iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** I **Month of publication:** January 2023

DOI: https://doi.org/10.22214/ijraset.2023.48662

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue I Jan 2023- Available at www.ijraset.com

Content-based Recommender System for Answering Web Based User Queries

Aayush Dani[1], Amitabh Dixit[2], Shailja Jadon[3], Shruti Dubey[4], Mr. Shailesh Sangle[5]
Computer Engineering Department, Thakur College of Engineering and Technology, Mumbai, India

Abstract: Online the idea of this paper is to give insights about a web-based project that aims to resolve the users query using machine learning technique of content-based filtering. The crux is to define the scope of the website. It becomes crucial to understand the business prospects of the project along with the relevance in the current technical world. As the paper goes about, various other papers and ideas will also be discussed that have been previously published and implemented. This paper will also break down the concept of content-based filtering for any individual who is a beginner in the field of machine learning.
Keywords: Content, recommendation, query, count, vectorization

I. INTRODUCTION

Early in the 1990s, recommender systems were developed, and they are now a big part of people's everyday life [1][2][3][4]. By examining client preferences and online activities, internet retailers like Netflix and Amazon.com propose products and services. Social networking sites could imply that you know some individuals or might want to make friends with someone. Additionally, interests-based social networking sites provide users suggestions for books, music CDs, movies, and articles as well as suggestions for individuals who may have similar preferences based on the ratings that users have given the aforementioned goods.

A recommender system mainly makes use of two types of data: user ratings for goods and/or user and/or item profiles. Only user ratings for things are used in memory-based recommendations, such as the user-item matrix, which lists user ratings for each entry. User-based recommendations and item-based recommendations are the two main categories of memory-based recommendations [6] [7]. In user-based recommendations, users with similar rating patterns to the current user are given larger weights, and these weighted users' utility ratings are used to determine the usefulness of new things.In order to determine an item's utility for a user, item-based recommendations first choose its most comparable things that the user has rated, after which they compute the utility as the weighted average of the ratings of these similar products. Item-based recommendation may be a workable solution since it computes the similarities between products offline. When the number of users in an online recommender system increases significantly, computing the similarity of every user pair's rating patterns would be exceedingly time consuming. A great, thorough overview of recommender systems is provided in Reference [8].

The vast amount of data available on the Internet has led to the development of recommendation systems. This project proposes the use of soft computing techniques to develop recommendation systems. It addresses the limitations of current algorithms used to implement recommendation systems, evaluation of experimental results, and conclusion. The goal of this project is to study recommendation engines and identify the shortcomings of traditional recommendation engines and develop a web-based recommendation engine by making use of context-based results for answering web-based user queries.

One prevalent approach for recommendation or recommender systems is content-based filtering. "Content" refers to the attributes or content of the items you enjoy. The algorithm utilizes your preferences to make suggestions for stuff you might enjoy. It takes the information you offer via the internet and the information they can obtain, and curates suggestions based on it.

The purpose of content-based filtering is to categorize products using keywords, learn what the client likes, search up those phrases in the database, and then suggest

comparable items. This form of recommender

The system is heavily reliant on human input, including Google, Wikipedia, and other well-known examples. When a user searches for a set of keywords, Google displays all the results that include those keywords.

The user-item matrix and the research papers' content details would both be used in the recommendation of research papers. We analyze research articles' subjects using topic model methodologies, and we categorize the similarities in topics as thematic similarity. We were able to produce highly relevant suggestions and significantly reduce the cold start issue by adding thematic similarity and a modified item-based technique.

©IJRASET: All Rights are Reserved

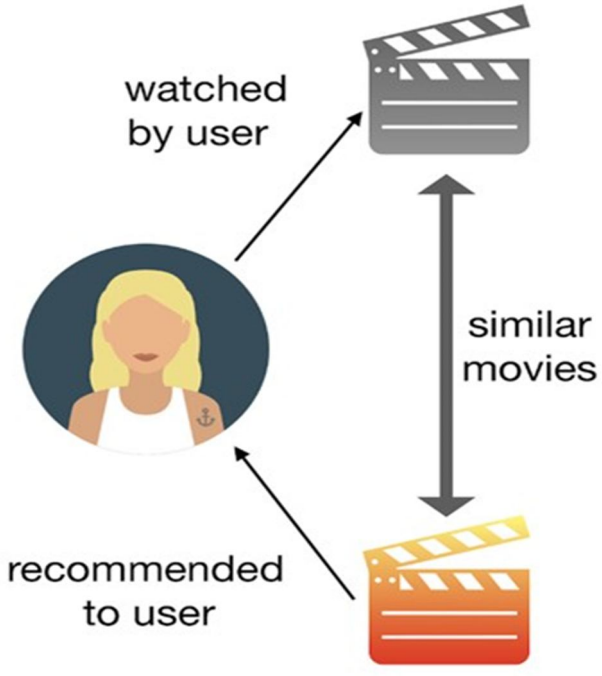


Fig 1: Content Based Recommendation

TABLE I:
OVERVIEW

Content-based recommendation system	
Advantages	*Disadvantages*
The methodology notloes any require ›ther the information about use are the because recommendatiomakes of personalized to individu This scaling a large gro individuals easier.[9]	As the feature representation of the objects is hand-engineered to some extent, this process needs a large lot of domain expertise. As a result, the model can only be as excellent as the hand-engineered attributes.
The model can understand a user's personal tastes and provide specialized recommendations that only a few other users are interested in.	Only depending on the user's current interests can the model make recommendations. To put it another way, the model's ability to capitalize on consumers' pre-existing passions is restricted.
In contrast to communal filtering, new things may be proposed before being reviewed by many people.	Content-based filtering delivers only a tiny degree of originality since it must fit the attributes of a user's profile with available items.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue I Jan 2023- Available at www.ijraset.com

II. BUSINESS PROSPECTS

A. Key Partners

As a business, this model can partner with organizations that focus on providing similar Information on their forums or websites. We can also find partners In Institutes who want to encourage students to develop projects. These can be viewed as prosperous partnerships as they would add value in terms of resources and customers.

B. Distribution Channels

To gain traction on the website, a digital marketing scheme can be put in order such that the site becomes prevalent and relevant to its purpose. Social media campaigns can also be done in collaboration with tech events via organizations and Institutes that favor our problem statement.

C. Unique Value Proposition:

Through this model, we aim to provide students, academicians, developers with the valuable resource material for a project that is powered by machine learning algorithms. The material provided would be recommended Users can further create a kit for such projects and can upload their source code as well. These kits can be private, open-source, or for sale. The model would create an Income stream from applying a nominal fee when one user buys a project kit from another user. Users will always be recommended to start projects based on their profile of the past project, projects they have bought/liked /searched. [10]

D. Key Activities

The project would be a self-sustaining one once it is deployed. Regular checks for content regulation and security and technological updates would be required.

E. Customer Segments

Consumers of this model include students, academicians, developers, and other tech geeks

F. Customer Relations:

This aims to be a passive site that provides the user with information based on their needs and profile. Customers can reach out to a help desk from there they can relay their
Issues and concerns via mail to the technical team.

G. Key Resources

Resources required by the model would be provided by the consumer itself. Initially, some data can be required to build the M.L. model of the project.

H. Cost Structure

In terms of investment, we might need a platform to host the website. We would also need a data management system such that the user data can be stored and retrieved with ease *I. Revenue Streams:*

In terms of revenue, we can consider multiple options such as the one mentioned in the value proposition section. We can top that up by providing sponsored resources on the top of the recommended kit. Ad spaces can also be a great method of generating a steady and passive Income if the model has a decent traction

III. LITERATURE SURVEY

A. Title: Design of aRecommender Systemfor Web-Based Learning.

Publisher & Year: *Lakshmi Sunil, Dinesh K Saini, WCE,*
London UK ISBN: 978-988-19251-0-7 ISSN: 2078-0958
July 2013

Key Findings: The paper[11] discusses the design of a recommender system based on a content ontology and learner profiles created in the system. The paper shows various types of recommender systems such as:

- PRS system
- QSIA system

©IJRASET: All Rights are Reserved

- CYCLADES system
- Learning Resource Recommending Systems.

The paper also discusses the issue of designing a recommender system for learning in the online environment. The CRS is a hybrid recommender system as it is based on learner profiles and content recommendations from user collaboration.

B. Title: Survey on Collaborative Filtering, Content-based Filtering and Hybrid Recommendation System

Publisher & Year: *Poonam B. Thorat, R. M. Goudar, Sunita Barve, Computer Engineering, MIT Academy of Engineering, Pune India, International Journal of Computer Applications (0975 – 8887) Volume 110 – No. 4, January 2015*

Key Findings: Content-based filtering (CBF) tries to recommend items to the active user based on similarity count which is rated by that user positively in the past.[12]

1) Advantages

a) Content-based recommender system provides user independence through exclusive ratings which are used by the active user to build their profile.

b) Content-based recommender systems provide Transparency to their active user by giving them an explanation of how the recommender system works.

c) Content-based recommender systems are adequate to recommend items not yet placed by any user. This will be advantageous for new users.

2) Disadvantages

a) It is a difficult task to generate the attributes for items in certain areas.

b) CBF advocates the same types of items because it suffers from an overspecialization problem.

c) It is harder to acquire feedback from users in CBF because users do not typically rank the items (as in CF) and therefore, it is not possible to determine whether the recommendation is correct. c. Challenges:

- Data sparsity
- Scalability
- Diversity

C. makeprojects.com

"makeprojects.com" is a website that helps its users explore and grow their projects. Here, makers share projects ranging across engineering, science, art, food, design and craft — and anywhere in between! The user can either upload their said project or explore different projects under the domain of Craft & Design, Digital Fabrication, Drones & Vehicles, Education, Science, and technology. It filters the content based on user preferences. It also has a huge community that helps and shares tips and tricks with other users of the community. All of this happens through their content-based recommendation system. It is smart to understand what domain a new user is interested in and slowly make the user a part of that community. [13]

D. Title: Online book recommendation system

Publisher & Year: 2015 Twelve International Conference on Electronics Computer and Computation (ICECCO), 2015, pp. 1-4, doi: 10.1109/ICECCO.2015.7416895.

Key findings: In this research[14], the authors describe a collaborative filtering-based recommendation system. The primary objective was to speed up the suggestion process, i.e., to design a system that can provide users with high-quality recommendations without requiring extensive profile information, browsing history, etc. The outcomes of the experiments demonstrate that the suggested strategy offers pertinent advice. The presented work may be used to suggest products like movies, music, and other media in various sectors. Different types of methodologies are used by recommendation systems to produce pertinent recommendations. Collaboration-based filtering and content-based filtering are common practices. According to the user's preferences as determined by his profile, the content-based filtering strategy learns the content of the item, or product, to classify it to the suitable user. Instead of matching things with users based on content, collaborative filtering matches items with users based on the assumption that users who have previously agreed would continue to do so. On the basis of the evaluations people give the objects, information about their preferences may be gathered. Amazon is one of the companies that successfully used collaborative filtering and used it to effectively propose a wide variety of its items. There is also a hybrid recommender system, which combines the two methods just mentioned.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

E. Title:GroupRecommendationAlgorithms for Requirements Prioritization

Publisher & Year: 2012 Third International Workshop on Recommendation Systems for Software Engineering (RSSE), 2012, pp. 59-62, doi: 10.1109/RSSE.2012.6233412.

Key Findings: The implementation of group decision heuristics in the context of needs prioritizing has been demonstrated and motivated in this paper. Discussions among stakeholders on alternate need prioritizations may be intensified by group recommendations, which would improve the decision's quality. Future research on the applicability and effects of recommendation techniques in other types of requirements engineering situations will be built on the findings of the study reported in this paper.

This paper makes it clear that the approach taken while building a recommendation for a group will be very different than building recommendation systems for an individual. The paper also aims to broaden the applicability of methodologies for group recommendations to needs prioritizing. We demonstrate fundamental group recommendation heuristics that are used in a needs prioritizing situation in the paragraphs that follow.

IV. PROPOSED SYSTEM

A. Content-Based Filtering with Count Vectorization Method

Since, we all are familiar with services like Netflix, Amazon, and YouTube. These services have developed their system in such a way that they make sure their users have the best experience and have complete satisfaction. Hence, here we used Content Based Filtration for our model.

Content-Based recommender system tries to read the behavior and liking or preferences of a user based on the item's features and thus the system makes a note of the positive reactions made by the user.

Once we know the likings of the user we can embed him/her in an embedding space using the feature vector generated and recommend him/her according to his/her choice. During recommendation, the similarity metrics (Cosine Similarity, in this case) are calculated from the item's feature vectors and the user's preferred feature vectors from his/her previous records. 14 Then, the top few are recommended. Also, Content-based filtering does not require other users' data during recommendations to one user.

The reason behind using cosine is that the value of cosine will increase with decreasing value of the angle between which signifies more similarity. The vectors are length normalized after which they become vectors of length 1 and then the cosine calculation is simply the sum product of vectors. This method helps to relate the user's interest and preference with the design of the features provided in the product. The feature which matches the most with the user's interest is the part which gets recommended, so always find the shortest distance between the vectors to ensure the maximum similarity.

In the Content based filtering basically two types of methods are used. Firstly, users are made to fill a form type including all the available features and the user selects the most preferable option for them. Secondly, the system can create a database of the preference or interests of the user and comparatively keep track of the user's connection with the feature. Moreover, users can be asked what features they believe identify with the products the most.

Once a numerical value, whether it is a binary 1 or 0 value or an arbitrary number, has been assigned to product features and user interests, a method to identify similarities between products and user interests needs to be identified. A very basic formula would be the dot product. To calculate the dot product the following formula should be used, $\sum pi\ d\ i{=}1\ ui$ (where pi is the product feature value and ui user interest value in column i). In the table given above, user interest level with Product 1 can be estimated to be 2*1 + 1*1 + 1*2, which equals 5. Similarly, interest in Product 2 will be 1*4 = 4 and will be 2*3 + 1*1=7 in Product 3. Hence, Product 3 will be the algorithm's top

recommendation to the user

V. TOOLS AND SOFTWARE

1) *Kaggle*: This will be useful for the development as well as making of the dataset which will act as the main base of the making of the model. This will also help to build an existing data and will help to analyze which all data columns are necessary for making of the model.
2) *Canva:* Canva is basically the primary tool which helps to initialize the project as it is used for designing the UI of the whole model and the app. Also planning of the admin and user diagram along with making charts for the project are fulfilled from this tool. This is also beneficial to add some of the images we want to add to our website to make it more user friendly.
3) *PyCharm/Jupyter Notebook/VS code:* The main coding where making of the actual model, training and testing of the dataset, building the model, coding the UI of the website that is the front-end is performed on this platform. The programming languages that we will be using are:

4) *Python:* Making and Building of the model as well as for the modification of the dataset.
5) *HTML:* It is useful for making the UI of the website.
6) *CSS:* For the Styling of the website.
7) *JavaScript:* For making the website responsive.
8) *MongoDB:* This platform is used for making the database that is the backend of the website. This helps to make a database which will help to develop the project.
9) *Heroku:* Using Heroku is a one-stop solution to host any website or server. So, this platform will be used for hosting of the website.



Fig 4 : Flow of project

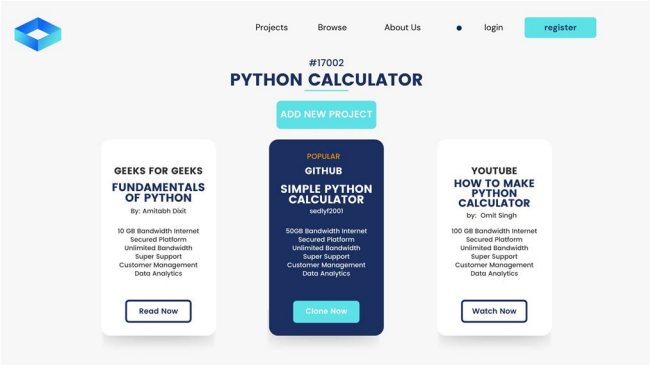


Fig 5 : Project preview UI

VI. WORKING MODEL OF PROPOSED SYSTEM

1) *Sign-in Page:* Helps a user to create an account and then be able to use our services.
2) *Form Menu Page:* This is a type of a menu or a cover page option for the form filing process. This is mainly used to get details from the user so the model can analyze and accordingly give recommendations to the user. This includes Personal details, Professional details, Preference of the user, etc.
3) *Home Page:* Basically, this page is divided into 2 parts giving the option of asking a query on one side and providing recommendations on the other side.
4) *Browser Page:* This page is basically displaying the option given by the recommendation model where the input query is given by the user and accordingly all the available sources are displayed to the user.
5) *Dataset:* Making of the dataset is done in two ways:

Using the platform of Kaggle and second manually making a .csv file in order to create data for the model. Dataset will include the details of the users followed by the category of queries they have and also a column where the dataset will record the questions asked by the users on that particular category. Queries will be noted and categories will help us modify the dataset in such a way that it helps us to develop the model with ease. Also, later when the app is developed and users start to upload their projects and ask for more queries, the model will itself learn and update the dataset. This helps to improve the working of the model as well.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

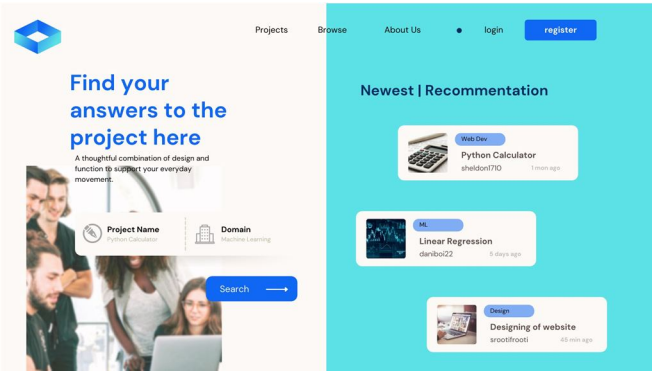


Fig 6 : Home page UI

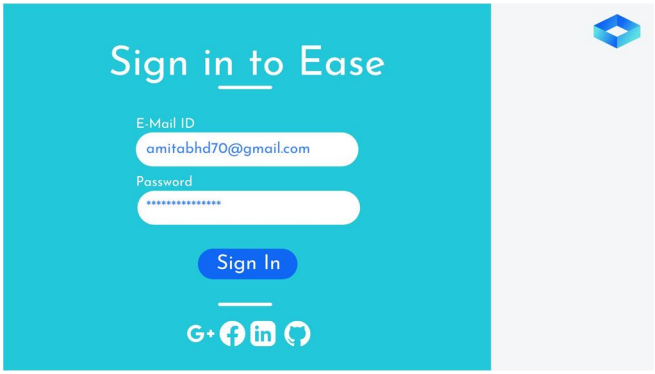


Fig 7 : Login/Sign-up UI

VII. RESULT AND CONCLUSION

A. *Result*

1) In this project Content-Based Filtering with Count Vectorization Method will be done.

2) The UI of this project is aimed to be as simplified as it can be. *B. Conclusion:*

A potent technology for social networks and internet commerce is recommender systems. They can help a company boost sales, assist customers in finding products they like, or assist individuals in making relationships with like-minded people. Using a content based recommendation system to solve user queries is one of the best approaches to make a portal for students and researchers to post & find projects.

B. *Future Scope*

Recommender systems that are based on content have their own set of constraints. Interdependencies and complicated behaviors are difficult to capture using them. For example, someone might prefer articles on Machine Learning if they incorporate both theory and actual application, rather than just theory. These recommenders are unable to capture this type of data. We plan on overcoming these issues and inconveniences in our project. Also, we would be planning to inculcate features such as giving users the option to upload their project and to have an option to keep it private or global.

VIII.ACKNOWLEDGMENT

We gratefully acknowledge the support, guidance, and encouragement of our Dissertation Guide Assistant Professor Mr. Shailesh Sangle for his novel work.

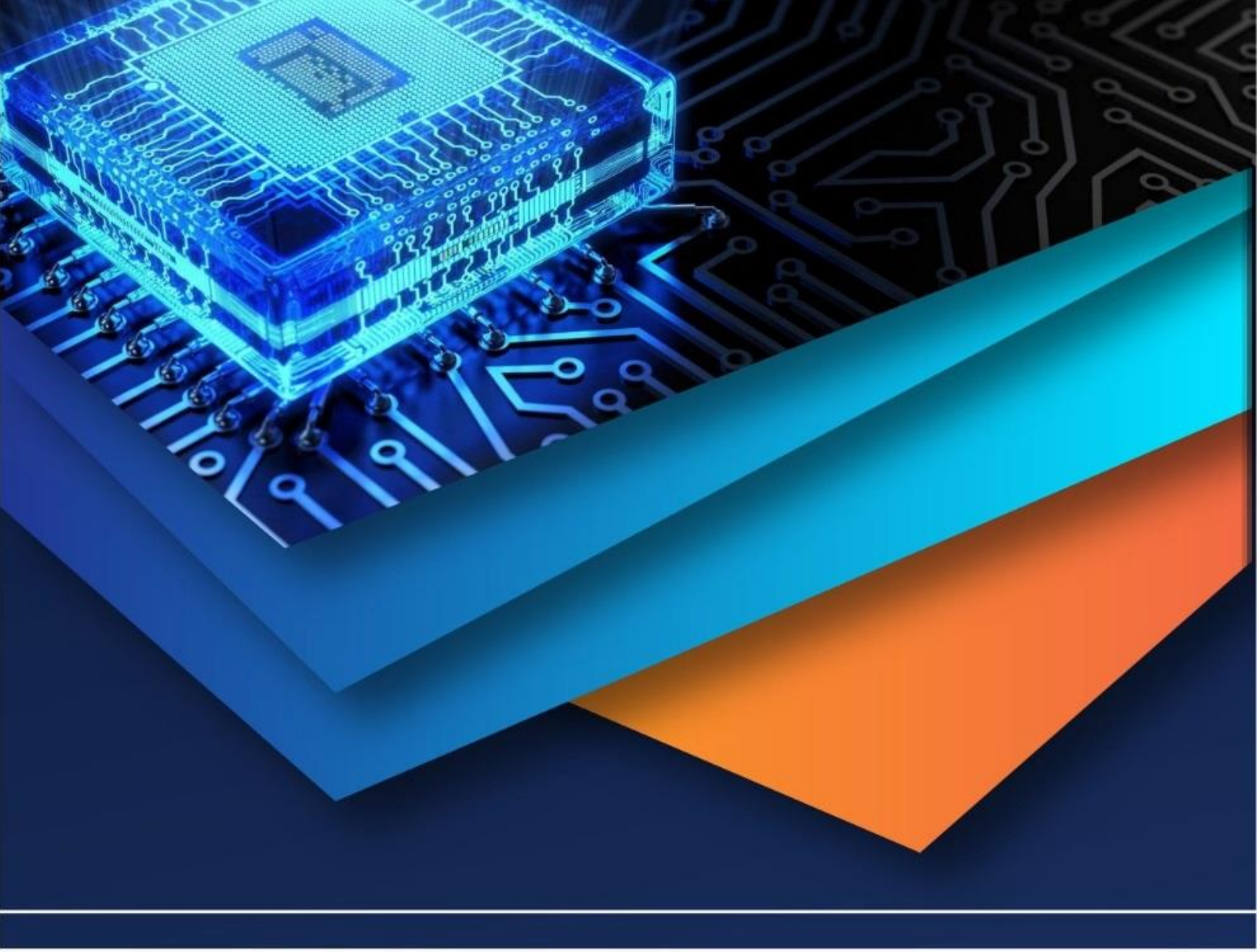
REFERENCES

[1] D. Goldberg, D. Nichols, B. M. Oki and D. Terry, "Using collaborative filtering to weave an information tapestry", Communication of the ACM 35 (1992), 61-70

[2] P. Resnick et a!., "GroupLens: an open architecture for collaborative filtering of netnews", Proc. ACM 1994 Conf Computer Supported Cooperative Work, ACM Press, 1994, 175-186

[3] UShardanand, P. Maes, "Social information filtering: algorithms for automating 'word of mouth'", ACM 1995 Conf Human Factors in Computing Systems, Vol I , 210-217

[4] lA. Konstan et a!. "Group Lens: applying collaborative filtering to Usenet news", Comm. ACM, VoI40,no.3, 77-87,1997
[5] Khera, Ankit, "Online Recommendation System" San Jose State University (2008). Master's Projects. 97.
[6] B.Sarwar, G. Karypis, lKonstan and I Riedl, "Item-Based collaborative filtering recommendation algorithms", Proc. 10th In!' I WWW Conf. 2001
[7] G. Linden, B. Smith and 1 York, "Amazon. com recommendations: item-to-item collaborative filtering", IEEE Internet Computing, Jan.lFeb. 2003
[8] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible Extensions", IEEE Trans. on Knowledge and Data Engineering 17, (2005), 634-749
[9] Isinkaye, Folasade Olubusola, Yetunde O. Folajimi, and Bolande Adefowope Ojokoh. "Recommendation systems: Principles, methods and evaluation." Egyptian informatics journal 16.3 (2015): 261-273.
[10] Thorat, Poonam B., Rajeshwari M. Goudar, and Sunita Barve. "Survey on collaborative filtering, content-based filtering and hybrid recommendation system." International Journal of Computer Applications 110.4 (2015): 31-36.
[11] Denny Abraham Cheriyan, "Personal recommender systems for learners in lifelong learning networks Springfield: UOS Press, 2004, pp.6-9.
[12] Sunil, Lakshmi, and Dinesh K. Saini. "Design of a recommender system for web based learning." Lecture Notes in Engineering and Computer Science 1 (2013): 363-368. [12] Sharma, Lalita, and Anju Gera. "A survey of recommendation systems: Research challenges." International Journal of Engineering Trends and Technology (IJETT) 4.5 (2013): 1989-1992.
[13] makeprojects.com
[14] N. Kurmashov, K. Latuta and A. Nussipbekov, "Online book recommendation system," 2015 Twelve International Conference on Electronics Computer and Computation(ICECCO), 2015, pp. 1-4, doi: 10.1109/ICECCO.2015.7416895.
[15] A. Felfernig and G. Ninaus, "Group recommendation algorithms for requirements prioritization," 2012 Third International Workshop on Recommendation Systems for Software Engineering (RSSE), 2012, pp. 59-62, doi:10.1109/RSSE.2012.6233412.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)