iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 5 **Issue:** XI **Month of publication:** November 2017

DOI:

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:6.887
Volume 5 Issue XI November 2017- Available at www.ijraset.com

Research Aspects of the Recommender System

Manish Kumar Singh[1], Dr. Dinesh Prasad Sahu[2]
[1, 2] *Department of Computer Science, University of Delhi,* [*] *Department of Computer Science, JNU*

***Abstract:** In today's age of information overload, companies incorporates a number of strategies to help people to make smart choices in various areas on the Internet including what to buy, how to buy, how to do some tasks, how to pass their leisure time, and even whom to date. Recommender systems are developed by companies for these reasons that can provide people personal, affordable and high-quality recommendations. For example, Google makes use of recommender system to maximize its target ads revenue. Many e-commerce websites make use of recommender system to advocate people such as "one who bought this can buy this also". Facebook has developed a recommender system to suggest people to tag friends in pictures. Various recommender system algorithms have been proposed so far to provide an affordable, personal and efficient recommendation. We have tried through this paper to put a light on the research aspects of the recommender system and then do analysis of a possible way to design a recommender system algorithm.*

***Keywords:** Recommender System, Content Filtering, Collaborative Filtering, Content-based Recommender System, Service-based Recommender System*

I. INTRODUCTION

In this era of information age where the amount of information on the World Wide Web (WWW) is on a significant rise along with the number of users of the Internet, the companies find it increasingly important to search, map and help the users to get the relevant amount of information as per their preferences, interest, and tastes. Companies use various recommender systems of their respective fields to filter, prioritize and deliver significant information efficiently so that the problem of information overload can be alleviated. The recommender systems are capable enough to do an efficient search through the large volume of information (that are dynamically generated) so as to provide personalized services and content to users. A recommender system is basically understood to be an information system that is meant to do information filtering and is generally used to do the prediction of the "preference" or the "rating" given by a user to a product. It has become quite popular in the recent times and is used generally in myriad areas including music, movies, books, news, social tags, search queries, research articles, and products.

Understanding the significance of recommender system in today's era of overgrowing data and their filtering to make a good decision in the field of social, commercial and education, we have presented this paper to illustrate the research aspects of the recommender system and then made the analysis of a possible way to design a recommender system algorithm. The section-II of this paper describes the recommender system types, section-III discusses the objectives of designing a recommender system algorithm, section-IV reveals the underlying principal behind designing a recommender system algorithm, section-V analyses the possible way to design a recommender system algorithm, section-VI unearths the challenges before designing a recommender system algorithm and section-VII warps up the paper with a certain set of conclusions.

II. RECOMMENDER SYSTEM TYPES

Recommender systems are used by companies to handle the problem of information overload that is usually encountered by the Internet users and provide them with personalized, exclusive service and content recommendations. In the recent past, various recommended systems have been developed to work either as content recommender system or service recommender system. For instance, Ringo is an online social information filtering system that uses the idea of the collaborative filtering to build users profile on the basis of their ratings on music albums. Amazon incorporates topic diversification algorithms to improvise its recommendation. The systems that use the idea of content-based techniques generally base their predictions on user's information, and they ignore contributions from other users as with the case of collaborative techniques. For instance, Fab is a famous content-based recommender system that relies mainly on the ratings of different users so as to create a training set. Letzia is another example that uses content-based filtering to help users find information on the Internet. Let us look at the various service based and open source recommender system build till today so as to understand what research has been made so far in the area of recommender system .

©IJRASET (UGC Approved Journal): All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor:6.887
Volume 5 Issue XI November 2017- Available at www.ijraset.com

A. Service based Recommender System

Some of the widely used service based recommender system are:

1) *Suggest grid*: A generic recommendation system based on Apache Spark.
2) *Mortar Recommendation Engine:* A kind of do-it-yourself recommender system.
3) *Peerius*: An e-commerce recommender system focused on live and email recommendations.
4) *Using Hadoop on Google Cloud*: An example use of Google cloud with benchmarks from recommender system
5) *IDots*: A tool to relate published content
6) *Amazon Machine Learning*: A machine learning platform to model data and create predictions
7) *Dressipi Style Advisor*: It is a clothing specific recommendations service.
8) *Sajari*: It is a search, recommendation, and matching (e.g. dating website) service.
9) *IBM Watson*: It uses cognitive computing to solve complex problems.
10) *Dlib: A recommender system as-a-service for academic organizations such as digital libraries and reference managers.*

B. Open Source Recommender System

Some of the widely used open source recommender system are:

1) Prediction IO: *A machine learning server that can be used to create a recommender system*
2) Mahout Hadoop: A linear algebra based data mining
3) Seldon: It is a Java based prediction engine built on technologies like Apache Spark. It provides a demo movie recommendations application here.
4) Lenskit: It is a Java based research recommender system designed for small-to-medium scale.
5) Oryx V2: It is a large scale architecture for machine learning and prediction (suggested by Lorand)
6) Crab: It is a python recommender based on the popular packages NumPy, SciPy, matplotlib.
7) Predictoris: It is a ruby recommender gem.
8) Surprise: It is a Python scikit for building, and analysing (collaborative-filtering) recommender systems.
9) Light FM: It is an actively-developed Python implementation of a number of collaborative- and content-based learning-to-rank recommender algorithms.

III. OBJECTIVES OF DESIGNING A RECOMMENDER SYSTEM ALGORITHM

A recommender system algorithm is build to benefit both user and service provider in the various aspects. Some of the aspects in which a recommender system works to benefit both user and service provider are discussed in this section.

A. At the user level

A recommender system algorithm is supposed to help users in finding things of their interest by narrowing down the set of choices. Further, it is supposed to benefit the user to explore the space of options available to them. Most importantly, the users are meant to discover new things.

B. At the service-provider level

A recommender system algorithm is supposed to benefit the service-providers in felicitating their customers to get additional and probably unique personalized service so much so that it would help them in increasing trust and loyalty of their customers. Further, it is supposed to help the service-providers in doing of sales, click through rates, conversion, etc. Through recommender system, the service providers are meant to grab opportunities for promotion and persuasion of their services among their customers. Most importantly, the service-providers are meant to obtain more knowledge about their customers.

	Item 1	Item 2	Item 3	...	Item n
User 1	2	3	?	...	5
User 2	?	4	3	...	?
User 3	3	2	?	...	3
...	...	...	...	...	...
User m	1	?	5	...	4

Fig. 1 A sample of users' information obtained through recommender system.

©IJRASET (UGC Approved Journal): All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor:6.887
Volume 5 Issue XI November 2017- Available at www.ijraset.com

IV. UNDERLYING PRINCIPLE BEHIND DESIGNING A RECOMMENDER SYSTEM ALGORITHM

The recommender system algorithm is supposed to provide a wide set of recommendations in a particular field. The system is supposed to implement one of the following methods to do the effective, efficient and personalized recommendations:

A. Collaborative Filtering

The recommender system can use this approach to build a model after following the past behaviour of a user including the previous search or selection of the related item and/or giving of numerical ratings to those items by the user. The model so obtained is then used to do the prediction of the items (or rating for those items) by the user in which they may have interest.

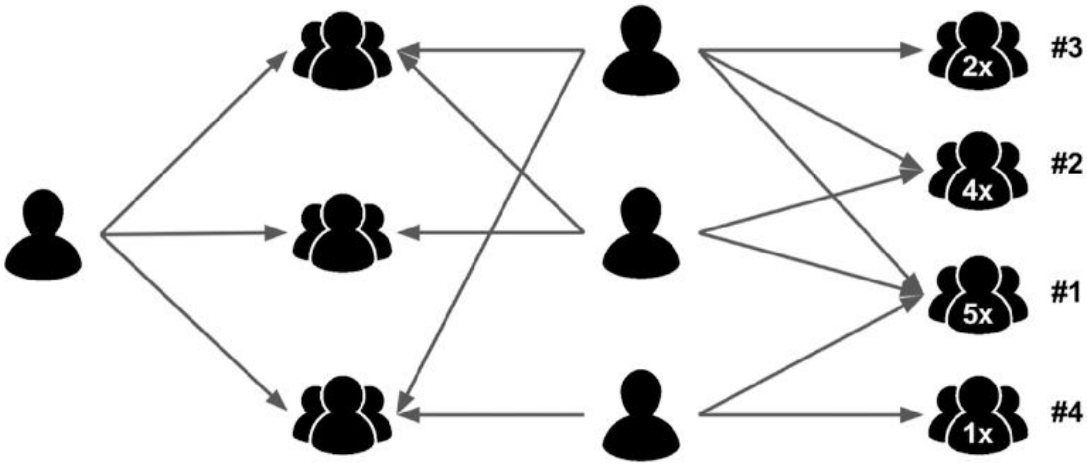


Fig-2: A user-based collaborative filtering algorithm to make an effective recommendation.

The collaborative filtering based recommender system uses the idea that if one has liked an item then one will also like a similar item. Based on this principle, the system provides the recommendations of the items and it is supposed to works quite well to determine the properties/context of each item.

B. Content -based Filtering

Content-based recommender system uses this approach to utilize a wide set of distinct characteristics of an item/content so as to provide additional items' recommendation with similar characteristics.

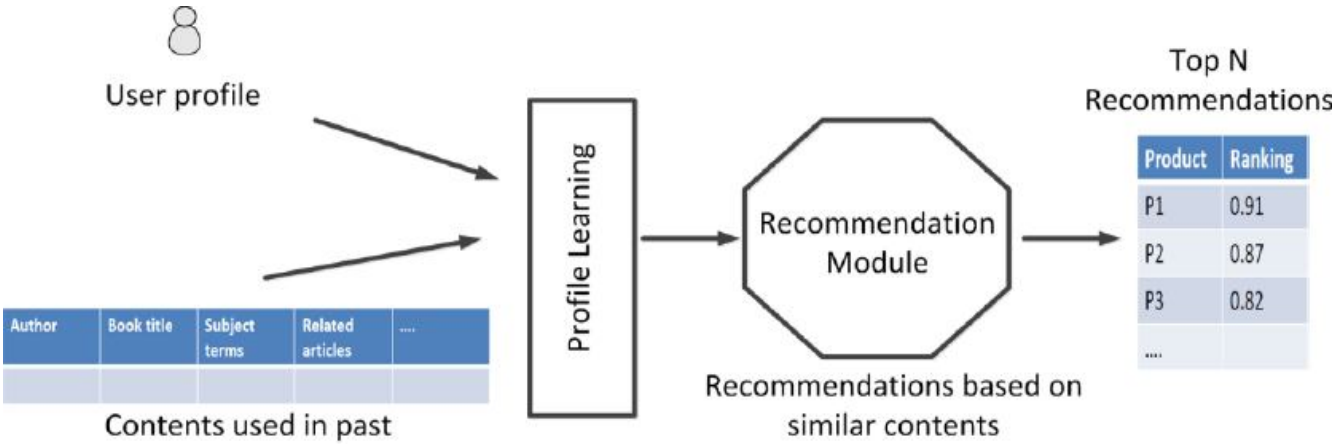


Fig-3: A content-based filtering algorithm to make an effective recommendation.

The recommender system based on content-based filtering uses the idea of content-based filtering in the manner that if a person, say P1, likes item I1, I2, I3 whereas another person, say P2, likes item I2, I3, I4 then both the person has similar interests and so it has to recommend that the person P1 should like item I4 whereas person p2 should like item I1. The system works entirely on the basis of the past behaviour under this philosophy rather than the context.

©IJRASET (UGC Approved Journal): All Rights are Reserved

C. Hybrid Filtering

The recommender system at most of the times combine the idea of both collaborative filtering and content-based filtering to provide effective recommendations of related information items/contents.

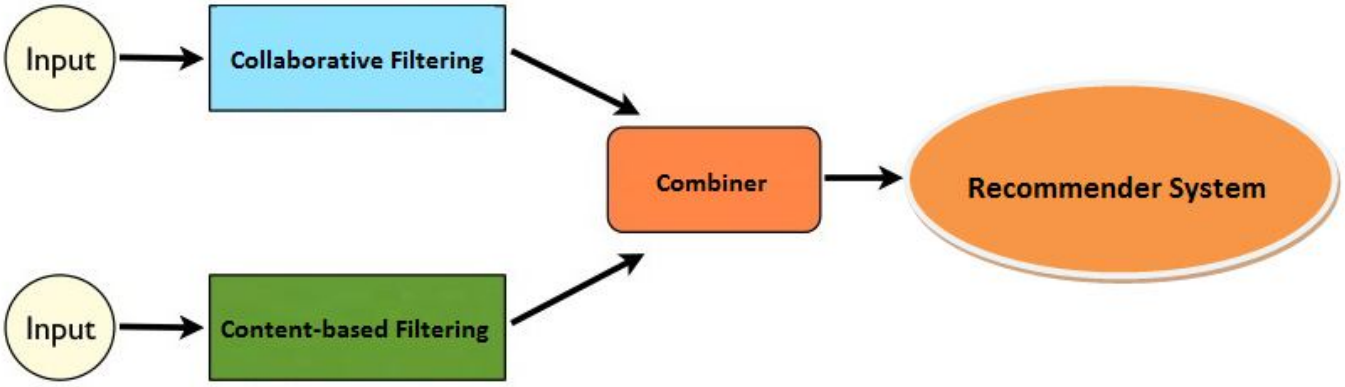


Fig-4: A hybrid filtering algorithm to make an effective recommendation.

A system that uses the hybrid filtering to overcome the issues associated with both kinds of the filtering approaches.

V. POSSIBLE WAY TO DESIGN A RECOMMENDER SYSTEM ALGORITHM

Let us use the mixed approaches of collaborative filtering and content-based filtering to analyse the possible way to design a recommender system to do recommendations of related items in a particular field.

A. First Step - Collaborative Filtering

Through collaborative filtering, the proposed system algorithm works on the idea that if one has liked an item in the past then one will also like a similar item in the future, and on the basis of one's like, others will be recommended to like the similar kinds of items. This assumption is used to build a model on the basis of user's behaviour that incorporates both kinds of data – explicit and implicit to make the proposed system work in an effective manner.

1) The general instances of collecting explicit data include:

a) To ask a user to do a rating of an item on a sliding scale.
b) To ask a user to do searching online.
c) To ask a user to do a ranking of a collection of items that ranges from favourite to least favourite.
d) To ask a user to choose the better item from the presented two item sets.
e) To ask a user to provide a set of items that liked it.

2) The general instances of collecting implicit data include:

a) To observe the items viewed by a user on an online store.
b) To analyse the viewing times item/user.
c) To maintain a record of the items purchased by a user online.
d) To obtain a list of items listened/watched by a user online/offline.
e) To analyse the likes and dislikes of a user on a social network site on which he/she makes a regular visit.

B. Second Step - Content-based Filtering

Through content-based filtering method, the proposed system algorithm works on the basis of a description made for the item and a profile build around the user's preference. Under this method, the system uses keywords that are used to describe the items and then to build a user profile so as to imply the type of item that the user likes. This simply means that the system makes use of the content-based filtering algorithms to recommend items that are very much similar to those liked by the user in the past. With the help of such comparison of the present searched item with the previously purchased item help to recommend the best-matching items.The proposed system can apply an item presentation algorithm to abstract the features of the items in the system. tf–idf representation or vector space representation is the widely used algorithm for item presentation.

©IJRASET (UGC Approved Journal): All Rights are Reserved

VI. CHALLENGES BEFORE THE RECOMMENDATION SYSTEM ALGORITHM DESIGNING

There are few major issues with the underlying approaches discussed in this paper to develop a recommender system algorithm that includes the handling of multi-tenancy, storing and processing a massive amount of data and other softer concerns like keeping the client's sensitive data safe on remote servers. The approaches suggested in this paper to design a recommender system algorithm have the following further limitations:

A. There are three major issues with collaborative filtering approach:

1) Cold start: It implies that the proposed system, most often, requires a huge amount of user associated existing data so as to make an effective recommendation.

2) Sparsity: It implies that although the e-commerce sites have extensively large number of items to be sold, the active users are able to rate a very small of items from the overall database. As a result of this, even the most popular items could achieve less ratings.

3) Scalability: It implies that a large amount of computation power is generally needed to make effective and personalized recommendation due to the extensively large number of users and products are available online.

The content-based filtering methodology has the major issue in the form of an uncertainty whether the proposed system is effective enough to learn user preferences from the actions of the users related to a particular content source and later on using them throughout the content types. For instance, news articles that are recommended based on the browsing of news are useful, but it would be much more useful when videos, music, products, and discussions from different services can be recommended based on news browsing.

VII. CONCLUSIONS

The field work of recommender system is quite fascinating, useful and so is the challenging. The filtering of a huge amount of data associated with items and users is the main characteristic of this area of research. In this paper, we have tried to give an outlook on the possible way to design a recommender system algorithm. We want to suggest that the use of machine learning algorithms and the technology of Python and R language would be quite beneficial to design an effective, personalized and efficient recommender system.

V. ACKNOWLEDGMENT

We want to acknowledge all the dear authors and co-authors of various research articles and conference papers that are described in the "References" section of this paper. We would also like to thank all our dear friends associated with the research work of machine learning and recommender system who have guided us throughout the writing of this paper to make the completion of the paper in the present form.

REFERENCES

[1] Jain, Arshay. Quick Guide to Build a Recommendation Engine in Python. JUNE 2, 2016. Web. https://www.analyticsvidhya.com/blog/2016/06/quick-guide-build-recommendation-engine-python/#comment-11173
[2] Jenson, Graham. A List of Recommender Systems and Resources. June 20, 2017. Web. https://github.com/grahamjenson/list_of_recommender_systems
[3] D Jannach. Recommender Systems. An introduction. PDF. Semantic Scholar
[4] Ricci and Lior Rokach and Bracha Shapira, Introduction to Recommender Systems Handbook, Recommender Systems Handbook, Springer, 2011, pp. 1-3
[5] Facebook, Pandora Lead Rise of Recommendation Engines - TIME". TIME.com. 27 May 2010. Retrieved 1 June 2015
[6] H. Chen, A. G. Ororbia II, C. L. Giles ExpertSeer: a Keyphrase Based Expert Recommender for Digital Libraries, in arXiv preprint 201
[7] H. Chen, L. Gou, X. Zhang, C. Giles Collabseer: a search engine for collaboration discovery, in ACM/IEEE Joint Conference on Digital Libraries (JCDL) 201
[8] Alexander Felfernig, Klaus Isak, Kalman Szabo, Peter Zachar, The VITA Financial Services Sales Support Environment, in AAAI/IAAI 2007, pp. 1692-1699, Vancouver, Canada, 2007
[9] Pankaj Gupta, Ashish Goel, Jimmy Lin, Aneesh Sharma, Dong Wang, and Reza Bosagh Zadeh WTF:The who-to-follow system at Twitter, Proceedings of the 22nd international conference on World Wide Web 2008-02-02
[10] Montaner, M.; Lopez, B.; de la Rosa, J. L. (June 2003). "A Taxonomy of Recommender Agents on the Internet". Artificial Intelligence Review. 19 (4): 285–330. doi:10.1023/A:1022850703159.
[11] Adomavicius, G.; Tuzhilin, A. (June 2005). "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions". IEEE Transactions on Knowledge and Data Engineering. 17 (6): 734–749.
[12] , J. L.; Konstan, J. A.; Terveen, L. G.; Riedl, J. T. (January 2004). "Evaluating collaborative filtering recommender systems". ACM Trans. Inf. Syst. 22 (1): 5–53. doi:10.1145/963770.963772

©IJRASET (UGC Approved Journal): All Rights are Reserved

[13] Beel, J.; Genzmehr, M.; Gipp, B. (October 2013). "A Comparative Analysis of Offline and Online Evaluations and Discussion of Research Paper Recommender System Evaluation" (PDF). Proceedings of the Workshop on Reproducibility and Replication in Recommender Systems Evaluation (RepSys) at the ACM Recommender System Conference (RecSys)
[14] Walia, P.; Singh, V.; Singh, M. (26 April 2016). "A Scientometric Analysis of Research in Recommender Systems" (PDF). Journal of Scientometric Research: 71–84. doi:10.5530/jscires.5.1.10
[15] S. Breese; David Heckerman & Carl Kadie (1998). Empirical analysis of predictive algorithms for collaborative filtering. In Proceedings of the Fourteenth Conference on Uncertainty in artificial intelligence (UAI'98)
[16] Sarwar, B.; Karypis, G.; Konstan, J.; Riedl, J. (2000). "Application of Dimensionality Reduction in Recommender System A Case Study
[17] Fig-2 by van der Goes, Maurits. Collaborative Filtering: Creating the Best Teams Ever. Web. 20 Nov, 2017
[18] Fig-3 by ReseaechGate. A Survey on Context-aware Recommender Systems Based on Computational Intelligence Techniques. Web. Feb 2015.

ABOUT AUTHORS

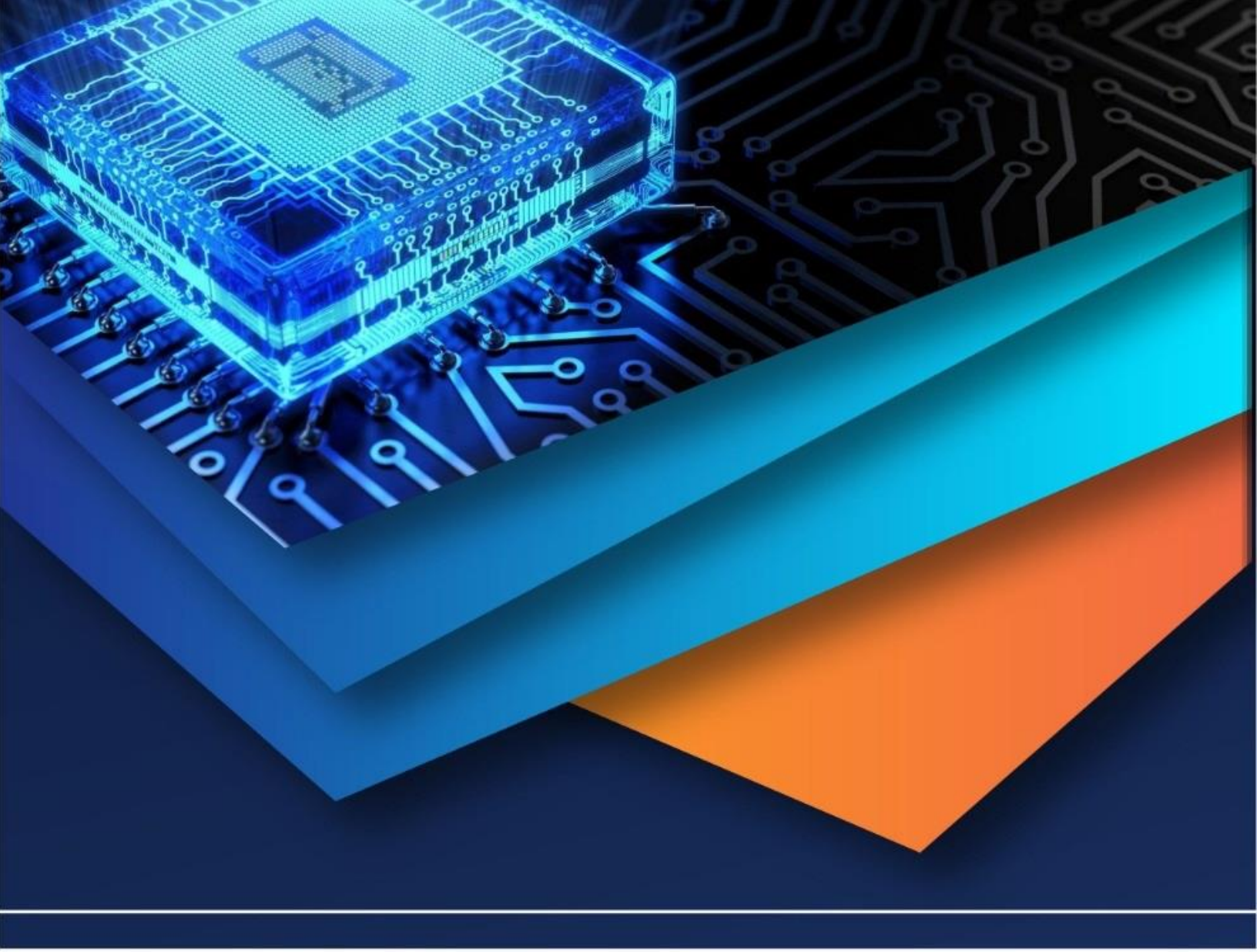


Manish Kumar Singh has published two research papers in international journals on web mining – chiefly “Web Mining: Penning an Era of Information Age" and “Understanding How Crucial Hidden Value Discovery In Data Warehouse Is?" and he has also published seven articles in international journal.



Dinesh Prasad Sahu received the Master degree (Computer Science & Application) M.Tech (Computer Science & Application) from Jawaharlal Nehru University, New Delhi, India. Currently, He is doing Ph.D. (Computer Science & Engineering) under the guidance of Dr. Karan Singh, from Jawaharlal Nehru University, New Delhi, India & is working in school of Computer & Systems Sciences, Jawaharlal Nehru University, New Delhi. His primary research interests include parallel and distributed system and Grid Computing. He has published 3 papers in proceedings of peer-reviewed Conferences.

©IJRASET (UGC Approved Journal): All Rights are Reserved



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)