iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 Issue: VII Month of publication: July 2021

DOI: https://doi.org/10.22214/ijraset.2021.36424

Online Recommender System

Prof. Ranjanroop Walia[1], Sahil Singh[2], Rishabh[3], Rohit Verma[4]
[1]*Professor (Mentor),* [2, 3, 4,] *Student, Inderprastha Engineering College, Ghaziabad, India*

Abstract: As the size of the e-commerce market grows, the consequences of it are appearing throughout society.The business Environment of a company changes from a product center to a user center and introduces a recommendation system. However, the existing research has shown a limitation in deriving customized recommendation information to reflect the detailed information that users consider when purchasing a product. Therefore, the proposed system reflects the users subjective purchasing criteria in the recommendation algorithm. And conduct sentiment analysis of product review data. Finally, the final sentiment score is weighted according to the purchase criteria priority, recommends the results to the user. Recommender system (RS) has emerged as a major research interest that Aims to help users to find items online by providing suggestions that Closely match their interest. This paper provides a comprehensive study on the RS covering the different recommendation approaches, associated issues, and techniques used for information retrieval.
Keywords: recommender system; issues; challenges; literature review; Filtering approach; filtering technique; information retrieval technique; machine learning; research trends; future direction.

I. INTRODUCTION

Almost every one of us has experienced that those friendly suggestions, even with their best intention, are not really effective in many of the cases as others' taste does not necessarily mean to harmonise with that of ours. These suggestions may often also be biased. The other anomalous options we can endure, such as be a decision science expert and try out the complex theories, plunge into the internet and waste hours going through the confusing reviews and suggestions, hire an expert, go along the herd, or simply listen to our soul. The point is that it is very arduous to highlight a precise suggestion on the items on which we might be interested. It would be of great help if we would have a personal advisor who would assist us by suggesting the best option whenever we have to make a decision.

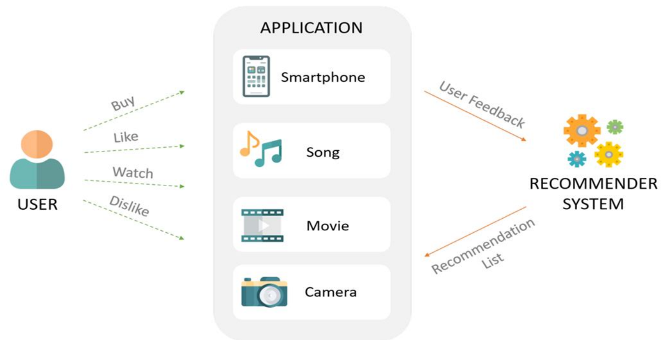
Thankfully, we have one such in the form of a web application known as the recommender system (RS). An RS is an intelligent computer-based technique that predicts on the basis of users' adoption and usage and helps them to pick items from a vast pool of online stuffs. Most internet users surely have happened upon an RS in some way. For instance, Facebook recommends us, prospective friends, YouTube recommends us the videos in accord, Glassdoor recommends us matching jobs, TripAdvisor recommends us suitable holiday destinations, Goodreads recommends us interesting books and so on. RSs have garnered phenomenal acceptance in the e-business scenario.

E-Commerce portals (e.g., eBay, Amazon, etc.) are using RSs to entice customers by heaving with the products that customers should, presumably, going to like.

This has helped them to attain a huge boost in sales. Not only the online business, but there are other applications also that take advantage of RSs, such as social networks, online news portals, entertainment sites, and other knowledge management applications. Actually, RSs have begotten a new dimension in the communication approach between users and online service providers. These days, many companies are adopting RS techniques as an added value to enrich their client services. Though, the implementation of an RS depends on the particular recommendation approach adopted by the application, the core working of RSs remain more or less the same for all applications.

The focal objective of RSs is to aid users in their decision making in order to pick out an online item, by supporting with in-hand recommendations of high accuracy . The potential of RS in different domains has attracted researchers to explore the possibilities exhaustively. A lot of work has been done by the research community to enhance the applicability and performance of RSs over the last few years . New methodologies and algorithms were developed to address many of the technological challenges such as producing more accurate recommendation while reducing online computation time. Several recommendation algorithms have been proposed and successfully implemented in different domains. These algorithms mainly follow demographic filtering (DF), content-based filtering (CBF), collaborative filtering (CF) and hybrid approaches.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VII July 2021- Available at www.ijraset.com



II. LITERATURE REVIEW

We use the term "idea" to refer to a hypothesis about how recommendations could be effectively generated. To differentiate how specific the idea is, we distinguish between recommendation classes, approaches, algorithms, and implementations We define a "recommendation class" as the least specific idea, namely a broad concept that broadly describes how recommendations might be given. For instance, the recommendation classes *collaborative filtering* (CF) and *content-based filtering* (CBF) fundamentally differ in their underlying ideas: the underlying idea of CBF is that users are interested in items that are similar to items the users previously liked. In contrast, the idea of CF is that users like items that the users' peers liked. However, these ideas are rather vague and leave room for different approaches. A "recommendation approach" is a model of how to bring a recommendation class into practice. For instance, the idea behind CF can be realized with user-based CF , content boosted CF , and various other approaches . These approaches are quite different, but are each consistent with the central idea of CF. Nevertheless, these approaches to represent a concept are still vague and leave room for speculation on how recommendations are calculated. A "recommendation algorithm" precisely specifies a recommendation approach. For instance, an algorithm of a CBF approach would specify whether terms were extracted from the title of a document or from the body of the text, and how terms are processed (e.g., stop-word removal or stemming) and weighted (e.g., TF-IDF). Algorithms are not necessarily complete. For instance, pseudo-code might contain only the most important information and ignore basics, such as weighting schemes. This means that for a particular recommendation approach there might be several algorithms. Finally, the "implementation" is the actual source code of an algorithm that can be compiled and applied in a recommender system. It fully details how recommendations are generated and leaves no room for speculation.

III. HISTORY AND BACKGROUND OF RSs

Though Graundy (Rich, 1979), a computerised librarian, may be considered as an early step towards automatic RS , the idea of accruing opinion of millions of online users in order to find more suitable and appealing contents have emerged in the early '90s. Tapestry , a manual CF system, allowed users to query for items in an online information domain. GroupLens has used a similar technique to identify the particular user's interest by using Usenet articles and based on the user's action to provide a personalised recommendation. In the late '90s, the RSs started to capture the attention of the researchers from the domain of human-computer interactions, machine learning and information retrieval, and other allied disciplines. As a result, many RSs for music, the bell core video recommender for movies, and Jester for jokes for different application domains have been developed. During the same period, the RS had been increasingly utilised in marketing to enhance sales, and customer experiences and many commercial applications of RSs were surfaced in the online realm . Gradually, recommendation approaches moved beyond the CF and many of the RS researchers' focus of interest shifted towards the content-based recommendation (CBR) approaches based on information retrieval, Bayesian inference, and case-based reasoning methods. In 2006, hybrid RSs , attracted much attention and Netflix launched the Netflix prize to improve the aptness of movie recommendations.

Nowadays social networking sites (such as Facebook, Twitter, etc.) have emerged as a substantial platform for applying RSs. These popular sites are considered to be the major source of information about people and hence becoming a great option to leverage novel and innovative approaches for the recommendation, leaving behind the old methods, to increase the accuracy .

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VII July 2021- Available at www.ijraset.com

IV. DIFFERENT RECOMMENDATION APPROACHES

Several recommendation approaches have been proposed and adopted in different applications. In this section, we present a brief overview of the popular recommendation/filtering approaches in RSs. Table 1 summarises these approaches along with the source of input data, the extraction method used, the limitations of each approach, and the research works emphasising the particular approach.

A. Different Recommendation Approaches

Several recommendation approaches have been proposed and adopted in different applications. In this section, we present a brief overview of the popular recommendation/filtering approaches in RSs. Table 1 summarises these approaches along with the source of input data, the extraction method used, the limitations of each approach, and the research works emphasising the particular approach while Table 2 lists the popular application domains of RS with their filtering techniques and related research works.

B. Content-based Recommender System (CBRS)

CBRS uses CBF to recommend items by matching user profile and item description. The user profile may include his previous search or purchase history. The system learns to recommend items that are similar to the ones that the user liked in the past. The

C. Collaborative filtering recommender system (CFRS)

This is the most recognised and widely implemented RS. CFRS follows the philosophy of "a man is known by his company he keeps." That means if CFRS believes that if two or more user's interests matched in the past, then it is likely that in future also their interests should match. For example, if the purchase histories of user1 and user2 strongly overlap then it is high on the cards that if user1 buys a product, then user2 will also buy the same or similar product. CF approaches to keep track of the user's past reviews and ratings on items to recommend similar items in the future. Even if the user did not deal with a particular item, it would be recommended to him if his peers have used the same. It is obvious that to achieve reasonable recommendation accuracy a large number of user groups are required to be considered. Trust is an important factor for reliable recommendation. has considered a trust-based CF approach to present a temporal-trust-based method to measure trust value.

1) *Memory-based collaborative recommender system (CRS):* similarity measure and the prediction computation are the two main steps used in the memory-based CRS, which is further categorised into two parts based on their similarity computation method as follows :
   a) *Item-based CRS:* similarity computation is performed on a set of items.
   b) *User-based CRS:* similarity computation is performed based on the similarity values of users.
2) *Model-based CRS:* in model-based CRS, different machine learning algorithms such as Bayesian network, clustering, Markov decision process, sparse factor analysis, dimensionality reduction techniques, and rule-based approaches, etc., are used to build a model for the recommendation.

D. Demographic Recommendation System (DRS)

DRS works based on the users' demographic profile such as age, sex, education, occupation, locality, etc. It generally uses clustering techniques to categorise target users according to demographic information. But in this RS if the demographic attributes remain unchanged, the user will receive the recommendation for the same set of items.

Thus, they might miss some new and worthwhile recommendation. Demographic information about a user can improve the accuracy of RS.

E. Hybrid Recommender System (HRS)

As the name suggests, hybrid RS is the product of the combination of multiple filtering approaches. The most popular pairing HRS is that of CBS and CFRS. The purpose of combining different filtering approaches is to improve the accuracy of recommendations while eliminating the limitations of the individual filtering approaches.

F. Knowledge-based recommender system (KBRS)

To recommend the items such as flat, bike, TV, etc., which are less frequently purchased by a user, sufficient information on the basis of which recommendation is made may not be available or relevant (even if available). For that, some additional information (e.g., the user's social network activity) is required.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VII July 2021- Available at www.ijraset.com

Knowledge-based RSs provide a recommendation based on additional knowledge model related to the relationship between the present user and items. Case-based reasoning technique is a common feature of KBRSs that divides the user's need into multiple cases, depending on various criteria and provide recommendations that closely matches to
user's likely preference. Another type of KBRS, known as constraint-based RS that works as per the user's preference and recommends items that match the preference. If no such item is available, then a set of alternative items that are close to the preferred item is recommended. Semantic web technology can help to establish a diversified knowledge base of the users and the items. It utilises ontologies, a formal knowledge representation the method that is used to express the domain knowledge of users and items. The similarity between items can be calculated based on domain ontology. Metadata of a user profile and item description are used to establish a proper matching for the recommendation.

G. Context-aware recommender system (CARS)

If the target user's contextual information is available, we can make the RSs ubiquitous. Various attributes like time, location, companion, mood, etc., can define a context. The difference between contextual information and demographic information is that demographic properties of a user generally remain the same for a longer period, whereas contextual information changes when the surroundings of the user change. Hence the mobile applications play a significant role in CARS. CARS plays an important role, especially in personalised and direct online marketing. The more relatable and specific recommendation can be provided by capturing the emotional context of the user.

TABLE 1

Recommendation	Source of data	Extraction	Limitations
Content-based (CBRS	Content related to product and user's information etc.	Document modelling, information filtering, and similarity metrics, etc.	Problems in content analysis, non-homogeneous items and no methods use in prediction of missing ratings, etc.
Collaborative (CFRS)	Feedback on products from the user's side.	K-nearest neighbor , cosine and correlation-based similarity, etc.	Cold start problem, sparsity problem, synonym problem and false user's rating problem, etc.
Demographic (DRS)	User's demographic information such as gender, age, date of birth, etc	Clustering and locating group interest etc.	It totally depends on demographic information, which can be less accurate and static in nature etc.
Knowledge-based (KBRS)	User's information from social networks, data used for search, or query of the product, etc.	Machine learning and Decision rule etc.	More cost is required for knowledge acquisition; preferences are not always independent from each other, etc.
Context-aware (CARS)	Data related to the different context of the user and product.	Text analysis, machine learning methods, classification, clustering, association rule, and neural network, etc.	Difficult to collect the desirable context for a recommendation because the user's context is dynamic in nature.
Hybrid (HRS)	Any source that is mentioned in this table.	Any extraction method that is used in this table.	Diversity, novelty and serendipity of recommendation

V. INFORMATION RETRIEVAL TECHNIQUES IN RSS

Numerous sources of information have crammed the digital world with unbounded data. The scenario has been exaggerated by the interactive participation of people. To deliver an effective and fruitful recommendation, the RS needs to study all possible zones of dealings to extract and analyse informative data to understand people's preferences and tastes.

A. Machine Learning

Machine learning provides an entity (machine) the ability to learn, artificially, without programming explicitly. It applies different algorithms like logistic regression, decision tree, association rule learning, cluster, Bayesian networks and support vector machine, etc.

B. Logistic Regression

Logistic regression is used for the prediction of discrete variables by using continuous and discrete data (Wang, 2011). To consider a collaborative tag RS, have utilised this technique to rank the meaningful tags in social networks. Logistic regression is also used in determining the trustworthiness of a user by identifying the probable attacks in CFRS.

C. Decision Tree

The decision tree is a powerful technique that helps in choosing an option among multiple alternatives. In RS, it is used to calculate and predict the missing preferences of
users. This technique in tackling the cold-start problem to provide a high-quality service recommendation for new items.

D. Association Rule Learning

Association rule learning is used to extract the frequent patterns, associations, correlations or causal structures from users and items dataset for recommendations.

E. Cluster Analysis

In RS, to make a group, among a large set of objects, based on similarity, structures, and patterns, cluster analysis (i.e., unsupervised learning technique) is used.

F. Bayesian Network

A Bayesian network classifier (i.e., a probabilistic model) is applied to solve classification problems in huge networks like social networks. To solve the user's cold start problem and improve accuracy in the recommendation,some proposed a trust-based probabilistic recommendation model for social networks .

G. SVM

Support vector machine (i.e., supervised learning) is used with an associated learning algorithm for analysing data using classification (linear and nonlinear) and regression analysis. Some have used this technique along with Hilbert-Huang transform to detect profile injection attacks in CFRS.

H. LDA

Extracting a common topic from various documents is called topic modelling. A topic is identified with the help of a different combination of words in a document. LDA (a probabilistic model of a corpus) used for topic modelling in RSs.

I. TF-IDF

TF-IDF (2017) is used to extract the important words from documents for identifying the topic. TF-IDF is used for document retrieval to recommend tourist spots. TF-IDF is used in to find similar user and tweets in hashtag recommendation for Twitter users.

J. Deep learning

Deep learning plays a major role in extracting hidden patterns from data and has opened up a new area in data mining research . It can be used in the building of effective and dynamic behaviour modelling in RSs. We can gather intrinsic details about the user by understanding the approaches of supervised and unsupervised learning in the deep neural network

VI. PROBLEMS ASSOCIATED WITH RSS

Filtering approach that suffers

Problems in	Content-Based	Collaborative	Demographic	Hybrid	Knowledge-Based	Context-aware	Research work
Contextual data requirement						✓	Feng and Tran (2019) and Singh et al. (2019a)
Extra knowledge modelling requirement					✓		Berghofer et al. (2003) and Zhang et al. (2013)
Long tail		✓		✓			Park (2013), Hamedani and
Black box		✓		✓			Kaedi (2019) and Millecamp et al. (2019)
Abbreviation	✓	✓	✓	✓			Shen et al. (2001) and Chatti et al. (2013)
Synonymy	✓	✓	✓	✓			
Scalability		✓		✓	✓	✓	Varudkar et al. (2019) and Koshti et al. (2019)
Sparsity		✓					Ahmadian et al. (2019)
Cold start		✓					Revathy and Pillai (2019) and Zhu et al. (2019)
Over specialisation	✓						Adamopoulos and Tuzhilin (2014) and Cazella and Alvares (2005)
Limited content analysis	✓						Zhang et al. (2012b), Chidlovskii et al. (2001) and Pereira (2019)

VII. FUTURE DIRECTIONS OF RSS

To date almost all of the RSs have been designed for sellers, producers, and service providers, i.e., they are designed to attract potential customers. We believe that future RSs will not only be limited to business, but they will have a much greater impact on our daily life (Bourke, 2015). These systems will become truly ubiquitous and become an

essential tool in every sphere of our life. The future RSs will not be bound merely to the applications for buying and selling products; rather it will become a sort of personal advisor which will assist in every sector of living by giving important suggestions and guidance. The ideal RS should be like someone who knows us better than we know ourselves. They should sense our need and will suggest instinctively, even if we do not express explicitly. Few other fields like the internet of things (IoT) internet of everything (IoE), cognitive computing , affective computing (cognitive science and psychology), etc., will play a significant role in future RSs. The RSs will be more intuitive and will continuously improve the quality of the recommendation by taking up user feedback loops from various sources. They will also be more flexible by supporting multi-criteria ratings. Future RSs will come up with innovative recommendation models using reinforcement learning or extensions of

recurrent neural networks (RNN) that will enable them to be accurately context, time, and mood-aware.They will be designed not only to recommend something but to understand when what to recommend and what not to recommend. Below some of the properties of future RSs and the application areas that will leverage the RSs are discussed.

A. Data-driven

The RSs will primarily be driven by IoT , IoE and big data. The major differentiating point of future RSs will be the intelligent use of ubiquitous data. Data will be captured, assessed and analysed literally from anywhere and for anything. Though tackling the ever-increasing data will be a great challenge for the future RS designers because the current algorithms may not be straightforwardly scalable to cope up the unforeseen amount of data.

B. No Cold Start Problem

Future RSs will be able to get rid of the ‘cold start problem’ by collecting suitable and implicit information from other online sources. Social networks, IoE and every possible way of pervasive connectivity will be the main enabler for this.

C. More Customer-Centric

Existing RSs are typically seller-centric, i.e., users get recommendations of only those products which the sellers intend to sell . This restricts the buyers’ independent preferences. Future RSs should serve buyers better by being more buyer-centric. Sophisticated data analytics tools will empower retailers by enabling them to analyse and find a valuable pattern in people’s online purchasing habits. They can capture the obvious tendency of the buyers for the products they are initially interested in and eventually what they purchase. They will also use the information of which products buyers put in their cart and among them which are eventually bought and which are not. Recommending based on these observations will offer buyers an optimised shopping experience. By means of IoT, the manufacturer or the service providers can obtain the usage metrics of the products or services for each user and attune their products or services and pricing strategies accordingly. A left-hander should get product recommendations that are suitable for him.

D. Sensing the emotional state of a user:

With the help of affective computing, RSs will be able to recognise the emotional state of a user and recommend services accordingly. For example, sensing the mood of the user, the RS will recommend appropriate music, movies or books. Going further, if the RS can sense the partner’s mood also, it can recommend both of them an ideal location for spending the evening together or a restaurant to dine.

E. Personalised Healthcare Recommendation

Thanks to the IoT and internet of nano things (IoNT) based ubiquitous and pervasive healthcare, RSs will play a major role in providing better and personalised healthcare. Suitable medicines, health supplements, required lifestyle changes, etc., will be recommended timely. For instance, if the sugar level goes high, then an insulin dose should be recommended. If the

user’s, who is suffering from depression, psychological health is read through affective computing, proper anti-depression medicine can be recommended. If the RS finds from other sources that the person has a back problem, then it might recommend a suitable ergonomic chair that will help in curing the back problem. A diabetic patient should get a recommendation of food products that are sugar-free.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VII July 2021- Available at www.ijraset.com

F. More Ethical

We expect, the businesses will apply RSs more ethically to recommend products to the users. Instead of recommending every possible item that will help them increase their sales, the products will be recommended to the users only if they really require and can afford to buy those. This will relieve users from likely distress.

G. Agricultural Recommendation

RS has a huge role to play in the agriculture sector, but it is not yet explored effectively. Analysing the soil type and other necessities for farming, and following the market and weather predictions, the suitable corps that should be harvested (season wise) will be recommended to the farmers which can help farmer greatly.

H. Education and Career

By studying and analysing interest, social activity, subject score and other parameters RSs should recommend the fitting course to the students. Similarly, a future job RS shall not only consider the biodata but will study other parameters such as both intelligence quotient and emotional quotient, geographic location (according to health) and recommend the job/sector where the candidate will have the maximum
chance of success.

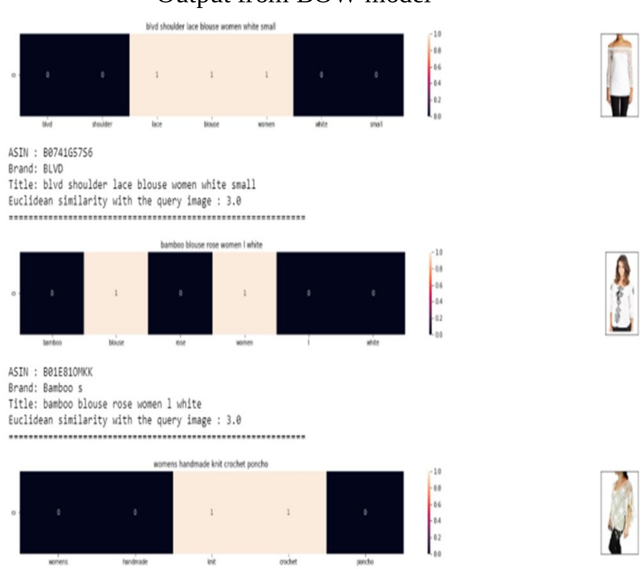
I. Cross-domain Recommendation

Future RSs will experience a major advancement in terms of cross-domain recommendations. Present RSs are somewhat good at assuming user's preference from a single domain, but they are not able to reflect the preference of one domain into some other domain (related or unrelated).
For example, the present RS cannot answer – if a person likes a thriller movie, then what kind of music or book he will prefer. The future RSs will have a unified model of preference for any user that will be effective across different domains.

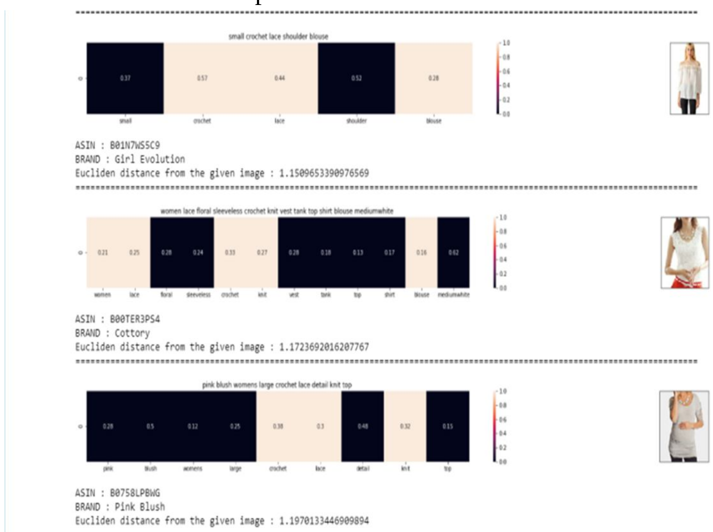
VIII. CONCLUSION

Recommender systems open new opportunities of retrieving personalized information on the Internet. It also helps to alleviate the problem of information overload which is a very common phenomenon with information retrieval systems and enables users to have access to products and services which are not readily available to users on the system.
We have collected dataset from various sources, then we preprocessed the data. After pre-processing we applied Machine Learning algorithms. In addition, we applied the algorithm on BOW model to get the desired recommendation, furthermore we did the same with TF-IDF model.

Output from BOW model



Output from TF-IDF model

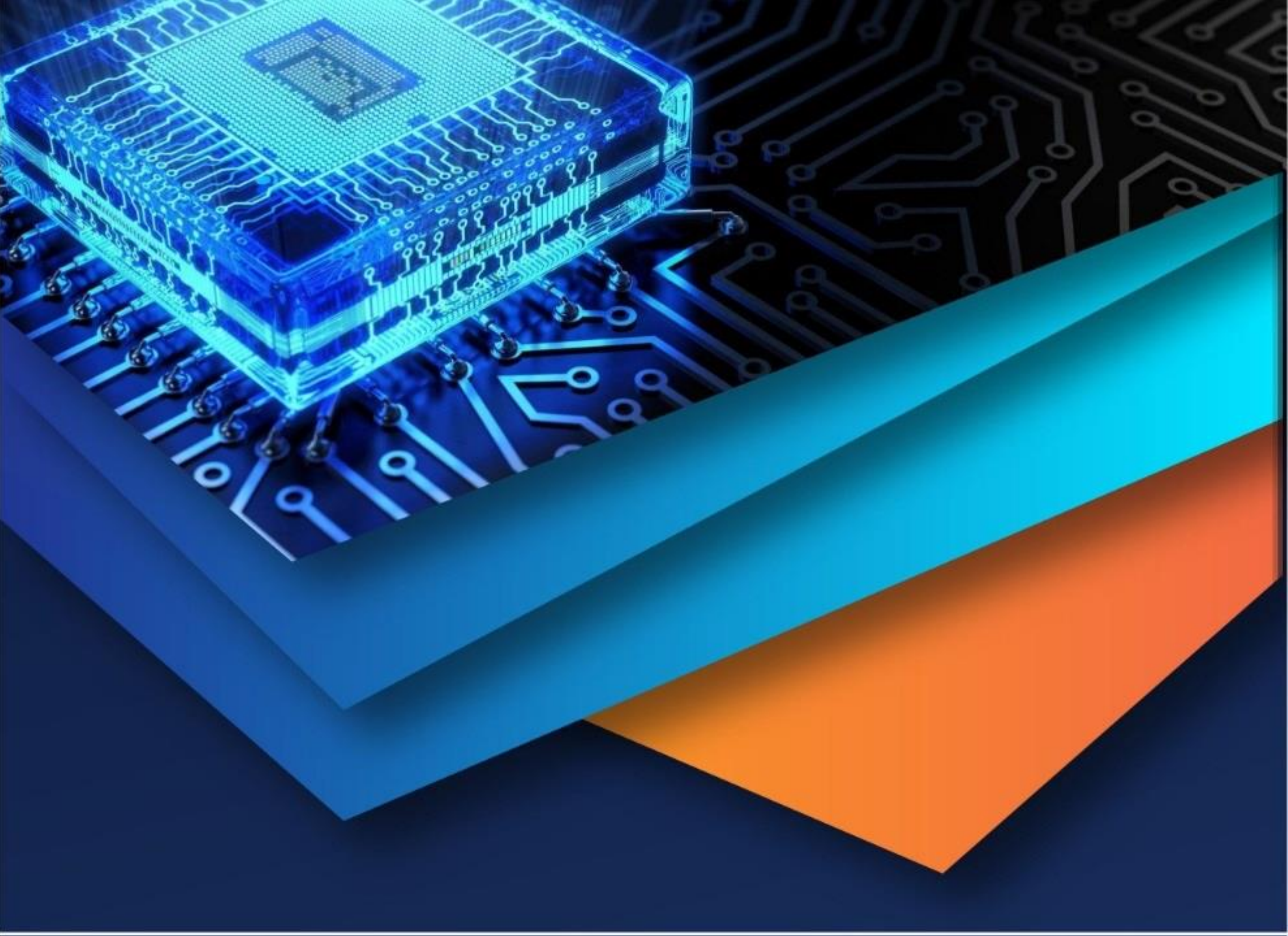


IX. ACKNOWLEDGEMENT

We would like to thank our project guide Prof. Ranjanroop walia for her enormous cooperation and guidance which helped us to develop a very good project idea. We have no words to express our gratitude towards the person who constantly and whole-heartedly supported us throughout the project. She has always given us her precious time and knowledge. The technical guidance provided by her was more than useful and made this project possible. She has always been an inspiration for us. We are also thankful to our HOD Dr Vijay Singh and all the members of the computer department for providing various facilities and support.

REFERENCES

[1] D. Jannach, M. Zanker, A. Felfernig, G. Friedrich Recommender systems – an introduction Cambridge University Press (2010)
[2] G. Linden, B. Smith, J. York Amazon.com recommendation: item-to-item collaborative filtering IEEE Internet Comput, 7 (1) (2003), pp. 76-80
[3] J.A. Konstan, J. Riedl Recommender systems: from algorithms to user experience User Model User-Adapt Interact, 22 (2012), pp. 101-123
[4] B. Pathak, R. Garfinkel, R. Gopal, R. Venkatesan, F. Yin Empirical analysis of the impact of recommender systems on sales J Manage Inform Syst, 27 (2) (2010), pp. 159-188
[5] P. Resnick, H.R. Varian Recommender system's Commun ACM, 40 (3) (1997), pp. 56-58
[6] M.J. Pazzani A framework for collaborative, content-based and demographic filtering Artific Intell Rev, 13 (1999), pp. 393-408
[7] R.J. Mooney, L. Roy Content-based book recommending using learning for text categorization Proceedings of the fifth ACM conference on digital libraries, ACM (2000), pp. 195-204



10.22214/IJRASET

INDEX COPERNICUS
45.98

IMPACT FACTOR:
7.129

TOGETHER WE REACH THE GOAL
IMPACT FACTOR:
7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY