International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 05 Issue: 12 | Dec - 2021 ISSN: 2582-3930

Product Recommendation System Using Machine Learning

Pratiksha Abhang[1],Shwetali Gorde[2], Rupali Saundade[3]

1-3 Students, Department of Computer Engineering, Jaihind College of Engineering, Kuran, Pune

---

Abstract - Recommendation systems have emerged rapidly in the past decade. Recommendation of products to attract customers that meet their requirements is very important for the vendors to survive in the global market. The approach proposed in this paper is novel and serves as a better alternative to rate a product based on its technical specification by analysing large number of user reviews which are extracted dynamically from several top e-commerce websites. This avoids the need, for the user, to search for opinions and comments online before making a purchase. The proposed approach in this study extracts specification list like battery, processor, camera etc. and customer reviews for a user specified product from different websites and identifies crucial terms corresponding to the technical features of the product in the review to determine polarity of the feature and classifying it under the specification list. Each specification is assigned a score based on polarity i.e. positive/negative feedback. Overall product rate is calculated by aggregating the score specific to individual features. This approach is very useful for those customers who target at specific features in a product.

***Key Words*:** Recommended System, Feature Extraction, Classification, Support vector Machine(SVM), Harr cascade, Machine Learning etc.

1. INTRODUCTION

Consumer reviews, opinions and shared experiences in the use of a product is a powerful source of information about consumer preferences that can be used in recommendersystems. Despite the importance and value of such information, there is no comprehensive mechanism that formalizes the opinions selection and retrieval process and the utilization of retrieved opinions due to the difficulty of extracting information from text data. In this paper, a new recommender system that is built on consumer product reviews is proposed. A prioritizing mechanism is developed for the system. The proposed approach is illustrated using the case study of a recommender system for digital cameras. E-Commerce is the easiest, most commodious way of organizing business over the internet for business experts and individuals. Managing business over the internet is simply surfing specific website for shopping products online or business related matters. One of the major objectives of this research work is to find out the reliable sales trend prediction mechanism which is implemented by using data mining techniques

to achieve the best possible revenue. Accurate predictions allow the organization to improve market growth with higher level of revenue generation. Data mining techniques are very effective in tuning huge volume of data into useful information for cost prediction and sales forecast, it is thebasic of sound budgeting.

1.1 Motivation

Recommendation systems help users find and select items (e.g., books, movies, restaurants) from the huge number available on the web or in other electronic information sources. Given a large set of items and a description of the user's needs, they present to the user a small set of the items that are well suited to the description. Recent work in recommendation systems includes intelligent aides for filtering and choosing web sites, news stories , TV listings, and other information. The users of such systems often have diverse, conflicting needs. Differences in personal preferences, social and educational backgrounds, and private or professional interests are pervasive. As a result, it seems desirable to have personalized intelligent systems that process, filter, and display available information in a manner that suits each individual using them. The need for personalization has led to the development of systems that adapt themselves by changing their behavior based on the inferred characteristics of the user interacting with them. The ability of computers to converse with users in natural language would arguably increase their usefulness and flexibility even further.

1.2 Problem Statement

Systems have emerged rapidly in the past decade. Recommendation of products to attract customers that meet their requirements is very important for the vendors to survive in the global market. The approach proposed in this paper is novel and serves as a better alternative to rate a product based on its technical specification by analyzing large number of user reviews which are extracted dynamically from several top e-commerce websites. This avoids the need, for the user, to search for opinions and comments online before making a purchase.

2. MODULES

i) **Preprocessing-**

The aim of pre-processing is an improvement of the image data that suppresses unwilling distortions or enhances some image features important for further processing, although geometric transformations of images (e.g. rotation, scaling, translation) are classified among pre-processing methods here since similar techniques are used.

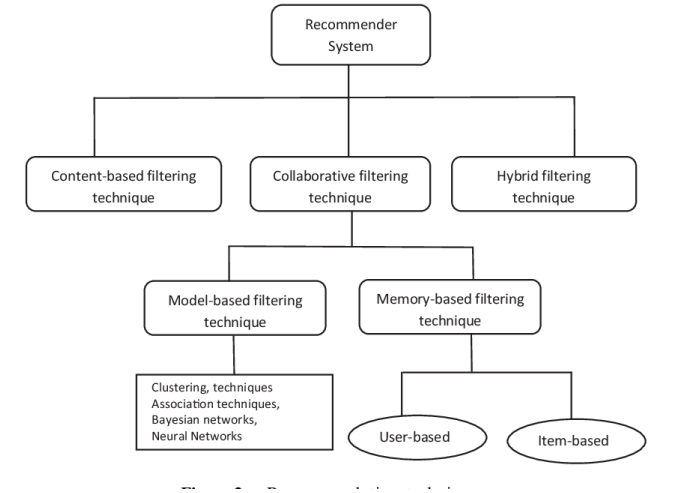
ii) Feature Extraction-

Feature extraction is a part of the dimensionality reduction process, in which, an initial set of the raw data is divided and analyze by SVM algorithm. So when you want to process it will be easier. The most important characteristic of these large data sets is that they have a large number of variables.

iii) Classification-

"Support Vector Machine" (SVM) is a supervised machine learning algorithm which can be used for both classification and regression challenges. However, it is mostly used in classification problems. In the SVM algorithm, we plot each data item as a point in n-dimensional space (where n is number of features you have) with the value of each feature being the value of a particular coordinate. Then, we perform classification by finding the hyper-plane that differentiates the two classes very well.

3. RECOMMENDATION TECHNIQUES



A. Content-based:

The system learns to recommend items that are similar to the ones that the user liked in the past. The similarity of items is calculated based on the features associated with the compared items. For example, if a user has positively rated a movie that belongs to the comedy genre, then the system can learn to recommend other movies from this genre.
CB filtering is most simple and natural approach to adopt as a recommendation it does not require any feedback from the user. Sometimes a single preference is enough to recommend many items to the user. This approach also extends naturally to the cases where item information is well organized and available such as movies, songs, products, and books. But at the same time, it also comes as a limitation of content-based filtering as item description is not always present in all cases that create difficulty in measuring the similarity between items. These recommendation systems have limitations to produce similar results and are static over the time.

B. Collaborative filtering:

The simplest and original implementation of this approach recommends to the active user the items that other users with similar tastes liked in the past. The similarity in taste of two users is calculated based on 12 Francesco Ricci, Lior Rokach and Bracha Shapira the similarity in the rating history of the users. This is the reason why refers to collaborative filtering as "people-to-people correlation." Collaborative filtering is considered to be the most popular and widely implemented technique in RS.

CF is the most popular and used recommendation technique. The basis for collaborative filtering is that users with similar interest are inclined to give same preference for the new and future items. This technique works on two points. First, it serves as a criterion to select a group of similar people whose opinions will be accumulated as a basis for a recommendation (nearest neighbors). Second, it also uses these opinions to form a bigger group and have a greater impact on the recommendation Collaborative filtering techniques involved very large data sets and applied in diverse application areas like finance, weather forecasting, environmental sensing, e-commerce.

1.Memory-Based Filtering
Memory-based methods are straightforward and easy to implement. The best-known technique used is memory-based neighborhood-based filtering, which predicts preferences by referring to users who are similar to queried user or to items that are similar to queried item. The accuracy and efficiency of neighborhood technique are profoundly affected by how the similarity between users or items is calculated. This technique can be further extended with default votes, inverse user frequency, and case amplification. These techniques are further classified as user-based filtering techniques and item-based filtering technique.
User-Based Filtering technique computes the similarity between users by comparing their preference over the same item and calculates the predicted preference for items for the active user .
Item-Based Filtering techniques calculate predictions using similarity between items. The technique works by retrieving all the items rated by an active user and determines similarities of retrieved items with target item. It then selects top N most similar items to predict the preference of the active user for the target item.

2.Model-Based Filtering
Model-based techniques make use of data mining and machine learning approaches to predict the preference of a user to an item. These techniques include association rule, clustering, decision tree, artificial neural network, Bayesian classifier, regression, link analysis, and latent factor models. Among these latent factor models are the most studied and used model-based techniques. These techniques perform dimensionality reduction over user–item preference matrix and learn latent variables to predict preference of the user to an item in the recommendation process. These methods include matrix factorization, singular value decomposition, probabilistic matrix factorization, Bayesian probabilistic matrix factorization, low-rank factorization, nonnegative matrix factorization, and latent Dirichlet allocation.

3.Hybrid Filtering
Some applications combine the advantages of memory-based and model-based approaches to form a hybrid filtering system. It results in better prediction and efficiency. A proper combination can overcome the limitation of collaborative filterings such as sparsity and diversity.

C. Knowledge-Based Filtering

Knowledge-based filtering uses back-end knowledge or information of users, items, and their relationship. These systems describe how a particular item meets the requirement of the user. The technique requires domain-specific knowledge of users and items. The most traditional knowledge-based system is the case-based system.

Knowledge-based systems tend to work better than others at the beginning of their deployment but if they are not equipped with learning components they may be surpassed by other shallow methods that can exploit the logs of the human/computer interaction (as in CF).

D. Hybrid Filtering

Hybrid filtering techniques is one which combines the advantages of two or more filtering techniques and overcomes their limitations. These techniques provide more effective and enhance results of recommendation. Hybrid techniques can adopt one of the following strategies to develop a hybrid filtering method:

1. Use content-based and collaborative-based filterings to produce separate recommendation, and then use a linear combination of this two recommendations to provide a single recommendation.

2. Collaborative filtering can be used with content-based characteristics to calculate the similarity between users and find out neighbors to predict the recommendation.

3. Content-based techniques can be added to collaborative filtering characteristics, such as latent factor model with the content-based approach.

4. A conventional probabilistic method for combining collaborative and content based technique to predict recommendation.

4. ALGORITHM

4.1 SVM Algorithm

"Support Vector Machine"(SVM) is a supervised machine learning algorithm which can be used for both classification or regression challenges. However, it is mostly used in classification problems. In the SVM algorithm, we plot each data item as a point in n-dimensional space (where n is number of features you have) with the value of each feature being the value of a particular coordinate. Then, we perform classification by finding the hyper-plane that differentiates the two classes very well (look at the below snapshot). Support Vectors are simply the co-ordinates of individual observation. The SVM classifier is a frontier which best segregates the two classes (hyper-plane/ line). It works really well with a clear margin of separation It is effective in high dimensional spaces.It is effective in cases where the number of dimensions is greater than the number of samples. It uses a subset of training points in the decision function (called support vectors), so it is also memory efficient.

4.2 Haar cascade algorithm

Face Detection, a widely popular subject with a huge range of applications. Modern day Smartphones and Laptops come with in-built face detection softwares, which can authenticate the identity of the user. There are numerous apps that can capture, detect and process a face in real time, can identify the age and the gender of the user, and also can apply some really cool filters. The list is not limited to these mobile apps, as Face Detection also has a wide range of applications in Surveillance, Security and Biometrics as well.It is an Object Detection Algorithm used to identify faces in an image or a real time video.

5. EXISTING SYSTEM

1. We consider all the reviews and ratings which is generated by the user and main purpose is to analyze semantic aspects and aspect level sentiments from review data as well as to predict overall sentiments of reviews.

2. All the time it is necessary to analyze and identify review. Sometime It become more fuzzy.

6. PROPOSED SYSTEM

1. In Our proposed system we will provide security for account accessing using face authentication.

2. We will extract customer reviews and specification list for the user selected product. Used Machine learning Algorithm to classify and trained datatset of Products Reviews/Rating.

3. After processing, the system will calculate score for each feature based on its polarity in the extracted review and overall rating of a product is calculated by aggregating individual feature scores.

4. Finally, the system will output the pros and cons of the user selected product in terms of category and subcategory results are displayed on screen for providing better understanding for the user.

7. CONCLUSIONS

Understanding the buying intention of customers on e-commerce sites. Understanding the satisfaction of the customer with specific products through online reviews
on e-commerce sites. Recommending products and places to customers based on collective opinions.

ACKNOWLEDGEMENT

We would like to take this opportunity to thank our project guide Prof. Pawar S.H for giving us all the help and guidance we needed. We are really grateful to them for their kind support. Their valuable suggestions were very helpful.
We are also grateful to thank, our other faculty members from the Computer Engineering Department, Jaihind college Of engineering, Kuran for allowing us to perform our project.

REFERENCES

[1]. J. McAuley, et al., "Image-based recommendations on styles and substitutes," in Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 43-52, 2015.
[2]. X. Ning, et al., "A comprehensive survey of neighborhood-based recommendation methods," in Recommender systems handbook, ed: Springer, pp. 37-76, 2015.
[3]. R. He and J. McAuley, "Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering," in proceedings of the 25th international conference on world wide web, pp. 507-517, 2016.
[4]. Bao, Jie, et al. "Recommendations in location-based social networks: a survey." GeoInformatica 19.3, pp. 525-565, 2015.
[5]. Wang, Hao, Naiyan Wang, and Dit-Yan Yeung. "Collaborative deep learning for recommender systems." Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2015.
[6].Recommendation Systems: Techniques,Challenges, Application, and EvaluationSandeep K. Raghuwanshi and R. K. Pateriya
[7]. S. S. A. K. L. Shunmuganathan, "Role of Agent Technology in Web Usage Mining : Homomorphic Encryption Based Recommendation for E-commerce Applications," Wirel. Pers. Commun., 2015.
[8]. L. Yan, "Personalized recommendation method for E-commerce platform based on data mining technology," Proc. - 2017 Int. Conf. Smart Grid Electr. Autom. ICSGEA 2017, vol. 2017-Janua, pp. 514– 517, 2017.
[9] . P. Aggarwal, T. Vishal, and A. Kathuria, "Comparing Content Based and Collaborative Filtering in Recommender Systems," Int. J. New Technol. Res., vol. 3, no. 4, pp. 65–67, 2017.
[10] . A. U. Martliong and N. M. S. Iswari, "Rancang Bangun Sistem Rekomendasi Restoran Menggunakan Metode AHP dan VIKOR pada Platform LINE," J. Ultim. Comput., vol. 10, no. 1, pp. 27–33, 2018.
[11] . Z. Wang, J. Liao, Q. Cao, H. Qi, and Z. Wang, "Friendbook: A semanticbased friend recommendation system for social networks," IEEE Trans. Mob. Comput., vol. 14, no. 3, pp. 538–551, 2015.
[12] . L. Cui, H. li, C. Wang, and B. Yang, "Personalized Book Recommendation Based on Ontology and Collaborative Filtering Algorithm,"OpenCybern. Syst. J., vol. 8, no. 1, pp. 632–637, 2018.
[13] A. Lawrence, "Five Customer Retention Tips for Entrepreneurs," Forbes, 2012. [Online]. Available: http://www.forbes.com/sites/alexlawrence/2012/11/01/five-customerretention-tips-for-entrepreneurs/210fc5e17b0b. [Accessed: 12-Apr2016].
[14] C. Shaw, "15 Statistics That Should Change The Business World – But Haven't,"
Linked in, 2013. [Online]. Available: https://www.linkedin.com/pulse/20130604134550-284615-15-statisticsthat-should-change-the-business-world-but-haven-t. [Accessed: 12-Apr2016].
[15] "The case for customer service training," White House Office of Consumer Af-fairs, 2009. [Online]. Available: http://aspiremarketing.com/the-case-for-customer-satisfaction/. [Accessed: 12-Apr-2016].
[16] J. Hauser and E. Dahan, "New Product Development," in Marketing Management:
Essential Marketing Knowledge and Practice, R. Grover and N. K. Malhotra, Eds. McGraw Hill, Inc., Columbus Ohio, 2007.
[17] M. D. Earle and R. L. Earle, Creating New Foods. The Product Developer's Guide.
The New Zealand Institute of Food Science Technology (Inc.), 2009.