongoing increases in computer processing speed and memory capacity, no program can yet match human adaptability across broader fields or in activities requiring a great deal of common knowledge. (Ertel, 2020)

Advantage of Artificial Intelligence

* **Reduction in Human Error**
* **Helping in Repetitive Jobs**
* **Digital Assistance**
* **Faster Decisions**
* **New Inventions**

Disadvantage of Artificial Intelligence

* **High Costs of Creation**
* **Making Humans Lazy**
* **Unemployment**
* **No Emotions**
* **Lacking Out of Box Thinking**

(Imran, December 23, 2019)

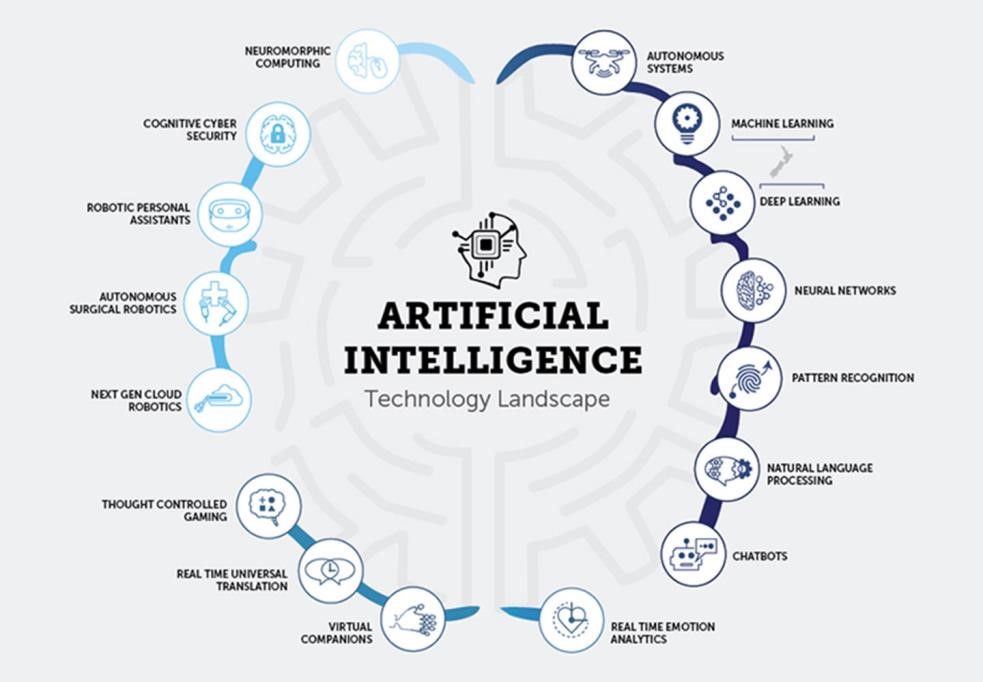


Figure 1: Artificial Intelligence

### Machine learning

Machine learning is used in artificial intelligence to replicate human intelligence. The computer must learn how to respond to specific activities, so it creates a propensity model using algorithms and historical data. After that, propensity models will begin to make predictions (like scoring leads or something).AI is capable of much more, but these are some of the most popular marketing applications and features. And, while it may appear like machines are poised to take over, humans are still required to perform much of the labor. We primarily utilize AI to save time by adding people to email automation and enabling AI to handle much of the heavy lifting while we focus on other activities.

Machine learning is a crucial part of the rapidly expanding discipline of data science. Algorithms are trained to generate classifications or predictions using statistical approaches, revealing crucial insights in data mining initiatives. Following that, these insights drive decision-making within applications and enterprises, with the goal of influencing important growth KPIs. As big data expands and grows, the demand for data scientists will rise, necessitating their assistance in identifying the most relevant business questions and, as a result, the data needed to answer them. (Kersting, 2018)

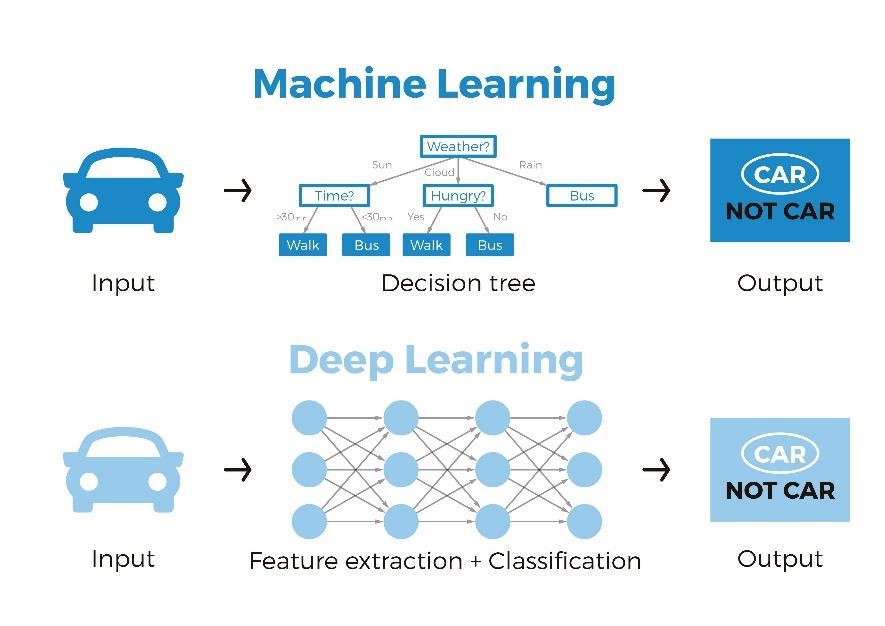


Figure 2: Machine Learning

* Advantage of Machine Learning
* Quick, accurate, and efficient
* Majority of applications are automated.
* Improved cyber security and spam filtering
* There is no need for human intervention.
* Disadvantage of Machine Learning
* Huge data requirement
* Error detection and correction are challenging
* Interpretation of data takes more effort and space
* Time consuming

(asquero, Aug. 16, 2020)

* Types of Machine Learning
* Supervised Learning

When you have input variables (x) and an output variable (Y), you can learn the mapping function from the input to the output using supervised learning.

Y = f(X)e and the concept of space

The goal is to estimate the mapping function to the point that you can forecast the output variables (Y) for new input data (x). Because the process of an algorithm learning from the training dataset can be thought of as a teacher supervising the learning process, it's called supervised learning. We know the correct answers, so the algorithm iterates through the training data, making predictions that are then corrected by the teacher. When the algorithm reaches a satisfactory level of performance, learning comes to an end.

* Unsupervised Learning

When using unsupervised learning, you just have input data (X) and no output variables. Unsupervised learning's purpose is to model the data's underlying structure or distribution in order to learn more about it. Unsupervised learning is so named because, unlike supervised learning, there are no correct answers and no teacher. Algorithms are left to their own devices when it comes to discovering and presenting the data's intriguing structure.

Some popular examples of unsupervised learning algorithms are:

* k-means for clustering problems.
* Apriori algorithm for association rule learning problems.

### AI concept used in recommendation system

Artificial intelligence (AI), specifically computational intelligence and machine learning methodologies and algorithms, has been used to construct recommender systems to improve prediction accuracy and tackle data sparsity and cold start difficulties. It carefully surveys various issues related to recommender systems that use AI, and also reviews the improvements made to these systems through the use of such AI approaches as fuzzy techniques, transfer learning, genetic algorithms, evolutionary algorithms, neural networks and deep learning, and active learning. (chua, Jun 26, 2019)

It is challenging for businesses in a competitive marketplace to offer products and services that appeal directly to an individual customer’s needs. Personalized e-services help to solve a major problem that of information overload thereby making the decision process easier for customers and enhancing user experience. The recommender systems used in these tailored e-services were built using techniques and theories from different artificial intelligence (AI) domains for user profile and preference finding, and they were initially established twenty years ago.

Various AI techniques have lately been applied to recommender systems, which has helped to improve the user experience and satisfaction. AI allows for a higher quality of recommendation than is possible with traditional methods. This has ushered in a new era for recommender systems, allowing for deeper insights into user-item relationships, the presentation of more complex data representations, and the discovery of comprehensive information in demographical, textual, virtual, and contextual data.

(Zhang, 2021)

## 1.2. Explanation/introduction of the chosen problem domain/topic

The recommendation engine is an excellent marketing tool, particularly for e-commerce, but it can also be used to boost profits, sales, and revenues in general. That's why tailored product recommendations are so popular in the retail industry, with eleven more studies proving the value of recommendation engines in the e-commerce sector. An information filtering system that uploads information according to users' interests, preferences, or behavioral history on an item is known as a recommendation engine. It can forecast a user's interest for a particular item based on their profile. Customers can easily and swiftly find the things they are seeking for thanks to product suggestion algorithms. So far, a few recommendation systems have been built to locate things that the user has already seen, purchased, or interacted with in some way. To be useful, a recommendation system must be adaptable to changing user behavior. It should be able to operate in a dynamic environment, providing users with quick information on special offers, assortment changes, and pricing modifications. (Isinkaye, 2015)

In general, algorithms for recommendation systems rely on previous purchases and page visits. Furthermore, many businesses now offer in-the-moment recommendations, as they employ artificial intelligence to analyses user interactions and choose visually appropriate products that will appeal to any given buyer. Recommendation engines use artificial intelligence to offer rapid and to-the-point recommendations that are personalized to each customer's needs and interests.

Artificial intelligence is helping to improve internet searches by making recommendations based on the user's aesthetic preferences rather than product descriptions. Artificial intelligence consulting engines appear to be posing a threat to search fields by assisting users in locating products or content that they might not otherwise be able to locate. As a result, today's recommendation engines are critical for sites like Amazon, Facebook, YouTube, and others.

Collaborative filtering needs a lot of data to create relevant suggestions. So, when you start using a platform with a collaborative filtering system, you start cold. The cold start problem in recommender systems is common for collaborative filtering systems. For example, when John visits YouTube for the first time, the system has to wait for him to watch several videos. Only then can it serve him relevant recommendations for other videos.

(rocca, jun 3 2019)

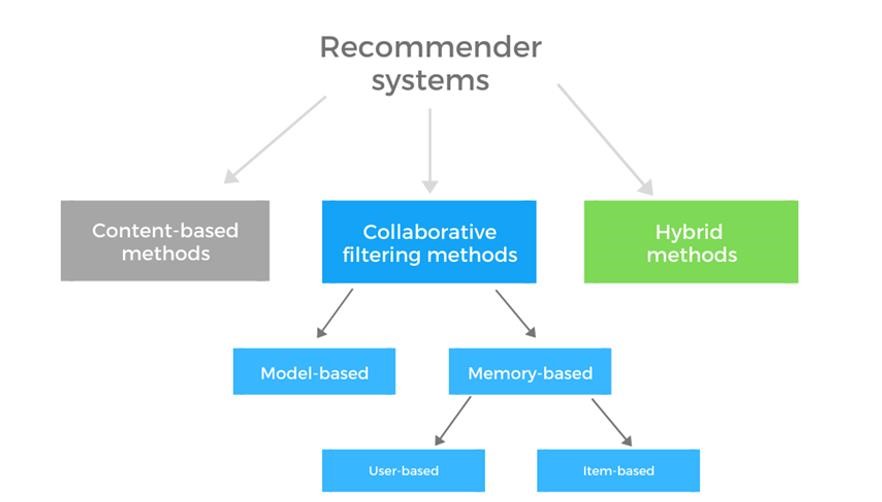


Figure 3: Recommendation system

# 2. Background

## 2.1. Research work done on the chosen topic/problem domain

#### Product Recommendation System

A product recommender system is a technology that uses machine learning to propose which products should be shown to people that interact with a brand's digital properties. Recommendation algorithms mine user, product, and contextual data both onsite and offline to provide each user with a personalized experience, which is fuelled by a number of algorithmic conclusions. Improving the discovery process, it helps consumers in finding what they are looking for and, in some cases, things they aren't even aware they are looking for. Businesses may learn more about each user's individual tastes and interests in this way, allowing them to optimize performance in real time while also refining their long-term testing roadmaps. When it comes to product suggestions, there is no one-size-fits-all approach that marketers should employ for each gadget. Depending on the quantity of information available about the client, their activity, and the context of products on a site, different methods must be used for different people. This contains information such as site activity, status, geolocation, time of day, previous transactions, and more. And, in order to understand customers, add value, and improve the overall relationship with a brand, recommendations can provide crucial insights and the opportunity to better grasp who a consumer is. (Retta, 2020)

The recommender system works with a vast amount of data by filtering the most important information based on the data provided by the user and other criteria such as the user's preferences and interests. It determines the compatibility of the user and the object, as well as the similarities between users and items, in order to make recommendations.

What is the purpose of the Recommendation System?

* Users benefit from being able to find items that are of interest to them.
* Assist item providers in getting their products to the correct people.
* Users will be able to identify products that are most relevant to them.
* Content that is tailored to the individual.
* Assist websites in increasing user engagement.

There are two methods to build a recommendation system:

1. Content Based Recommendation system
2. Collaborative Filtering Recommendation system

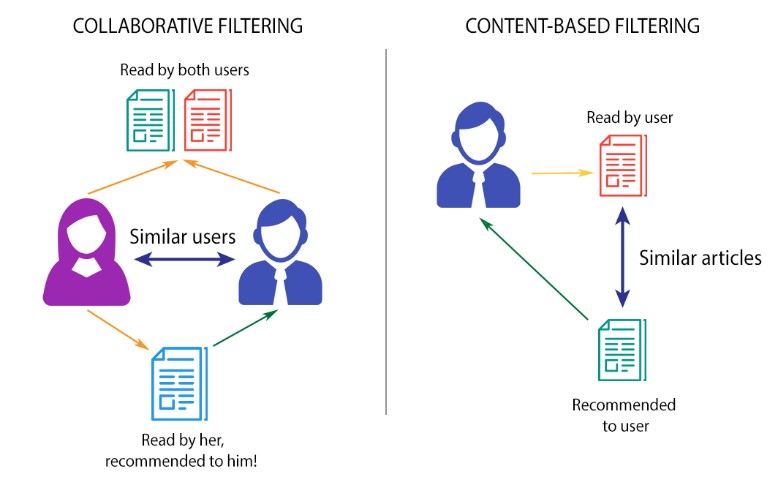


Figure 4: method Recommendation System

In my project I will be using Collaborative Filtering System.

### 2.1.1. Collaborative Filtering System

The most well-known application suggestion engine, Collaborative Filtering, is based on calculated guesses: people who liked the product will like it again in the future. A productbased collaborative shift is another name for this type of algorithm. Users are filtered and associated with each User instead of items in this Filtering. Only the conduct of users is taken into account in this system. It is insufficient to rely solely on their content and profile information. The behaviour of a User who gives a positive rating to a product will be linked to the behaviour of other Users who give a similar rating. (Nandi, July 14, 2017)

A list of items based on the user's previous choices is required for collaborative filtering.

To function, this system does not necessitate a large number of product characteristics. Each object and User are described by an embedding or feature vector, which sinks both the items and the users in the same embedding place. On its own, it generates enclosures for goods and users.

collaborative filtering is the suggestion of well-known, interesting, or popular material based on the people's location. The published stories will show a better communal interest as the collection of people becomes more diverse. Wikipedia is another example of collaborative filtering in action. Content extraction and analysis are not required for collaborative filtering. When compared to counting functioning, people will be able to appropriately analyse information. Complex things or multimedia, such as music, movies, and graphics, begin to function properly. (Nguyen, Feb 10, 2021)

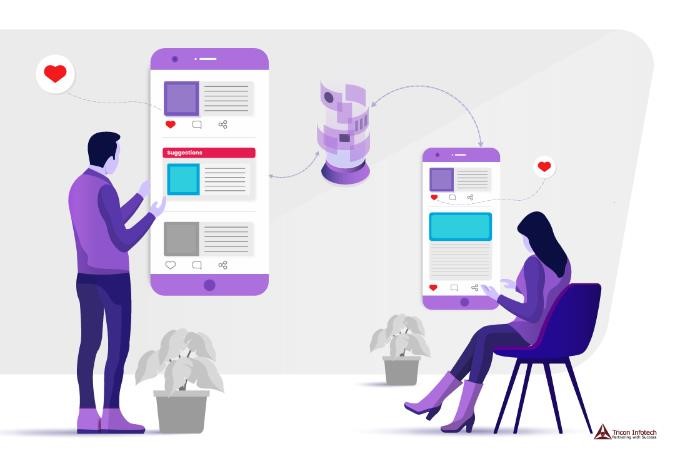


Figure 5: collaborative filtering system

### 2.1.2. Example of Collaborative Filtering System

In the field of e-commerce, one of the best examples of collaborative filtering can be found. When we visit an e-commerce website, we'll notice that it suggests various products for us. Some of the goods there are almost identical to what we're looking for. A question may now arise for us that how the website determines our interests. All of this is due to collaborative filtering.

Friends' recommendations are also highly prevalent on social networking platforms. For example, on Facebook, there is a section called People you may know, which is a really useful function that displays a list of people who can be added as friends. This method educates and estimates missing edges based on social connection data, such as if we are friends with the 10th out of 11th densely associated people, we must be friend with the 11th. Collaborative Filtering algorithms are used to create social relationships.

Let's look at another case. Ram and shyam enjoy playing the same games. Ram played it and had a great time with it. Although Ram has not yet played that game, the system has decided that shyam and Ram share a common interest, thus the system suggests that shyam play it. Using the same product, recommender systems can perform collaborative filtering. The item will appeal to the other buyer as well. (Ekstrand, 2011)

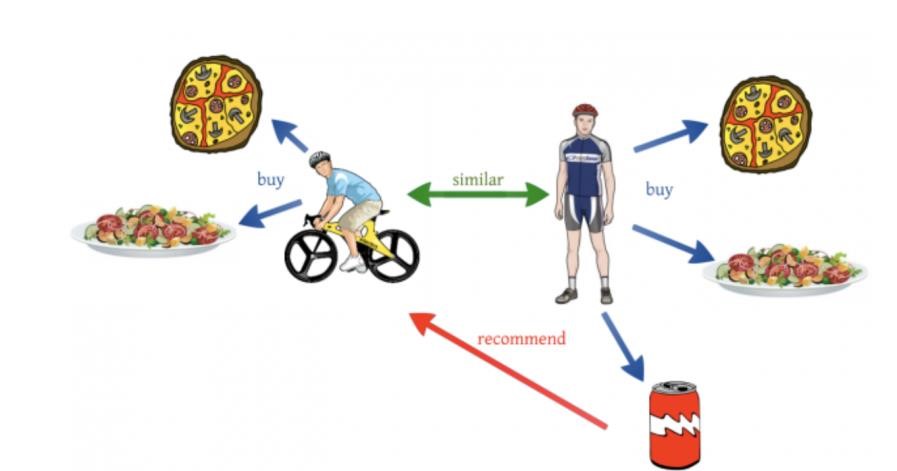


Figure 6: Example of collaborative filtering system

### 2.1.3. Types of collaborative filtering

There are two types of collaborative filtering. They are:

• Memory based collaborative filtering

Memory-based approaches leverage the data you already have (likes, votes, clicks, etc.) to develop correlations (similarities?) between people (Collaborative Filtering) or things (Content-Based Recommendation) in order to recommend an item, I to a user u who hasn't seen it before. We acquire recommendations from items seen by the users who are closest to you in the case of collaborative filtering, hence the word collaborative. Content-based recommendation, on the other hand, attempts to compare objects based on their qualities (movie genre, actors, book publisher or author, etc.) in order to suggest similar new goods. memory-based techniques rely heavily on simple similarity measures

(Cosine similarity, Pearson correlation, Jaccard coefficient… [etc)](http://inside.mines.edu/~ckarlsso/mining_portfolio/similarity.html) to match similar people or items together. If we have a huge matrix with users on one dimension and items on the other, with the cells containing votes or likes, then memory-based techniques use similarity measures on two vectors (rows or columns) of such a matrix to generate a number representing similarity.

There are two types of Memory based algorithm. They are:

* Item-Item based algorithm
* User-User based algorithm

• Model based collaborative filtering

Model-based methods, on the other hand, attempt to fill in the gaps in this matrix. They take on the challenge of "guessing" how much a user will like something they haven't seen before. They do this by using a variety of machine learning methods to train on a user's vector of objects, after which they can create a model that can anticipate the user's rating for a new item that has just been introduced to the system. Models are constructed using various data mining and machine learning algorithms to anticipate user ratings of unrated things in this method. A variety of model-based Collaborative Filtering techniques exist. Bayesian networks, Bayesian networks, Bayesian networks, Baye Singular value decomposition, clustering models, and latent semantic models’ Multiple multiplication factor, probabilistic latent semantic analysis, latent Dirichlet Models of allocation and Markov decision-making.

There are two types of Models based algorithm. They are:

* Matrix factorization based
* Clustering based

( yasserebrahim, October 13, 2012)

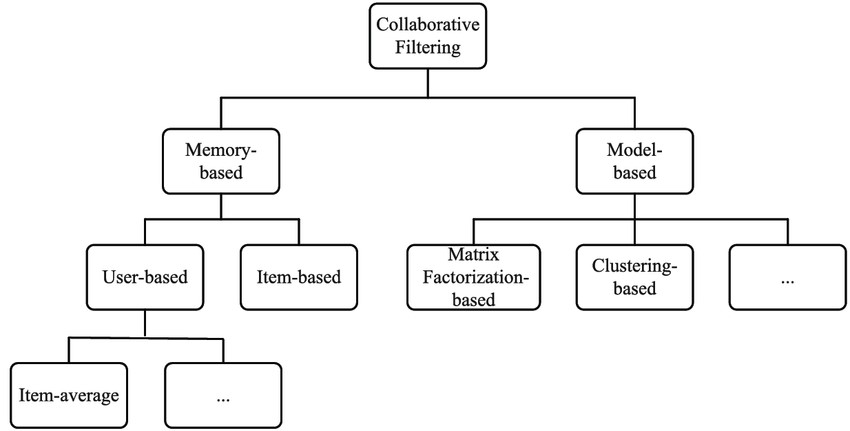


Figure 7: types of collaborative filtering

### 2.1.4. Advantage and Disadvantage of collaborative filtering

Advantage

* Because the embeddings are automatically taught, we don't need domain expertise.
* The model has the potential to assist users in discovering new hobbies. Even if the ML system does not know the user is interested in a certain item, it may recommend the product since other users are interested in it.
* Even when no information on an item is available, we still can predict the item rating without waiting for a user to purchase it.

#### Disadvantage

* Needs more data
* Problem with new users and new product.
* Usually recommend more popular items.

(iteratorshq, 2021-07-15)

### 2.1.5. Dataset

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

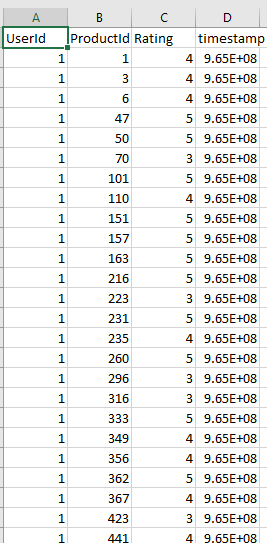


Figure 8: Dataset of product recommendation system

## 2.2. Analysis of existing work

Recommender systems are computer programs that make recommendations to users based on a variety of parameters. These systems forecast the most likely product that users will buy and that they will be interested in. Netflix, Amazon, and other companies employ recommender systems to assist their users find the right product or movie for them.

Collaborative Filtering is an important machine learning technique that helps a computer to filter information based on past interactions and data recorded on the user's end. collaborative filtering algorithms produce similar results based on the user's historical data. collaborative filtering algorithms produce similar results based on the user's historical data. (Su, 2009)

The role of collaborative filtering recommendations system at various real-world is described below:

1. Facebook

Facebook, a social networking site that was created in 2004, pioneered the field of social networking, which attempts to link people from all over the world. Facebook, which is currently managed by Mark Zuckerberg, employs a variety of AI techniques to enhance the social networking site. Collaborative Filtering, on the other hand, is one of the most remarkable strategies employed by this social media behemoth. "For various portions of the site, Facebook employs various recommender systems. For example, one algorithm may be used in the user timeline, while other recommender systems are used in the News and Marketplace sections to provide data it thinks is useful to the user. So, [Collaborative Filtering is used in Facebook.](http://rstudio-pubs-static.s3.amazonaws.com/269655_408f0b783aad496b82f4ae1da18ba239.html)

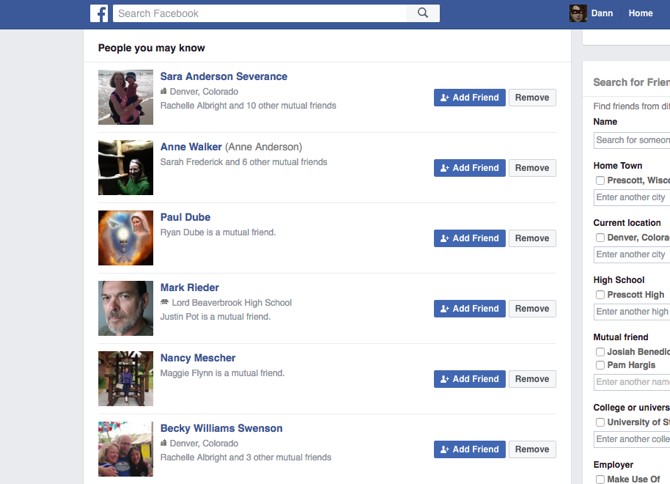


Figure 9: recommendation of Facebook

1. Amazon

Amazon, which was founded in 1994, used to simply sell books. Amazon uses the collaborative filtering technique into its recommendation system by utilizing a number of the top machine learning techniques for improved performance and enhanced user participation. What's more, this platform opts for item-based collaborative filtering more than a user-based approach in order to produce high-quality recommendations. Overall, the platform's item-based collaborative filtering has shown to be an effective approach that has increased the company's profit potential. At first, collaborative filtering had only one approach - a user-based approach. However, it was Amazon that developed an item-based approach that began to look at items rather than users.



Figure 10: recommendation of amazon

1. Netflix

Netflix one of the most renowned OTT platforms worldwide - Netflix. Known for its humongous entertainment collection and latest OTT content, Netflix was founded in 1997.With millions of users from around the world, the platform offers various recommendations to its users, done by collaborative filtering movie recommendation system. Recommendation algorithms are at the core of the Netflix product. They provide members with personalized suggestions to reduce the amount of time and frustration to find something great to watch. With so much to watch and learn from, platforms like Netflix have brought in the best of all worlds as they use recommender systems empowered by collaborative filtering approaches that narrow down options, and work according to our preferences.



Figure 11: recommendation of Netflix

1. Spotify

The team at Spotify appears to be constantly thinking up new ways to sort and recommend different types of music to its users. To get there, first they have to take the different kinds of data that they collect and build models that can analyse, compare, contrast, sort, and group the variety of information they’re getting. for example, is powered by close analysis of the songs a user has recently listened to and a scan of all playlists that may contain those songs or songs like them. Spotify uses a machine learning tool called the approximate nearest-neighbour [search algorithm](https://github.com/spotify/annoy) to group songs and users together based on shared attributes or qualities.

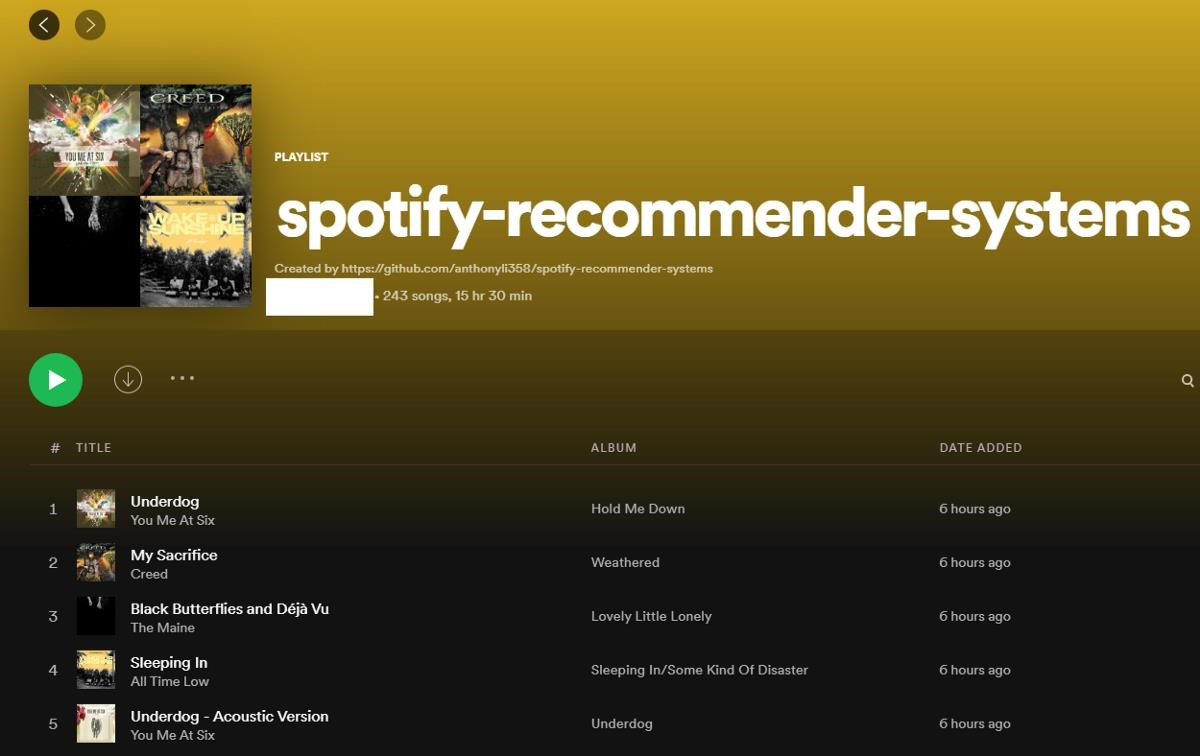


Figure 12: recommendation of Spotify

1. LinkedIn

A ranked list of candidates is provided by the LinkedIn Recruiter product in response to a search request in the form of a query, a job ad, or a suggested candidate. Candidates who match a search request are selected and then ranked using machine-learned models in multiple passes based on a variety of factors (such as the similarity of their work experience/skills with the search criteria, job posting location, and the likelihood of a response from an interested candidate). According to various papers on LinkedIn recommender, LinkedIn uses content matching and collaborative filtering to recommend companies or jobs a user might be interested in. LinkedIn collects data on positions, education, summary, specialty, experience and skills of a user from the profile itself.

Then, from member’s activity - data about member’s connections, the groups that the member has joined, the companies that the member follows amongst others are collected.

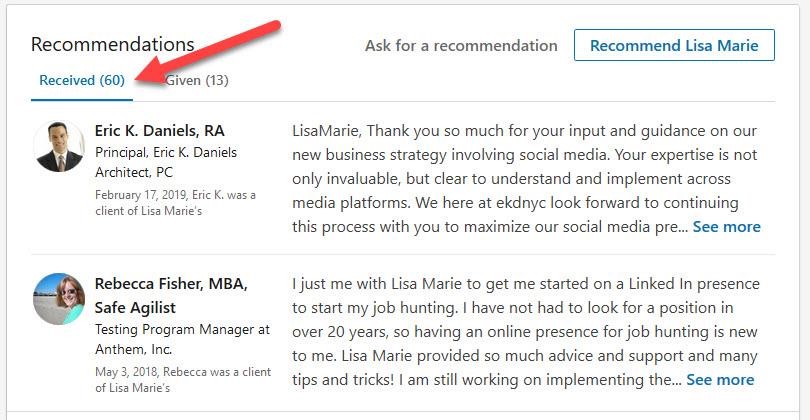


Figure 13: recommendation of LinkedIn

(Lokesh, 2019)

# 3. Solution

## 3.1. Explanation of the proposed solution/approach to solving the problem

As previously noted, the Recommendation may be quite useful in a variety of industries, like ecommerce, movies, and so on. These industries will benefit from the technology that will be developed. The technology will be able to recommend films to customers. The experiment will be built on a dummy dataset, and the data will be processed and the tests generated using an algorithm. I'll be working within this framework. In this framework I will be using Collaborative Filtering Recommendation System.

The spread of technology through internet is increasing day by day. With the increasing technology, use of the recommendation systems is coming into force. The recommendation systems solve many problems of customers by providing them recommendation based on their choice of products. Many of the collaborative filtering algorithms have been used for this purpose. The recommender system helps to suggests which text should be read next, which movie should be watched, and which product should be bought, creating a stickiness factor to any product or service. Its unique algorithms are designed to predict a users’ interest and suggest different products to the users in many different ways and retain that interest till the end.

* It narrows down the variety of choices, so people can focus only on those products which they are really interested in.
* Nowadays, many organizations use recommendation system algorithm to increase customer higher retention rate that leads to higher profit
* According, to Monetate Report, using a product recommendation system can lead to a 70% increase in sales, and that’s a lot. Customers make purchases even in the case initially they didn’t know what to buy.

Recommendation system solve the problems of users in different organization like Many online sellers implement recommender systems to generate sales through machine learning (ML). Many retail companies generate a high volume of sales by adopting and implementing this system on their websites. The pioneering organizations using recommenders like Netflix and Amazon have introduced their algorithms of recommendation systems to hook their customers. All these organization has implemented the recommendations system for the easiness to customer and to increase the marketing of business.

A solution to the cold start problem in recommender systems is clustering data with attribute similarities. Let’s go back to our YouTube example. John visits YouTube for the first time. The first video he selects is a Beyonce video. As mentioned before, the platform will cluster John with other users who watched the same video. It could also add him to other clusters. Let’s say the video belongs to the “pop song” cluster. Needless to say, the pop song cluster is populated with pop songs, such as “Hit Me Baby One More Time” by Britney Spears*.*

Now, the system can recommend other songs based on the following criteria:

* Other Beyonce Listeners’ Choices
* Other Songs in the Pop Cluster

(iterators, 2021-07-15)

## 3.2. Algorithm

An algorithm is a series of instructions telling a computer how to transform a set of facts about the world into useful information. The word algorithm is frequently used in conjunction with adjectives that describe the activity for which a set of rules has been developed. A search algorithm, for example, is a method for determining what type of information may be extracted from a vast amount of data. An encryption algorithm is a set of rules for encrypting data or messages so that only authorized people can read them. (AEDT, October 16, 2020)

Item-Item and User-User Memory Based Collaborative Filtering are the two forms of Memory Based Collaborative Filtering. I'd like to use an Item-Item Based Memory Collaborative Filtering technique for this project. Reasons why the User-User Based Memory Collaborative Filtering Approach was not chosen is in the market, the percentage of users who rate a product is extremely low, especially if a new item is available on the website. will be released with a low rating and no easy way for users to approach the product. If we work together, we can make a study of the current market scenario Many businesses believe in item-based pricing rather than user-based pricing.

So, for my project I have chosen item-Item based memory collaborative filtering.

#### Item-Item based algorithm

Amazon was the first to invent and implement it in 1998. Item-to-item collaborative filtering links each of the user's purchased and rated things to similar items, then combines those similar goods into a recommendation list, rather than matching the user to similar customers. The very first step is to build the model by finding similarity between all the item pairs. The similarity between item pairs can be found in different ways. One of the most common methods is to use cosine similarity.

Item-based collaborative filtering is a recommendation-based algorithm based on a model. The program determines the correlations between different elements in the dataset using one of several methods. These similarity values are then utilized to estimate user-item pair scores using a number of similarity tests. The value of similarity between objects is determined by looking at all of them. The similarity between two items depends on the ratings given to the items by users who rated both items. (Saumya, May 25, 2021)

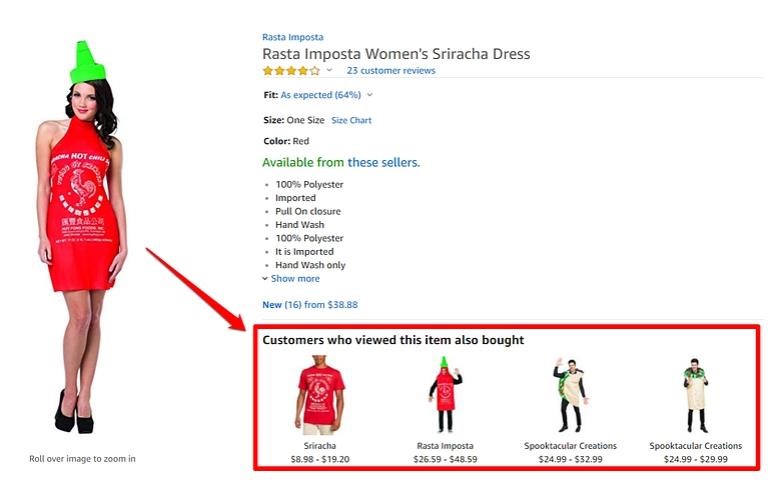


Figure 14: Item-Item based algorithm

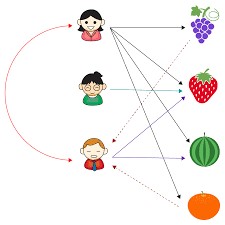


Figure 15: Item-Item based algorithm

Following is the process of working mechanism of item-item based algorithm.

#### Item to Item Similarity

The first step is to build the model by finding similarity between all the item pairs. The similarity between item pairs can be found in different ways. One of the most common methods is to use cosine similarity.

Formula for Cosine Similarity:



Figure 16: formula of cosine similarity

#### Prediction Computation

The second step involves executing a recommendation system. It uses the items

(Already rated by the user) that are most similar to the missing item to generate rating. We hence try to generate predictions based on the ratings of similar products. We compute this using a formula which computes rating for a particular item using weighted sum of the ratings of the other similar products.



Figure 17: formula of prediction computation

Example of item-item based algorithm using collaborative filtering.

Given below is a set table that contains some items and the user who have rated those items. The rating is explicit and is on a scale of 1 to 5. Each entry in the table denotes the rating given by an ith User to a jth Item. In most cases majority of cells are empty as a user rate only for few items. Here, 4 users and 3 items are taken. We need to find the missing ratings for the respective user.

|  |  |  |  |
| --- | --- | --- | --- |
| User/Item | Item\_1 | Item\_2 | Item\_3 |
| User\_1 | 2 | - | 3 |
| User\_2 | 5 | 2 | - |
| User\_3 | 3 | 3 | 1 |
| User\_4 | - | 2 | 2 |

*Table 1: Example of item-item based algorithm*

Step 1: Finding similarities of all the item pairs.

The item pairs are (Item\_1, Item\_2), (Item\_1, Item\_3), and (Item\_2, Item\_3). Selecting each item to pair one by one. After this, we find all the users who have rated for both the items in the item pair. calculating the similarity between the two items using the cosine formula stated above.

#### Sim (Item1, Item2)

In the table, we can see only User\_1 and User\_2 have rated for both items 1 and 2.

Thus, let I1 be vector for Item\_1 and I2 be for Item\_2. Then,

I1 = 5U2 + 3U3 and,

I2 = 2U2 + 3U3



Figure 18: Value of s (I1, I2)

So, the value of s (I1, I2) is 0.9

#### Sim (Item2, Item3)

In the table we can see only User\_3 and User\_4 have rated for both the items 1 and 2.

Thus, let I2 be vector for Item\_2 and I3 be for Item\_3. Then,

I2 = 3U3 + 2U4 and,

I3 = 1U3 + 2U4



Figure 19: Value of s (I2, I3)

So, the value of s (I2, I3) is 0.869

#### Sim (Item1, Item3)

In the table we can see only User\_1 and User\_3 have rated for both the items 1 and 2.

Thus, let I1 be vector for Item\_1 and I3 be for Item\_3. Then,

I1 = 2U1 + 3U3 and,

I3 = 3U1 + 1U3



Figure 20: Value of s (I1, I3)

So, the value of s (I1, I3) is 0.869

Step 2: Generating the missing ratings in the table

Now, in this step we calculate the ratings that are missing in the table.

#### Rating of Item\_2 for User\_1

Putting the value of s (I3, I2) and s (l1, l2)



Figure 21: rating of User 1 Item 2

#### Rating of Item\_3 for User\_2

Putting the value of s (I1, I3), s (l2, l3)



Figure 22: rating of User 2 Item 3

#### Rating of Item\_1 for User\_4

Putting the value of s (I1, I2), s (l1, l3)



Figure 23: rating of User 4 Item 1

Therefore, the ratings that are missing in the table of U1I2 is 2.49, U2I3 is 3.43 and U4I1 is 2.0

(geeksforgeeks, 16 Jul, 2020)

## 3.3. Pseudocode of the solution

Pseudocode is a haphazard and contrived method of developing programs in which we portray the sequence of actions and instructions (also known as algorithms) in a fashion that humans can understand. Computers and humans are vastly different, and here is where the trouble resides. A computer's language is quite strict: not permitted to make any errors or vary from the rules. Even with the creation of high-level, human-readable languages such as JavaScript and Python, an ordinary human developer still finds it difficult to reason and program in those coding languages. Pseudocode is a great way of getting started with software programming as a beginner.

(freecodecamp, Jul 26, 2021)

**START**

IMPORT pandas

IMPORT cosine\_similarity

READ dataset

PROCESS dataset

CREATE pivot table

DEFINE similarity

COMPUTE similarity with cosine similarity

CREATE DataFrame from calculated similarity array

CREATE function returning similarity product

CREATE an array of user having value product and rating

RECOMMEND 5 similar products

**END**

## 3.4. Diagrammatic representations of the solution (Flowchart)

A flowchart is a diagram that depicts an algorithm. It is frequently used by programmers as a problem-solving technique. It employs symbols that are linked together to represent the flow of information and processing. Flowcharting is the process of creating a flowchart for an algorithm. A graphical representation of an information system's or program's operation sequence. Flowcharts for information systems depict how data moves from source documents to the computer and then to users. The sequence of instructions in a single program or function is shown in a program flowchart. Each sort of flowchart has its own set of symbols. (geeksforgeeks, 20 Nov, 2020)

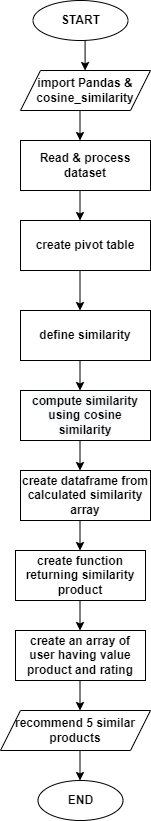


Figure 24: flowchart of product recommendation system

## 3.5. Explanation of the development process (with explanation of the used tools and technologies/libraries)

Here, I have developed the product recommendation system. This product recommendation system is developed only through tools and libraries. So, tools and libraries used for development are:

### Tools

I have used jupyter lab tools for implementing the development code in this product recommendation system project. JupyterLab is a free, open-source web application that primarily serves as a user interface for Jupyter Notebook. It has many of the same capabilities as the latter, such as a text editor, web browser support, and so on, but it has better support for third-party extensions. The conda and pip software packages can be installed using simple Python scripts. JupyterLab requires the packages, which are available for Windows, Mac, and Linux operating systems (OS). Jupyter Notebook is a free and open-source web tool that allows me to work with data in an interactive way. It generates papers (notebooks) that have both inputs (code) and outputs in one file. Jupyter notebooks support over 40 programming languages, with Python being the most popular. Anyone can use it for their data science projects as it is a free and open-source tool.

(SHARMA, Mar 5, 2018)

### Libraries

.

* Pandas

The Panda is a BSD-licensed open-source library. In the field of data science, the library is commonly employed. They are mostly used to analyze, manipulate, and sanitize data. Panda allows for easy modeling and data analysis without the need to convert to another language like R. Pandas is an open-source library designed to make it simple and natural to work with relational or labeled data. It includes a number of data structures and methods for working with numerical data and time series. This library is based on the NumPy Python library. Pandas is quick and has a high level of performance and productivity for its users. Pandas is a widely used open-source Python library for data science, data analysis, and machine learning activities. It is built on top of NumPy, a library that supports multi-dimensional arrays.

(activestate, October 9, 2020)

* Scikit-learn

Scikit-learn is undoubtedly Python's most helpful machine learning library. Classification, regression, clustering, and dimensionality reduction are just a few of the useful capabilities in the sklearn toolkit for machine learning and statistical modeling. Scikit-learn is a Python library that includes a variety of supervised and unsupervised learning techniques. Scikit-learn is a free Python machine learning library. It supports Python numerical and scientific libraries like NumPy and SciPy, as well as algorithms like support vector machine, random forests, and k-neighbors. In Python, Scikit-learn (Sklearn) is the most usable and robust machine learning library. It uses a Python consistency interface to give a set of efficient tools for machine learning and statistical modeling, such as classification, regression, clustering, and dimensionality reduction.

(dataquest, November 15, 2018)

Development Process

I have given description of all the cell that I have done in jupyter lab in this development process. While going through the development process I have used function, model, method which is described individual below.

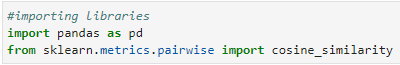


Figure 25: Importing Libraries

I import necessary packages that has used for product recommendation system.



Figure 26: Loading Dataset

Here, read\_csv is special method in pandas’ library that enable to extract data from csv file. Beauty\_file is a variable that extract ratings\_beauty.csv file.

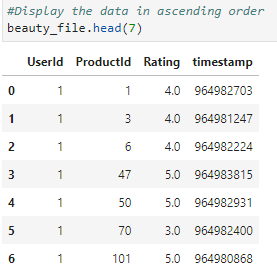


Figure 27: Displaying data

I have used head () function to display records of the dataset. Here, I have pass 7 in parenthesis. So, it displays 7 of the data in ascending order from dataset. We can pass any custom number in parenthesis and it display according to number if we don’t pass any number then, it displays default number data i.e., 5.

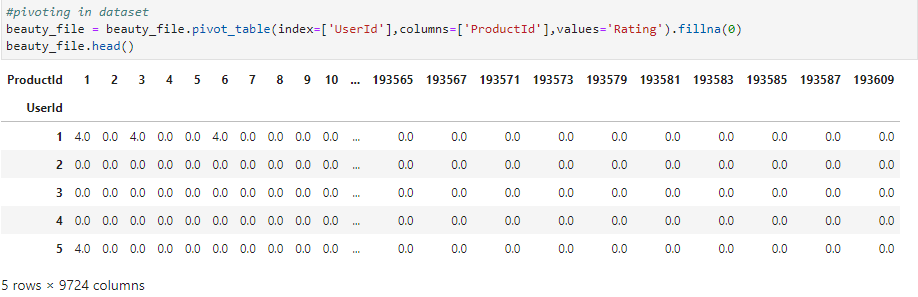


Figure 28: pivoting in dataset

Here pivot table is used using pivot\_table() function and the feature should be in the rows and columns using the index and columns parameters respectively. In the rows i.e., in index DataFrame of UserId is needed i.e., who rated the user, in the columns product is needed i.e., user has rated and in the value for each column we need rating that the user gave for particular product. In all the product user won’t give the ratings so, there will be lot of nan values hence, fillna can be used to replace the NaN values in the grouped table with the values that is provided in parenthesis here, 0 is provided so, its nan values are replaced by 0. and it is store in beauty\_file variable.

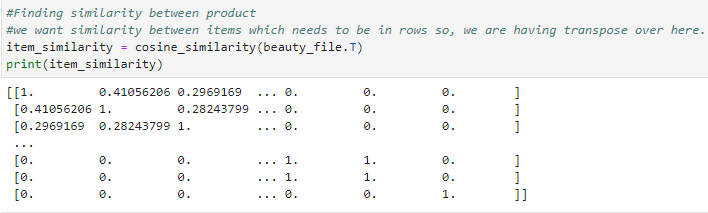


Figure 29: Finding Similarity between product

Here, computing the similarity between product using cosine similarity. I find the similarity between product using item-item collaborative filtering algorithm. Angular distances have been used between the ratings over here. Cosine similarity method calculate the similarity between row and in the rows, I have user over there but I want to calculate the similarity between item so, I transpose this matrix using T and print it. Then, Similarity matrix have been created and this is model based on the recommendation for new user.

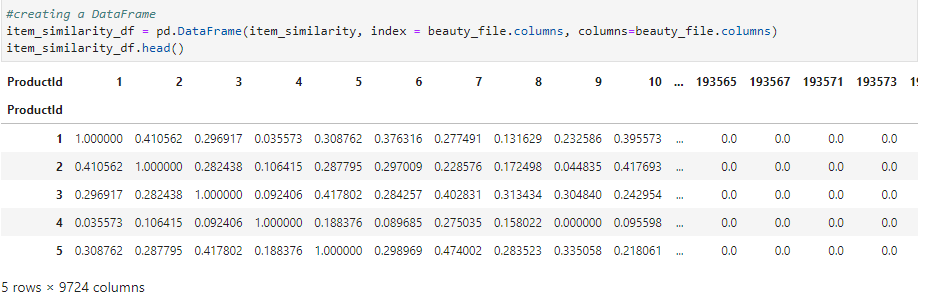


Figure 30: Creating DataFrame

Here, creating a dataframe from calculated similariy array. Before starting recommendation, I have created DataFrame which is in the form of NumPy array. For the convenient and easiness, I have converted the DataFrame into pandas DataFrame. Here pandas DataFrame create a DataFrame from item\_similarity numpy array. Both rows and column need the product so, both index and columns has beauty\_file.columns i.e. beauty\_file.columns indicates the product. Here first productId is 1 and is similar to 1 itself by 100% similarly, productId 1 is similar to 2 by 40%. These similarity value of product to product is needed while calculating the similar rating.

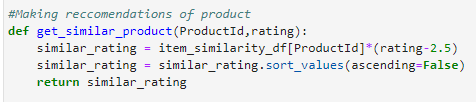


Figure 31: Recommendation of product

Here, at first function is created that returns the similiratity value of product. get\_similar\_product method is created that takes productId and the rating that user has given from the product that they have used in past. This method get\_similar\_product returns the similar\_rating for all the product that are similar to this particular product. If the user doesn’t like product and if they give the rating as 1 because of the fact that the user has rated the product other product will come on top if the user has rated the product bad then, we want other similar product should go down in the list so, user\_rating i.e., rating given by the user is subtracted by 2.5. If the ratings are below 3 it will push towards the negative sign and if the ratings are above 2.5 it will keep on the top of list. Rating that has been given for product by user gets subtraction with 2.5 and multiplies with item\_similarity\_df[ProductId] values that has find in before cell and it returns the similar\_rating.

## 3.6. Achieved results (screenshots of the application/screenshots of the results attained)

After completing the development process of application, finally achieved result is attained which is described below.

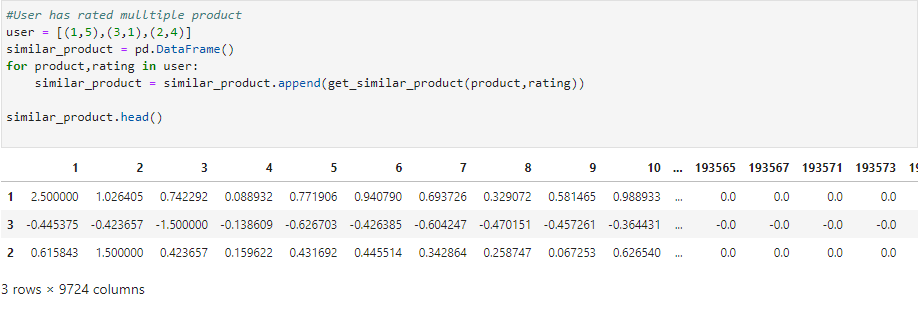


Figure 32: Rating multiple products by user

In this cell an array of user is created that contains product and rating in values. user has rated the multiple products; the user has rated 3 products where first element is the name of product and second element is rating that user gave to that particular product. Empty DataFrame is created to collect all the similar product for the suggestion to user. Append method is used to append the get\_similar\_product with variable that has been used in for loop. This appends method doesn’t return new list but it makes changes in the original list by adding the item. In the result of above, each row indicates each product that user has rated and the values in each column are the similarity score for particular product.

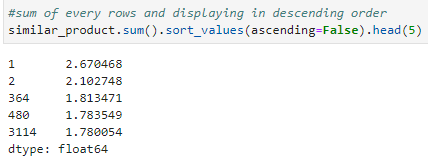


Figure 33: sum of every row of values

Based on the observation of 3 rows, it has sum up in 1 row for the suggestion of product for user. Sum method sum all the values in row wise and sort those values in descending order. As we see, product 1 is on the top because user has liked the product 1 most. Along, with the rating of product 1 other product id has recommended with its rating. It recommends the 5 similar products.

# 4. Conclusion

## 4.1. Analysis of Work Done

To summarize, collaborative filtering is more than just a mechanism for making recommendations. It's all about how the internet perceives us and provides us with the greatest possible service. It could appear invasive to some. However, a machine learning our consumer behaviors and recommending things based on who is using them and anything that mimics our own decisions is completely brilliant.

The topic I've chosen is Product Recommendation System, which is a widely used AI concept in the market. A simple prototype for assisting customers with product recommendations or similar products that they have searched for. for example, After the consumer has searched for clothes, the user will be presented with additional different clothes with their brands, color and size. So, here recommendation system is applied as it recommends different clothes after one clothes with its different features.

Recommender Systems have been widely used to exhibit the most appropriate items to users given their past consumption preferences. Recommendation systems are achieving great success in e-Commerce applications, during a live interaction with a customer; recommendation system may apply different techniques to solve the problem of making a correct and relevant product recommendation. Deep learning techniques with different neural network architectures can be applied to the recommendation systems to identify the different patterns and behaviors of the customers in different applications.

## 4.2. The Solution to Real-World Problems

We all interact with product recommendation algorithms almost every day as internet users through Google searches, while using movie or music streaming services, when shopping online, when reading social media, and when using dating applications.

As a result, product recommendation systems are one of the most successful and widely used machine learning applications in the corporate world. According to Monette’s recent research, product suggestions can result in a 70 percent boost in purchase rates, both in the first session and in subsequent sessions, as well as a 33 percent rise in average order values. Customers that click on product recommendations had 4.5x higher basket rates, 4.8x more product views per visit, and a 5x higher per-visit expenditure, according to a Salesforce study. (Mrukwa)

E-commerce platforms employ collaborative filtering. For example, if a customer is browsing products on an e-commerce website such as Amazon, he or she may receive product recommendations tailored to the user's preferences. This can greatly assist customers in making informed product choices and receiving the correct product. In addition, ecommerce businesses benefit from the recommendation system because they can increase their sales by a significant amount year after year. When a customer gets a recommended product, they will most likely go for it and buy it which gives the solution to increase the profit of unbalanced business.

* It makes the user to find the desirable product easily
* It shows the varieties of the options for the user
* It helps in the comparison on price, quality and materials etc
* It uses the algorithm and recommend the most relevant items to particular user

## 4.3. Further Work

In this report, we discuss what is artificial intelligence (AI) is and its importance in recommendation system, as this course work is based on sentiment analysis of product reviews. Research was conducted to determine the problem, types of technology, and algorithms should be used, flowchart, pseudocode of product recommendation system and the results were researched and presented above. At first primarily documentation is done about the introduction, problem, research, and so on are presented which is presented above in document.

After the completion of research and documentation, the following step is to put the findings into implementation. To do so, first of all dataset should be imported into the application. Research for this project has been completed till date and the research for development phrase has done very deeply. After having different researches development phase of product recommendation has done in jupyter lab using different libraries. While developing method, model, function is used and also input should be given by user. So, as per by research and as per by dedication I gave my 100% to this coursework and I have done the best for my project.

# **Bibliography**

activestate, October 9, 2020. s.l.: s.n.

asquero, Aug. 16, 2020.

Imran, M., December 23, 2019. s.l.: s.n.

iteratorshq, 2021-07-15.

SHARMA, A., Mar 5, 2018. [Online]   
Available at: https://analyticsindiamag.com/jupyterlab-what-is-it/

activestate, October 9, 2020. s.l.: s.n.

asquero, Aug. 16, 2020.

dataquest, November 15, 2018. s.l.: s.n.

Imran, M., December 23, 2019. s.l.: s.n.

iteratorshq, 2021-07-15.

SHARMA, A., Mar 5, 2018. [Online]   
Available at: https://analyticsindiamag.com/jupyterlab-what-is-it/

yasserebrahim. (October 13, 2012). Retrieved from https://yasserebrahim.wordpress.com/2012/10/13/memory-based-vs-modelbased-recommendation-systems/ (2019). Retrieved from witan.world. AEDT. (October 16, 2020). *theconversation*.

chua, R. (Jun 26, 2019). *medium*. Retrieved from https://medium.com/voice-techpodcast/a-simple-way-to-explain-the-recommendation-engine-in-aid1a609f59d97 Ekstrand, M. (2011).

Ertel, W. (2020).

freecodecamp. (Jul 26, 2021).

geeksforgeeks. ( 16 Jul, 2020). Retrieved from https://www.geeksforgeeks.org/item-toitem-based-collaborative-filtering/

geeksforgeeks. (20 Nov, 2020). gunjal. (2014). *database system.* bhojraj gunjal. gupta, s. g. (2012). *database management system.* khanna.

Haton, J. ( 2006).

Isinkaye, F. (2015). *sciencedirect*.

iterators. (2021-07-15).

Kersting, K. (2018). *frontiersin*.

Lokesh, A. (2019).

Mrukwa, G. (n.d.). *Jan 10, 2019.*

Nandi, M. (July 14, 2017). Retrieved from https://blog.dominodatalab.com/recommender-systems-collaborative-filtering Nguyen, T. (Feb 10, 2021).

*researchgate.net*. (2003). Retrieved from database management.

Retta, L. (2020). *dynamicyield*. Retrieved from https://www.dynamicyield.com/lesson/introduction-to-product-recommendersystems/

rocca, b. (jun 3 2019). In *Towards Data Science.*

Saumya. (May 25, 2021). In *analyticsvidhya.* singh, d. (2014). *information system.* delhi: khanna.

Su, X. (2009). *hindawi*.

Zhang, Q. (2021). *link.springer*.