iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** V **Month of publication:** May 2020

DOI: http://doi.org/10.22214/ijraset.2020.5441

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue V May 2020- Available at www.ijraset.com

Recommendation Systems using Machine Learning Approach

Farhana Siddiqui[1], Uzma Unagar[2], Sadaf Abdul Qadir[3], Sara Shaikh[4]
[1]Professor, [2, 3, 4]B.E. Student, Department of Computer Engineering, M.H Saboo Siddik COE, Maharashtra, INDIA

Abstract: In the advanced age, each association is transforming its business into online business. Online organizations utilize proposal systems to channel the data and prescribe most relevant things to its clients. In this paper we study different types of recommendation systems and then propose a model of a product-based recommendation engine for a business.
Keywords: Recommendation, Collaborative filtering, Model based, Memory based, Content based, Hybrid

I. INTRODUCTION

The rapid expansion of the World Wide Web is leading to increasing complexity of knowledge present online. Today, the online world contains an outsized amount of knowledge, most of it doesn't relate to the user, it is present either as unwanted information or as content irrelevant to his interests. Every user of the online world has unique interests which can correspond with some percentage of web's content. To assist users find the knowledge that's in accordance with their interests the web will be personalized, using recommender systems. Recommendation systems are a special kind of information filtering systems. They recommend items of interest to the users supported preferences they need expressed, either explicitly or implicitly.

II. TYPES OF RECOMMENDER SYSTEM

A. Collaborative Recommender System:

Collaborative filtering systems work by way of accumulating user statements in the form of ratings for objects in a given discipline and exploiting similarities in score actions amongst numerous customers in figuring out how to propose an object.[1] Collaborative filtering structures recommend an object to a user based on evaluations of different users. Like, in a film recommendation application, Collaborative filtering device attempts to locate different like-minded users and then recommends the movies which might be most favored via them. There is a vast variety of collaborative filtering techniques. These techniques can be classified majorly into two categories:

1) *Memory based Approach:* Memory-based algorithms utilize the whole user-item database to generate a prediction. These structures employ statistical techniques to find a group of customers, referred to as neighbors, which have records of agreeing with the goal user. Here, 'agreeing' means that they either rate specific gadgets similarly or they tend to shop for comparable sets of objects. Once a neighborhood of customers is formed, these structures use exclusive algorithms to mix the possibilities of neighbors to produce a prediction or top-N recommendation for the active user. The strategies, also referred to as nearest-neighbor or user-primarily based collaborative filtering are more popular and widely used in practice.
2) *Model based Approach:* In model-primarily based CF algorithms, a theoretical model is proposed of user score behavior. Rather than use the uncooked rating records immediately in making predictions, rather the parameters of the version are predicted from the to-be had score information and the model is used to make predictions. Many model-primarily based CF algorithms had been studied during the last years. For example, discusses two probabilistic models, namely, clustering and Bayesian networks. In 4 partitioning-based totally clustering algorithms are used to make predictions, main to higher scalability and accuracy in evaluation to random partitioning [2].

B. Content Based Recommender System

Implementation of content based recommendation systems require analysis of sets of documents or descriptions of items that the user has rated in the past, then building a profile or model of interest of the user based on the characteristics of objects rated by the user. The advice process basically is composed in matching up the attributes of the consumer profile against the attributes of a content object. The end result is a relevant judgment that represents the consumer's level of hobby in that object. If a profile efficiently displays person preferences, it is of tremendous advantage for the effectiveness of an information access process [3].

Content based feature selection can be done through wrapper methods, filter methods or embedded methods. Wrapper strategies evaluate one-of-a-kind subsets of features by schooling a model for each subset after which evaluating every subset's contribution on a validation dataset.

As the number of all viable subsets is factorial inside the wide variety of capabilities, one-of-a-kind heuristics are used to choose “promising" subsets (forward-selection, backward-elimination, tree-induction, etc.) Wrapper strategies are independent of the prediction algorithm .Filter techniques are normally primarily based on heuristic measures, which include Mutual Information or Pearson Correlation, to score features based totally on their information contents with recognition of the prediction task. Similar to wrapper strategies, filter strategies are also independent of the algorithm in use. However, they do not require training many models and therefore scale well for large datasets. [4] Embedded techniques are a circle of relatives of algorithms wherein the characteristic selection is performed within the direction of the training phase. Unlike wrapper techniques, they're not based totally on cross-validation and therefore scale with the dimensions of the data. However, in view that the feature choice is an inherent asset of the algorithm, an embedded technique is tightly coupled with the unique model.

The techniques of content based approach includes TF-IDF, Naive Bayes, etc. The terms that occur frequently in one document ie TF or term-frequency, but have less frequency within the remainder of the corpus i.e. IDF or inverse-document-frequency, have a higher chance to be relevant to the subject of the document. Additionally, normalizing the resulting weight vectors prevent longer documents from having a higher chance of retrieval. [4] Naive Bayes is a probabilistic technique to inductive learning, and belongs to the overall class of Bayesian classifiers. These procedures generate a probabilistic model based totally on previously located data.

C. Knowledge Based Recommender System

This sort of recommender framework endeavors to propose objects dependent on inductions about a client's needs and inclinations. Information put together proposal works with respect to utilitarian information: they know about how a specific thing meets a specific client need, and can in this way reason about the connection between a need and a potential suggestion or recommendation.

D. Hybrid Recommender System

The previously mentioned recommendation engine techniques have different strengths and some limitations. Problems of cold start and sparsity affect the collaborative filtering techniques, whereas narrowness effects content-based approaches and hence they require descriptions. The hybrid recommender systems combine the strengths of multiple algorithms, so as to overcome the weakness or limitations of previous algorithms. These systems are designed to make predictions where the other fails, resulting in a more robust recommender system [5]. There are a few manners by which the frameworks can be consolidated.

1) *Weighted Hybrid Recommender:* In this system the score of an advocated object is computed from the consequences of all of the to-be had recommendation strategies present within the system. For example, P-Tango machine combines collaborative and content based totally recommendation systems giving them same weight in the starting, but gradually adjusting the weighting as predictions about the user scores are showed or disconfirmed. Pazzani’s combination hybrid doesn’t use numeric scores but as a substitute treats the output of every recommender as a hard and fast of votes, which might be then combined in a consensus scheme.
2) *Switching Hybrid Recommender:* These Hybrid Recommender, switches between the proposal methods dependent on specific models. Assume that we consolidate the substance and community based recommender frameworks, at that point, the switching hybrid recommender would first be able to convey content based recommender framework and in the event that it doesn't work, at that point it will send collaborative based recommender framework.
3) *Mixed Hybrid Recommender:* Where it's conceivable to make an enormous number of proposals all the while, we ought to opt for mixed recommender frameworks. Here suggestions from more than one strategy are introduced together, so the client can look over a wide scope of proposals. The PTV framework, principally a prescribed program to propose clients for TV seeing, created by Smyth and Cotter is utilized by most of the media and amusement organizations.

Hybrid recommender systems face some issues. The first problem is to reflect the collaborative and content-based statistics or data while making recommendations. A smooth solution is to apply collaborative and content-based methods in parallel or in cascade. However, such a technique has certain drawbacks. Although Meta recommender structures were proposed to pick out a recommender system among traditional ones on the premise of certain satisfaction measures, the disadvantages of the selected recommender are inherited. Moreover, the heuristics-primarily based integration treated in other research lacks a principled justification [6]. The next issue is the efficient adaptation of the recommender system with respect to the increase in ratings or scores and users. The simplest suggested solution is the use of memory based recommender systems, as they utilize entire data to make recommendations. However, this solution results in the late responses, trying to overcome this disadvantage by using a probabilistic method in a pure collaborative filtering context.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue V May 2020- Available at www.ijraset.com

III. PROPOSED SYSTEM

The proposed system deals with a new customer for the business and a collaborative technique for the regular customers of an online store. When a new customer accesses the online store for the first time, the recommender system will not have any information about the customer's likes and dislikes. In this scenario, the proposed system will use the popularity based approach and will recommend those items to the customer which are very popular in the specific store and are liked and purchased by the other regular users of the store. System will recommend items to the regular users based on their history of purchased items and the ratings provided to various items by similar users. The similarity between users will be computed using correlation. This will make use of collaborative filtering techniques. These techniques help in making predictions of recommended items for a particular user by identifying patterns based on preferences from multiple user data. If the business is new and does not have any user-item history of purchase, a text-based recommendation engine will be used. The item recommendations will be based on textual clustering analysis of the description of item.

IV. DESIGN DETAILS

The design describes a plan to implement the requirements. Various diagrams can be made to achieve a clear idea about the implementation.

A. Use Case Diagram

Use case diagrams describe a set of actions or use cases that a system can perform in collaboration with one or more external users.

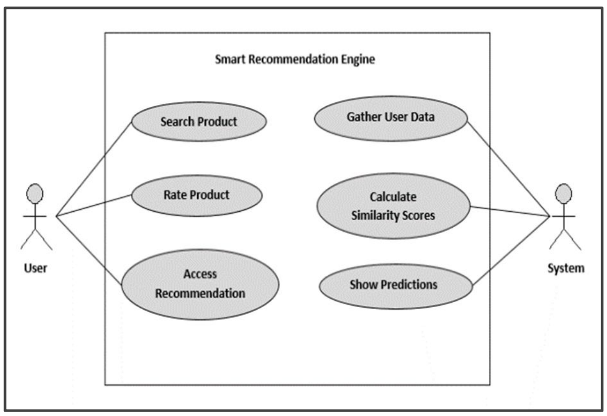

Fig. Use Case Diagram

Here the actors are the user and the system. Each use case comes with the description, actors, pre/post conditions and the flow of control. Flow of control would be the inherent activities of each use case like the processing of the request as soon as the user enters a query for searching the product, the compiling of the list generated.

B. Activity Diagram

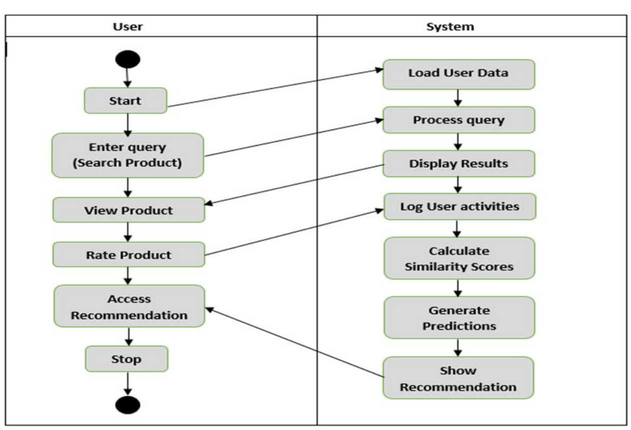

Fig. Activity Diagram

V. CONCLUSION

Through the proposed work, the customers can be assured with a better satisfaction as the related products are suggested as soon as they select a product to purchase as the recommendation algorithm contains various techniques of finding the similar products. This enhances customer's fulfillment and is also an improved way of marketing the products. This measures individuality of rating items with the reference of experienced users with various factors.

REFERENCES

[1] Tewari A.S., Kumar A., and Barman A.G.,. Book Recommendation System Based on Combine Features of Content Based Filtering, Collaborative Filtering and Association Rule Mining, IEEE, 978-1-4799-2572-8, pp 500-503, 2014.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
[2] Zunping Cheng, Neil Hurley," Effective Diverse and Obfuscated Attacks on Model-based Recommender Systems" 2009.
[3] Meenakshi Sharma and Sandeep Mann, "A Survey of Recommender Systems: Approaches and Limitations" ,ICAECE2013.
[4] A Review Paper on Machine Learning Based Recommendation System, 2014 IJEDR | Volume 2, Issue 4 | ISSN: 2321-9939.
[5] Svetlin Bostandjiev, John O"Donovan, Tobias Höllerer," TasteWeights:A Visual Interactive Hybrid Recommender System" 2012 ACM
[6] Kazuyoshi Yoshii, Masataka Goto, Kazunori Komatani, Tetsuya Ogata, Hiroshi G. Okuno," An Efficient Hybrid Music Recommender System Using an Incrementally Trainable Probabilistic Generative Model",IEEE 2008