**User Product Recommendation System Using**

**KNN-Means and Singular Value Decomposition**

**S.Akansha1, G Srikar Reddy2, C.N.S.Vinoth Kumar3\***

1,2 B.Tech Student, Department of Networking and Communications, College of Engineering and Technology (CET), SRM Institute of Science and Technology, Kattankulathur, India.

3\*Associate Professor, Department of Networking and Communications, College of Engineering and Technology (CET), SRM Institute of Science and Technology, Kattankulathur, India.

[sa7429@srmist.edu.in](mailto:sa7429@srmist.edu.in), [gr5249@srmist.edu.in](mailto:gr5249@srmist.edu.in), [vinothks1@srmist.edu.in](mailto:vinothks1@srmist.edu.in)

\*Corresponding Author: Dr. C.N.S.Vinoth Kumar

**Abstract:** - In the course of our regular lives, each and every one of us frequently makes use of the e-commerce websites in order to purchase a variety of goods. These results in a rising diversity of consumer demand, which makes it difficult for a retail company to deliver the appropriate products in accordance to the tastes of its customers. These e-commerce websites make use of a variety of recommendation systems in order to give the consumer a satisfying experience when they are shopping online. A technique that may be used to overcome this obstacle is called a recommender system. With the help of product recommendation, it is possible to meet the requirements of customers, which assists in sustaining loyal clients while also attracting new customers. In this project, we propose a system for product based recommendation by using hybrid machine learning algorithms which consists of the combination of KNN-Means and Singular Value Decomposition for the purpose of improving the overall excellence of our rating and recommendation system, that improvises the disadvantages of collaborative filtering on their own. This method has the benefit that there is no necessity for the development of a brand new algorithm in order to calculate the forecasts. The trials demonstrate that our strategy can provide superior outcomes, and our integrated model provides an exceptionally high level of accurate predictions.

***Keywords:*** *Cold Start, Content Filtering, Sparsity, Recommender System, Privacy, Collaborative Filtering, KNN-Means, Singular Value Decomposition, Hybrid Filtering.*

**1. Introduction**

Recommender systems, also referred to as recommendation systems, are a class of information filtering systems that use a model created from an item's characteristics (content based approaches) or the user's social environment to try and predict the "rating" or "preference" that a user will give to a social element (such as people or a group) or an item (such as music, books, or movies) that they have not yet considered. This is done by employing a model built. This is accomplished by comparing the characteristics of the collaborative filtering approaches. A system that makes suggestions or recommendations to one who makes recommendations is known as a recommender system. The purpose of the recommender system is to provide relevant product suggestions provided to a group of people for items or products that have the potential. Recommendation systems are approaches that are used to predict the rating that one person would give to another for a product or social institution. This is how these techniques are defined. These things can be anything from books to movies to restaurants to other things on which different people have varying interests. These preferences are being anticipated by employing two different methods: the first method is called a content based method, and it involves the properties of a particular item. The second method is called a collaborative filtering method, and it involves taking into consideration previous user behaviour to make decisions. Within the framework of collaborative filtering, partners are selected to provide recommendations on the basis that they have a rating history that is comparable to that of the target user. There is a possibility that a single partner with ratings that are comparable to those of the intended user are not a trustworthy prediction for a particular item. Because of this, the history of success enjoyed by the partner in question in terms of making solid recommendations additionally, it is necessary to take this into account, as this is determined by the partner's trustworthiness. Those who really issue reputation ratings to the partners are the ones who come into play while thinking about how to implement a recommender reputation system for the purpose of maintaining a record of previous interactions.

1. **Related Work**

In [1], ratings are kept in matrices whose sizes grow exponentially, rather than being calculated by an algorithm. This avoids the need for a new algorithm to calculate the predictions of each individual user's score. [2] offers a thorough analysis of current works that incorporate review texts and explores how these review texts might be

used to overcome some of the major problems with conventional collaborative filtering algorithms. [3] have used the dimensionality reduction provided by the SVD to reduce the data's sparsity so that they can offer a list of movies based on the information provided.. The goal of [4] is to deliver a location-based agricultural product recommendation system. The great circle distance is used to determine how far the vendor and buyer are from one another based on their locations. As a result, customers are given product recommendations based on similar search results from the closest merchants. [5] employed sentiment analysis on a recommendation system based on hybrid approach of context based engine and stochastic learning. They were able to significantly reduce the size of the search space by adopting a hybrid method, allowing users to search for any items anywhere and at any time. With the proposed approach, ratings, emotions and review are categorized to negative and positive sentiments. [6] merged embedded users and goods with deep neural networks. The authors suggest that in order to improve the quality of suggestions, further deep learning techniques should be investigated. In this study, the MovieLens dataset which is divided into 80% and 20% and is utilized to carry out the evaluations.

Similar recommendation algorithms employing bipartite graphs exist as well [7][8][9]. Bipartite graphs were used in the academic article recommendation model developed by Ohta et al. A label-based similarity calculation method is one that looks for similarities between many articles that include the same term. The most straightforward and readily implementable TF-IDF approach was adopted for this strategy, despite the fact that researchers have proposed other ways in previous studies. Numerous additional research [11][12][13][14][15] employ similar methodology for the same objective as ours.

In addition to this, there are strategies that integrate two sets of data and utilise them as a way for recommending academic articles, as done by Zhao et al. [16]. [17] discussed a few issues and difficulties with the current recommender systems and comes to the conclusion that new methods should be used to construct recommender systems. Search the books list based on their rating and content using cooperative filtering and content-based filtering. The recommendation mechanism mostly relies on book ratings and recommendations from current users. [18][19] uses the SVD algorithm on the movie lens data set to get the best movie suggestions. In [20], collaborative filtering and context-based filtering algorithms are combined. He offered many methods for developing this system. Their research' findings indicate that hybrid techniques are more effective than non-hybrid ones. The programme generates a quotation by using the supplied text as a parameter.

**3. Problem Statement**

The idea behind a popularity based recommendation system is to explicitly suggest things that are popular or in style. The popularity based recommendation system is not personalized and would suggest the same kind of products or items that are purely determined by how popular they are among all other users, despite the fact that it does not have difficulty starting up when it is cold and that it is able to make product recommendations based on a wide number of filters without the need for the user's historical data. This is because the system can recommend products based on popularity without the requirement for access to the user's previous data. Even if there are many users and many things that need to be suggested often in collaborative filtering, issues with sparse user and rating matrices might occur, making it difficult to identify users who have rated the same item. Systems for making recommendations increasingly contain models to enhance, and make efforts to fix concerns with memory based methods (KNN-Means). Model based algorithms (Singular Value Decomposition) are also based on historical user evaluations, but instead of generating predictions directly, this method classifies people or creates models from their data. This strategy divides people into categories or builds models from their data instead of making predictions directly.

Considering the benefits and drawbacks of the two aforementioned strategies, it is clear that many systems are built on their mixture, giving rise to the term "hybrid filtering systems". Multiple suggestions strategies are used in hybrid recommender systems in order to create a synergy between them.

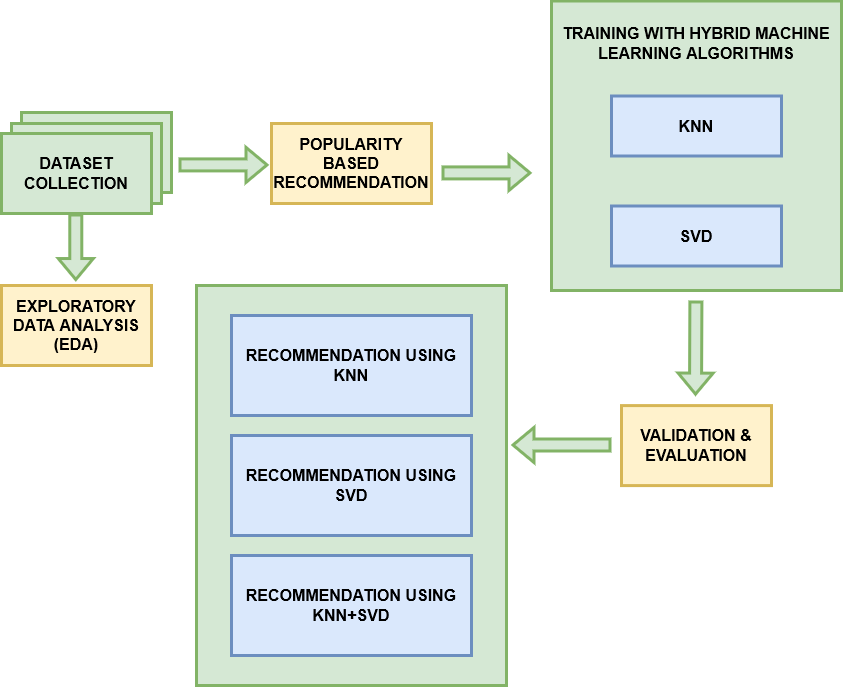


Fig 1 Proposed System Architecture

1. **Data Analysis and Executing the**

**Hybrid Model**

The dataset is collected and then it will be used as the basis for training going forward with the hybrid deep learning algorithms. Kaggle website will be used for the dataset collection. Exploratory Data Analysis sometimes known by its acronym EDA, is a type of data analysis that places an emphasis on the utilization of various visual tools. It is possible to use it to find patterns and trends, as well as validate assumptions, with the assistance of statistical summaries and graphical representations. Methods of data visualization are utilized quite frequently within the context of these procedures. EDA is used primarily to investigate what the data may disclose beyond the work of formal modelling or testing hypotheses. In addition to this, it can help determine whether the statistical approaches that are being examined right now for the analysis of the data are appropriate or not.

Popularity based recommendationis a kind of recommendation system that functions in accordance with the public opinion rule in addition to whatever else is presently in style. Specifically, it is a form of recommendation system. After conducting research on the product or movie that is currently in style or that is the most well-liked among the clients, these systems quickly recommend the researched product or movie to the customers. For instance, if the majority of people frequently purchase a certain product, the system will figure out that this product is the most popular because of the high volume of purchases. This directly leads to the system recommending buying that product to each new user that registers for it, increasing the likelihood that the new user will also buy that product. Because it does not encounter the issues that are associated with a cold start, it may begin making product recommendations based on a wide variety of filters on the very first day that it is in business. The previous information that the user has provided is not necessary at this time. As the name suggests it recommends based on what is currently trending/popular across the site. This is particularly useful when we don't have past data as a reference to recommend product to the user. It is not tailor fit for any particular group of audience or movie. The algorithm would recommend the exact same things or movies to each and every user, and the criteria for these recommendations would be limited solely to how popular the products or movies are.

Following the completion of the data collection, it will be used for training with the hybrid machine learning algorithms such as KNN and SVD. The abbreviation KNN stands for “K-Nearest Neighbour”. It is a form of machine learning known as supervised machine learning. The algorithm can be utilized to answer problem statements that pertain to classification as well as regression. A new unknown variable that needs to be predicted or categorized is denoted by the letter 'K', which stands for the number of variables that are the nearest neighbors to the new variable. Its purpose is to identify all of the data points that are located in the immediate vicinity of an unidentified data point in order to determine which category the point belongs to. It is a technique that is based on distance. KNN operates by first calculating the distance between all of the points that are close to the unknown data, and then excluding from consideration the points that have the shortest distances to the unknown data. This is done in order to find the best candidate for the unknown data.

The K-Nearest Neighbors approach is a non-

parametric methodology, which means that in order to carry out its computations, it makes no assumptions about the data that it is studying. It is sometimes referred to as a lazy learner algorithm since it stores the dataset rather than learning directly from the training set. Then, when the time comes for classification, it performs an action on the dataset. Because of this, it performs less well than other learning algorithms. During the training phase, the algorithm does nothing more than store the dataset. The moment it obtains a new data, it promptly places that data into a category that is quite similar to the category in which it was previously stored.

This technique operates under the presumption that data points with comparable characteristics can be found in close proximity to one another. Because of its user-friendliness, its applicability to problems involving classification and regression, and the simplicity with which its own results may be interpreted, it is frequently put to use in a variety of fields for the purpose of problem solving.



Fig 2 KNN Algorithm node setup

A factorization of one matrix into three additional matrices is referred to as the Singular Value Decomposition (SVD) of that matrix. It reveals fundamental geometrical and theoretical insights regarding linear transformations, in addition to having some fascinating algebraic features, and it does so in a concise manner. Additionally, it possesses a number of essential applications in the field of data science.

The Singular Value Decomposition method begins with a rectangular matrix of gene expression data, which is denoted by the letter A and is in the form of a n by p matrix. Within this matrix, the n rows stand for the genes and the p columns stand for the experimental circumstances. The SVD theorem states:

Anxp= Unxn Snxp VTpxp ……………….(1)

Where

UTU = Inxn

VTV = Ipxp (i.e. U and V are orthogonal)

The columns of V and U are generated using the eigenvectors of their respective vectors, ATA and AAT. The rows of VT include the right singular vectors, and S has singular values and is diagonal (mode amplitudes). If the matrix A is a real matrix, then the matrices U and V will also be real if the same thing is true for matrix A. A representation of an expansion of the primary data set in a coordinate system in which the covariance matrix is diagonal is called the singular value decomposition (SVD).

After applying the hybrid machine learning algorithms, the dataset has been separated into two parts, with the first half devoted to training and the second to testing. The entirety of the dataset is divided as follows: 75% goes to the testing set, and 25% is allocated to the training set. The model should first be trained with the help of the training set, and then it should be verified with the assistance of the test set. After the conclusion of the data training, the trained model file is formed, and then the trained model file is provided with the testing data. Validation in machine learning or deep learning is like an authorization or authentication of the prediction done by a trained model. While on the other hand, evaluation in machine learning refers to assessment or test of entire machine learning model and its performance in various circumstances. It involves assessment of machine learning model training process, deep learning algorithms performance and how accurate is the predictions given in different situations.

**Recommendation using KNN (Memory Based Collaborative Filtering**

The information is typically provided in the form of a "User x Item" matrix when it is being processed by a filtering system that is memory based and collaborative based. Users are represented in the rows, and the items themselves are displayed in the columns. In order to formulate suggestions, the memory based method type makes use of user feedback on various things (often in the form of reviews). This type places a heavy emphasis on the use of statistical approaches in order to discover neighbouring users who have rated the same set of goods in a manner that is equivalent to the ratings provided by the active user. Memory based approaches make use of a variety of algorithms to aggregate the thoughts and experiences of users' neighbors in order to provide predictions for active users after the neighbors have been discovered. The majority of this strategy is based on ranks. The amount of weight that is given to the rating of each user is determined by the degree of correlation that exists between that user and the other users in the system, including the person for whom we want to make the proposal. Because systems typically need to be able to accommodate a significant number of users, making suggestions that are based on ratings from millions of users might have severe ramifications for the performance of the system. In addition, when the number of users reaches a certain level that has been established beforehand, a decision needs to be taken regarding which neighbors are the "best". Pearson's similarity is built on the foundation of the calculation of correlations, and the only user currents that are taken into consideration are those that are correlated. In the realm of information retrieval, vector space models

are extremely common. For the purpose of this discussion, we will concentrate on numerical similarity. These techniques represent each thing as a feature vector in dimensions space. The angle formed by their vectors will be (U, V), and the cosine of this angle will represent the degree of similarity.

**Recommendation using SVD (Model Based Collaborative Filtering)**

The use of models has been introduced into recommendation systems as a means of improving upon and addressing issues with memory based methods. This method does not explicitly generate predictions; rather, it divides users into categories or trains models from their data based on how they interact with the system. These groups and models are determined by how the users interact with the system. The model serves as the foundation for the development of algorithms that take into account the historical assessments (profiles) of users. Several different approaches are used in the process of building the model. The majority of the time, machine learning techniques including clustering, matrix factorization, Bayesian networks, and decision trees are used in the processes built on the model. The main focus of our method will be matrix factorization, sometimes referred to as matrix decomposition. The process involves disassembling a matrix into a number of distinct new matrices. It will be sufficient to perform the product of these matrices between each other in order to discover the initial matrix. Matrix factorization has proven to be effective in recommender systems, with positive outcomes.

**Recommendation using Hybrid Model (KNN+SVD)**

It is acknowledged that many systems are based on the mixture of the two ways described above, which is what makes them so called hybrid filtering systems. This is because both approaches have their positive and negative aspects to consider. In most cases, hybridization can be broken down into two distinct stages. Generate candidate recommendations by independently applying collaborative filtering and other filtering approaches. To create the recommendations that consumers will ultimately utilize, combine these various early ideas using a range of strategies, including weighing, mixing, cascading, and switching. In a broader sense, hybrid systems are in charge of keeping user profiles that are content-oriented, and the comparison of these profiles results in the creation of user communities that make collaborative filtering possible. Hybrid systems allow these communities.

**5. Proposed Method**

In order to smooth their drawbacks and improve performance when proposing items to customers based on their interests, we have built an universal unifying model within the scope of this work that includes both the collaborative based KNN-Means and SVD based algorithms. The suggested method also resolves the sparsity and cold-start issues and suggests and recommends the appropriate products according to the customers’ interests.

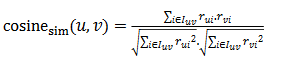
In this project we suggest a solution for recommending system for products with the assistance of hybrid machine learning algorithms. The electronic datasets have been collected from Kaggle website. After considering the dataset the next step is to perform Exploratory Data Analysis (EDA) as it helps in providing the context that is needed for building a good model. EDA gives information which is used for identifying any errors in the dataset. Then, popularity based recommendation system makes use of the things that are popular at the moment. It ranks products based on its popularity i.e., the rating count. If a product is highly rated then it is most likely to be ranked higher and hence will be recommended. As it is based on the products popularity, this cannot be personalized and hence same set of products will be recommended for all the users. In the next step we have to choose an algorithm to generate the recommendations. In this project, it is decided to use the KNN and SVD algorithms. This is mainly because there is a significant amount of data for training purposes that is available in comparison to features or columns of the dataset. The hybrid algorithms are used to find similarities on the other products features and their ratings, and similar purchases done by the other users. Then, the recommendation using KNN and SVD separately. In order to make an accurate prediction of the users' opinions regarding a particular film, we shall employ a form of collaborative filtering known as hybrid collaborative filtering. To be more specific, we will integrate the results that were acquired using the KNN algorithm and the model based SVD technique. We are going to conduct an analysis on the hybrid model to determine whether or not the combination of a model based (SVD) and a memory based (KNN) approach produces superior outcomes compared to either of the techniques when used on their own.

**6. Experimental Setup**

In the following paragraphs, we will discuss the technique for setting up the assessment of frameworks, as well as the datasets and frameworks that were utilized.

Popularity Based Recommendationis a sort of advice that operates on the basis of popularity as well as everything else that is in vogue at the present time. These systems conduct research to determine which items, such as movies or products, are currently popular or are the most well-liked among its clients, and then they make direct recommendations for those items. As the name suggests it recommends based on what is currently trending/popular across the site. This is particularly useful when we don't have past data as a reference to recommend product to the user. It is not tailor fit for any particular group of audience. Every single user would receive recommendations for the exact same things, and the algorithm's criteria for making these recommendations would be solely based on how popular the products are.

Collaborative filtering using KNN’s goal is to leverage user evaluations of particular documents (content) to promote such documents to other users without actually assessing the content of the individual documents. This is accomplished through the utilization of user ratings. The calculation of correlations is the foundation of Pearson's similarity; the only user currents that are considered are those that are statistically significant.

 ……….(2)

In most cases, we make use of the K-Nearest Neighbour, or K-NN, method to figure out which neighbors are the most pertinent ones to select and create credible suggestions. This technique lets us choose only the k best neighbors who have the highest correlation value able to differentiate between two different approaches of collaborative filtering that are based on memory: one method that is item-centered memory, and another method that is user-centered memory based.

Collaborative Filtering using SVD method does not explicitly generate predictions; rather, it divides users into categories or trains models from their data based on how they interact with the system. These groups and models are determined by how the users interact with the system. The model serves as the foundation for the development of algorithms that take into account the historical assessments (profiles) of users. During the course of the construction of the model, a number of distinct methodologies are utilized. In most cases, the procedures that are based on the model make use of the machine learning strategies of clustering, matrix factorization, bayesian networks, and decision trees.

Hybridization can be broken down into two distinct stages: In order to create candidate recommendations, you should independently perform collaborative filtering as well as any other filtering approaches. To develop the final suggestions for consumers, combine these early sets of recommendations by employing some strategies, such as weighing, mixing, cascading, and switching. In a broader sense, hybrid systems are responsible for managing content-oriented user profiles. The contrast between these profiles ultimately results in the formation of user communities, which in turn makes collaborative filtering possible.

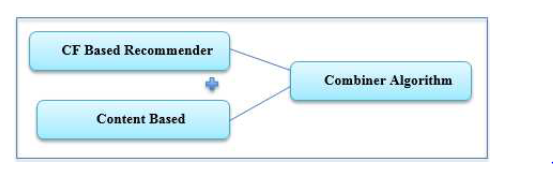


Fig 3 Hybrid Filtering Techniques

In order to make an accurate prediction of the users' opinions regarding a certain film, we are going to make use of a hybrid collaborative filtering strategy. Specifically, we are going to integrate the findings obtained by the model based SVD method and the k closest neighbour technique. The collaborative filtering algorithms have the benefit of not requiring any prior knowledge about the features of the items being filtered. Therefore, we are free to disregard the movie tags as well as the demographic statistics, and place more of an emphasis instead on the users and the ratings that they have given. We are going to conduct an analysis on the hybrid model to determine whether or not the combination of a model based (SVD) approach and a memory based (KNN) method produces superior outcomes compared to either of the techniques when used on their own.

1. **Result and Analysis**

The act of collecting datasets includes the process of gathering datasets regarding product recommendations. The dataset needed for the project has been compiled, and the results can be seen in the figure that can be found below:

This figure shows the rating distribution of the products.

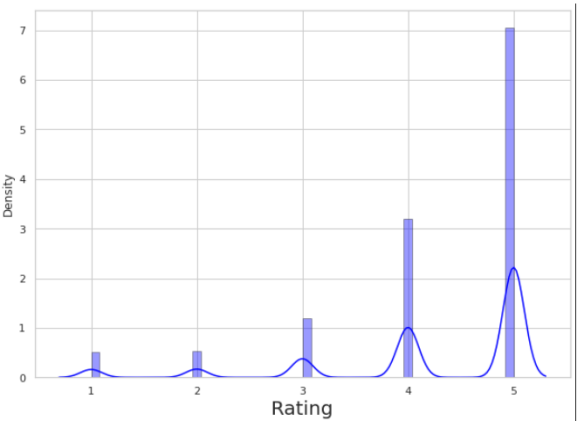


Fig 4 Rating Distribution

This figure shows the top rating count distribution grouped by all the products.

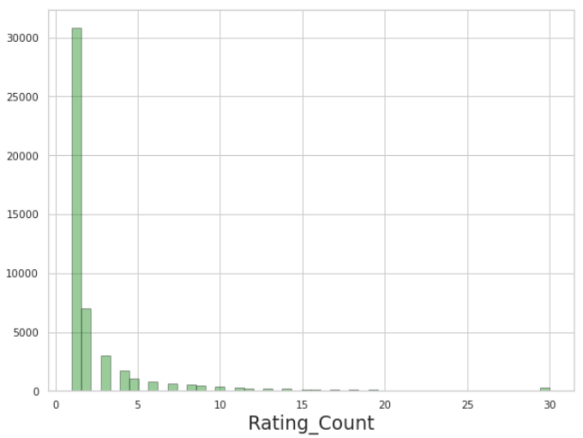


Fig 5 Top Rating Count Distribution

This figure shows the top rating count distribution grouped by all the users.

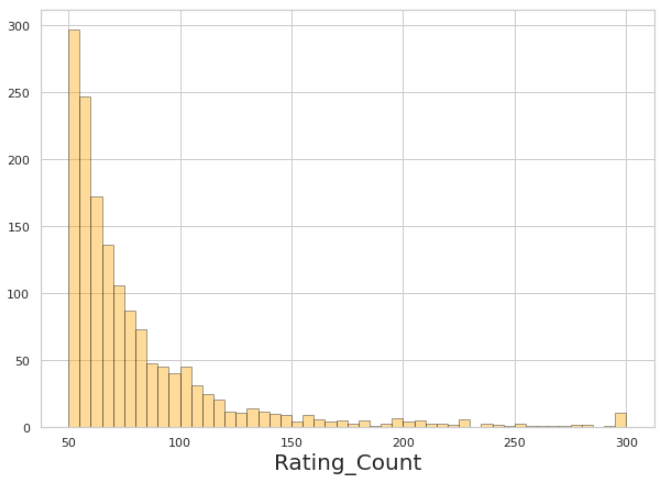


Fig 6 Top Rating Count Distribution Grouped By Users

This figure shows the mean rating distribution grouped by all the Products. There are many products which were given a Rating of 4 or 5.

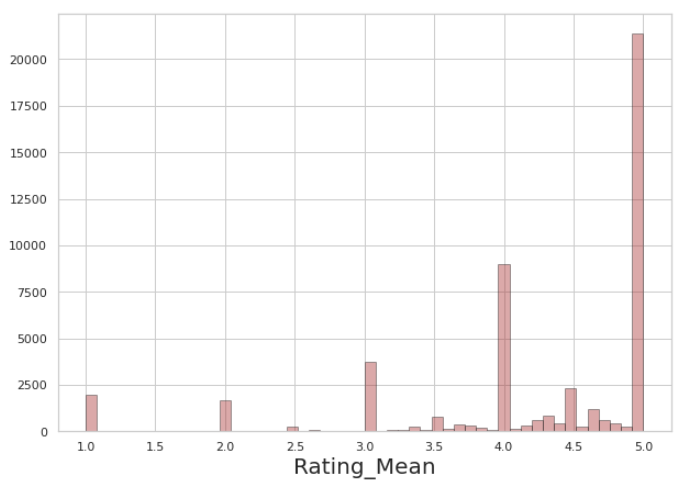


Fig 7 Mean Rating Distribution Grouped By Products

This figure shows the mean rating distribution grouped by the entire product. Data is clustered between mean ratings of 3.5 to 5.

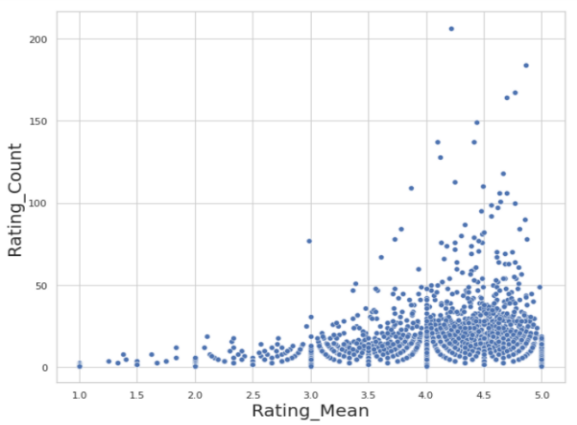


Fig 8 Rating Count Distribution Grouped By Products

This figure shows mean rating distribution grouped by all the Users. This looks like slightly left skewed normal distribution. Data is uniformly distribution from rating 3 to 5 with many around 4 to 4.5. Many users have an average rating of around 4.5.

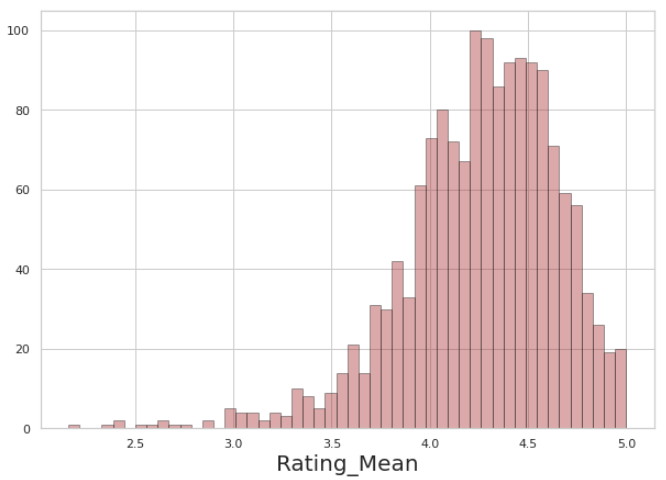


Figure 9 Mean Rating Distribution Grouped By Users

This figure shows the mean rating count distribution grouped by all users.

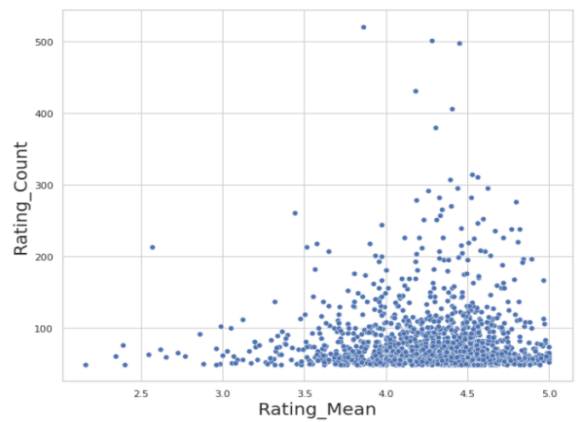


Fig 10 Mean Rating Count Distribution Grouped By Users

This figure shows the popularity based recommendation system uses the items. It ranks products based on its popularity the rating count. If a product is highly rated then it is most likely to be ranked higher and hence will be recommended.

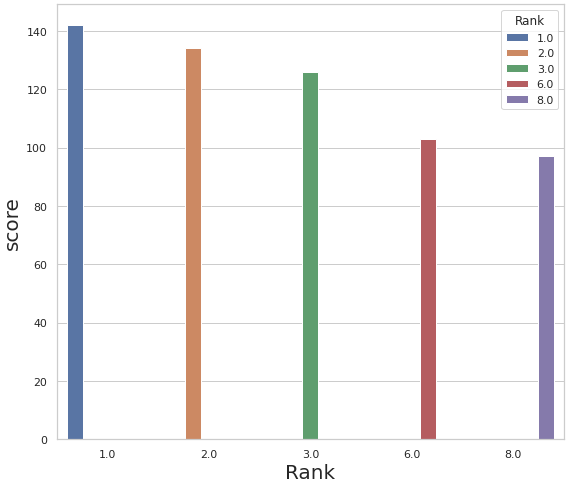


Fig 11 Popularity Based Recommendation

This figure shows the product recommendation based on the particular user id.



Fig 12 Popularity Based Recommendation Output

This figure shows the collaborative filtering is a method of making predictions about interests of the users. Here we used as KNN with Means algorithm.

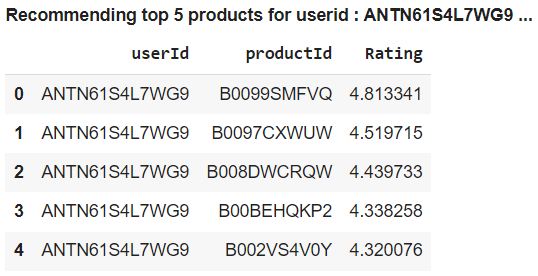


Fig 13 KNN With Means Memory Based Collaborative Filtering

This figure shows the approach of developing predictions about the interests of the users is referred to as collaborative filtering. Here we used an SVD algorithm.

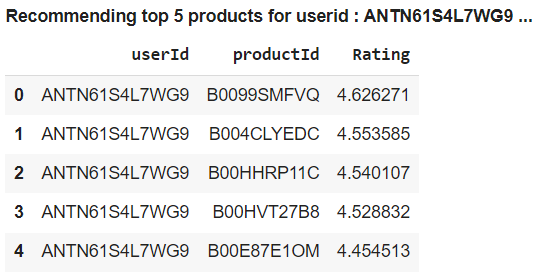


Fig 14 SVD Model Based Collaborative Filtering

This figure shows the we also used other algorithms (Base line, KNN BaseLine, KNNWithZScore, CoClustering, KNNBasic, NMF) for product recommendation based on user rating.

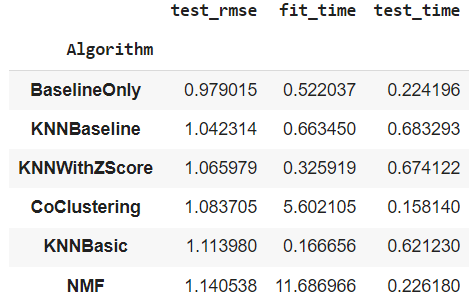


Fig 15 Other Surprise Algorithm

This figure shows the hybrid recommendation model (KNN With Means and SVD) for product recommendation based on the user rating.

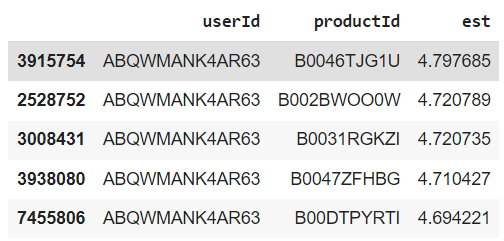


Fig 16 Hybrid Recommendation Model (KNN with Means and SVD)

**8. Conclusion**

In conclusion, our integrated model provides a very high degree of precision, and trials suggest that superior outcomes can be obtained from using it. This indicates that the vast majority of the suggestions made are applicable. There have been many different recommendation systems proposed, each of which is utilizing techniques such as collaborative filtering, content analysis, and hybrid recommendation approaches such as KNN and SVD are the finest techniques which ensures accuracy in the achieved output and algorithm has a highest quality guesswork, and one may be certain that it will be applied for provide recommendations to the user.

KNN (K-Nearest Neighbours) With Means model has a test RMSE value of 1.04 and cross validation RMSE value of 1.037. Each user will have different products recommended to them as they are inferred based on the ratings provided by the similar users. SVD (Singular Value Decomposition) model has a test RMSE value of 0.99 and cross validation RMSE value of 0.981. According to the result obtained, SVD performed better compared to KNN or Popularity with a better RMSE value of 0.981. We can also see the products that are recommended in SVD is different to that of KNN With Means as SVD uses matrix factorization. This is more useful when the data is sparse with many missing ratings. But when we combine both KNN-Means and SVD by hybrid filtering, the rmse value is even more reduced as compared to SVD which then recommends precise product suggestions to the users.

The future work of the project includes improving the efficiency of the system. And it should also be able to give appropriate recommendations to the users who don’t have any previous purchase history or to the new users. We have reviewed the application of recommending the products in the e-commerce field with a higher degree of precision. There is a greater potential to grow or transform this idea into many different forms within this discipline. Utilizing a variety of time-saving strategies and algorithms will result in an increase in the prediction's level of precision.

**References**

[1]Badr Hssina, Abdelakder Grota, Mohammed Erritali(2017), “Recommendation System Using the K-Nearest Neighbors and Singular Value Decomposition Algorithms”, International Journal of Electrical and Computer Enginnering.

[2]”Recommender Systems Based on Collaborative Filtering Using Review Texts” Mehdi Srififi, Ahmed Oussous, Ayoub Ait Lahcen, Salma Mouline (2016), International Journal of Electrical and Computer Enginnering.

[3]”Building a Movie Recommendation Systems Using SVD algorithm” (2018), Asoke Nath, Adityam Ghosh, Arion Mitra, International Journal of Electrical and Computer Enginnering.

[4]”Location Based Agricultural Product Recommendation System Using Novel KNN Algorithm” (2019), Sachin J, Geethatharani P, Surya M K, Kavin K V, International Journal of Scientific Research in Computer Science Enginnering and Information Technology.

[5]”Sentiment Analysis for Product Recommendation System Using Hybrid Algorithm” (2019), Umamaheswari, G.Kanimozhi International Journal of Scientific Research in Computer Science Enginnering and Information Technology.

[6]”RecDNNing: A Recommender System Using Deep Neural Network with User and Item Embeddings” (2019), Hafed Zarzour, Ziad A. Al-Shairf and Yaser Jararweh, [International Conference on Information and Communication Systems (ICICS)](https://ieeexplore.ieee.org/xpl/conhome/8794646/proceeding).

[7]A. A. Müngen and M. Kaya, “A Novel Method for Event Recommendation in Meetup,” in 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 2017.

[8] E. Gündoğan, B. Kaya, and M. Kaya, “Prediction of SymptomDisease Links in Online Helath Forums,” in IEEE/ACM International Conference on Advances in Social Networks and Mining (ASONAM, 2017), 2017.

[9] S. Aslan and M. Kaya, “Link prediction methods in bipartite networks,” in Computer Science and Engineering (UBMK), 2017 International Conference on, 2017.

[10]W. Zhao, R. Wu, and H. Liu, “Paper recommendation based on the knowledge gap between a researcher’s background knowledge and research target,” Inf. Process. Manag., vol. 52, no. 5, pp. 976–988, 2016.

[11]W. Zhang, T. Yoshida, and X. Tang, “A comparative study of TF\*IDF, LSI and multi-words for text classification,” Expert Syst. Appl., vol. 38, no. 3, pp. 2758–2765, 2011.

[12]J. Ramos, “Using TF-IDF to Determine Word Relevance in Document Queries,” Proc. first Instr. Conf. Mach. Learn., pp. 1– 4, 2003.

[13]T. Kenter and M. de Rijke, “Short Text Similarity with Word Embeddings Categories and Subject Descriptors,” Proc. 24th ACM Int. Conf. Inf. Knowl. Manag. (CIKM 2015), pp. 1411– 1420, 2015.

[14]C.-H. HUANG, J. YIN, and F. HOU, “A text similarity measurement combining word semantic information with TF-IDF method,” in Chinese Journal of Computers, 2011.

[15]S. Albitar, S. Fournier, and B. Espinasse, “An Effective TF/IDFBased Text-to-Text Semantic Similarity Measure for Text Classification,” Web Inf. Syst. Eng. - Wise 2014, Pt I, vol. 8786, pp. 105–114, 2014.

[16]ZhongqiLu, ZhichengDou, JianxunLian, XingXie and QiangYang(2015),” Content-Based Collaborative Filtering for News Topic Recommendation”, Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence.

[17]Narsinga rao,M.R.,Venkatesh Prasad ,V.,Sai Teja,P.,Zindavali,M.&phanindraReddy,O.(2018),”A survey on Preventio of Overfitting in convolutional neura networks using machine learning techninques”International journal of engineering and Technology

[18]ParulAggarwal, Vishal Tomar, AdityaKathuria(2017),” Comparing Content Based and Collaborative Filtering in Recommender Systems”,International Journal of New Technology and Research.

[19]Surarchita,V.,VenkateswaraRao ,P.,battacharyya ,D.,Kim,T.(2016) ,”Classification of penaeid prawn species using radial basis probabilistic neural networks and supportvectormachines”International Journel of Bio-Science and Bio-technology.

[20]R. Torres, S. M. McNee, M. Abel, J. A. Konstan, and J. Riedl, “Enhancing digital libraries with TechLens+,” Proc. of the 4th ACM/IEEE-CS Joint Conference on Digital Libraries, pp.228-236, 2004.