iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** III **Month of publication:** March 2020

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: 7.429
Volume 8 Issue III Mar 2020- Available at www.ijraset.com

A Survey on different Citation Recommendation Algorithms using Machine Learning

Rejoice Wilson E[1], Praveen. A. N[2]
[1]*Dept. of information technology Government Engineering College, Idukki, Kerala*
[2]*Assistant professor in IT Government Engineering College, Idukki, Kerala*

Abstract: ***Machine learning a supplement using which ma- chines can be made intelligent, this means they can take decisions on their own, classify things, predict things, recommend things based on your likes. The Recommendation systems use machine learning algorithms to provide users with product or service recommendations. Recently, these systems have been using machine learning algorithms from the field of artificial intelligence and data science. A recommendation method for citations simply focuses on locating the appropriate publications.***
Recommender systems are divided into Three main categories, depending on the information used to drive the recommendations. Collaborative, Content-Based, Graph based filtering. Using these algorithms citation recommendations systems recommend the related publications.
This paper discusses different machine learning algorithms for recommend the related publications and performance analysis of these algorithms in various studies.
Index Terms: Citation, Recommendation, Filtering

I. INTRODUCTION

Throughout recent years, more and more research papers have been published, and it is difficult for scholars to track proper papers on time.

One approach for solving this problem is to use keywords to collect papers, but researchers still need to manually go over the papers they have obtained and determine which paper should be selected. Another approach is to trace reference papers from other scholarly publications that fascinate researchers immensely, but it restricts reference papers within a specific research submitted. Recommendation for citation, which can effectively locate relevant research papers based on a given manuscript or text fragment, is a critical technique for addressing this problem.

II. CITATION RECOMMENDATION

A. Recommendation for citation during the Modern Scholarly

Data Age is of great importance to scientific research. The role deals with different types of artifacts (e.g., text, author, etc.) and types of relationships, which naturally form a heterogeneous information network (HIN).Approaches for citations are known as global citation recommendations and local citation

Identify applicable funding agency here. If none, delete this. recommendations. The role of global recommendations is to find relevant articles concerning a given corpus as a whole. The role in local recommendations is to identify local contexts (e.g., sentences, paragraphs in a paper) for which the literature includes relevant articles. In this paper we review guidelines for global citations.

B. Global Citation Recommendations

Approaches for global citation recommendations machine learning algorithms are listed as: content-based filtering (CBF) collaborative filtering (CF) and graph-based approaches

1) *Content-Based Filtering (CBF):* Recommendation Framework with a content-based approach to filtering bases its recommendations on the item data they can access. CBF recommends scientific papers close to those that have been of interest to researchers in the past. Once researchers browse a particular paper, the RS collects information about that computer and searchers in a computer database with similar attributes, such as the name of the article, author and position. The result of this search will then be returned as suggestions to the user.

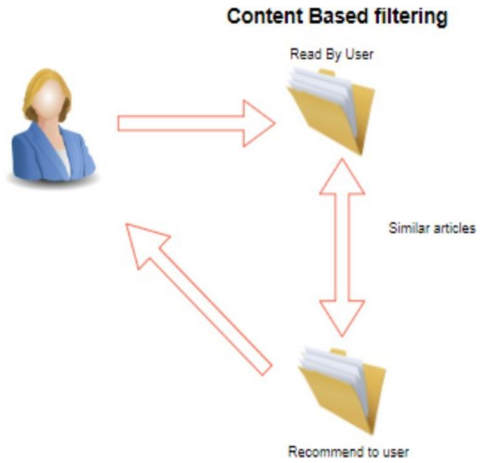


Fig. 1. Content-Based Filtering (CBF).

2) *Collaborative Filtering (CF):* Using a collaborative method Recommendation Framework considers user data when storing information for recommendation.CF relies on finding researchers with similar interest in research and making use of their ratings to offer recommendations. The artificial intelligence and recommendation system will classify users who share the same paper interest and then suggest that similar users search for paper.

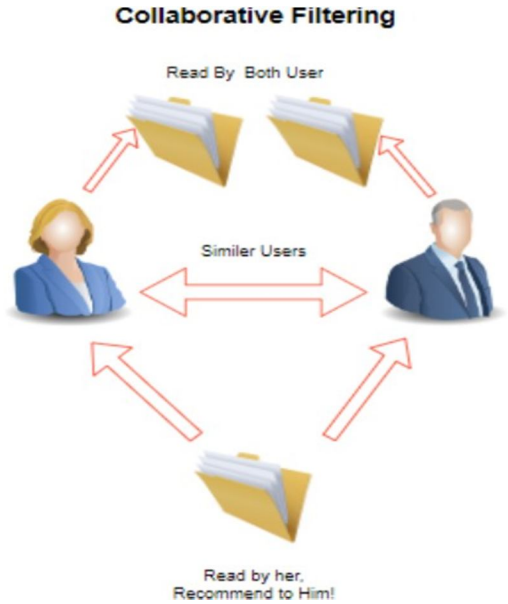


Fig. 2. Collaborative Filtering (CF) .

3) *Graph-based filtering (CF):* Graph-based recommender framework inherently merging content-based approaches with collaborative approaches. recommending items based on the consumer and the item details. For example, a recommendation system on a social network can suggest user-like profiles by comparing their interests. The program can then find the suggested profiles as objects and thus access their data to search for new profiles of similar nature. In the end, it returns both sets of profiles as recommendations.

Fig. 3. Graph-based filtering (CF) .

III. LITERATURE REVIEW

This literature review focus on various machine learning algorithms used for citation recommendation and to compare their performance

Haifeng Liu, Xiangjie [1] Proposed a Citation Framework (CCF) for citation of relevant papers as target document references. CCF measures similarities between the citations by comparing pairly paper representations obtained from citation context. Collaborative Filtering(CF) method is used in this proposed model. The reasoning behind this measurement of similarity is that citing papers is assumed to be identical if they co-occur with the same citing papers.

Xiaoyan Cai, Junwei Han [2] Propose a citation recommendation approach by reinforcing each other on a three-layer graph in which each paper author or venue is represented as a vertex in the paper layer, author layer and venue layer respectively. Graph-based citation filtering is used in this pro- posed model. Three-layered dynamic clustering on a three- layer graph built to solve the graph's enormous size problem. Shakila Shaikh and Dr. Sheetal Rathi [3] This paper also focuses on how graph algorithm can be used in ecommerce websites to improve recommendations. The system proposed compares image suggestion flickr.com with the process pro- posed. The technique incorporates semantonal recommendations using graph-based overlap methodology.

Mr Arif Muhammed [4] Propose a method of citation recommendation using citation relationships and comparison be- tween many other articles. The basic method consists of cross- reference references for recommendations. When one paper is co-occurred in two or more citing papers then to some degree they are identical. After that, to get parallels between them, these citing papers are linked in pairs to their contents. Bansari Patel, Palak Desai [5] propose a new recommendation structure applying mutual reinforcement rules on a multi- layered graph that consists of all available multiple types of relationships between author, paper and key word entities from the collection of literature. Importantly, to achieve a query-focused and mutually validated recommendation, the customized query information is also combined with multi- layered mutual reinforcing rules.

Waleed Waheed, Muhammad Imran [6] proposed recommender method called citation Paper Network and Authors ' Relationship Network (CNRN) is based on Multilevel Simultaneous Citation Network (MSCN) and overcomes issues when either old or new papers are selected as interest papers. The suggested approach to CNRN was contrasted with both the Google Scholar and MSCN benchmarks using NDCG, MRR and AP metrics.

Marco Gori and Augusto Pucci [7] present a random-walk scoring algorithm that can be used to suggest papers based on a small collection of relevant articles selected by the user. They checked the algorithm on a data set collected by ACM Portal Digital Library and it performed very well, in fact target papers got a ranking within the first 20 positions of the score.

Cai et al. [8]Proposed an embedding heterogeneous information network (HINE) method, it can capture inter-relationships between heterogeneous vertices, intra-relationships between homogeneous vertices and correlations between vertices and text contents at the same time, representing different types of objects as vectors in a continuous and common vector space. They also suggested a model for profound network representation

Yang et al.[9] proposed an approach based on the LSTM. They first studied vector representations of citation contexts and scientific papers separately, then evaluated significance between them based on the representation of the learned vectors. Finally, high scoring scientific papers are produced as recommended list of reference papers.

Mamoulis [10] studied the issue of heterogeneous network embedding for proximity based on meta direction, which can completely leverage the network heterogeneity.

Dong et al.[11] Built a meta-path-guided random walk strategy in a heterogeneous network that includes multiple node types within the network. Chang et al. [12] Deep embedding algorithm designed for networked data. The algorithm represents the architectures of local and global networks and makes the resulting embedding useful for a number of data mining tasks.

Gupta and Varma [13] Addressed the issue of scientific paper recommendations by means of a new method combining network representation from the bibliographic network with information on content

Wang et al. [14] A predictive network representation approach was presented but node content information was ignored which reduces network representation efficiency.

Kobayashi et al. [15] Proposed a context-based approach, integrates distributed representations of text and citation maps, and learns distributed vector representations of articles, with each vector capturing facets of discourse within an article.

Hu et al.[16] proposed a deep neural network to obtain distributed representations of users, objects and meta-path at the same time Chen and Sun [17] proposed an embedding architecture for a directed and path-augmented heterogeneous network.

Jacob et al.[18] proposed a heterogeneous social network with embedded node classification algorithms

Tao et al.[19] Proposed a HIN2Vec model that learns both node vectors and relationship vectors by maximizing the likelihood of predicting node relationships jointly

IV. COMPARISION BETWEEN DIFFERENT ALGORITHMS

Here is the comparison of that different methods or tech- niques after the survey of different methods or techniques in the recommender system Graph based filtering technique used in recommender method is most widely used from survey. It's learned more over other techniques.

There are three filtering methods in the recommender system for offering the user suggestions / recommendations for things not yet seen by the user. Table 1 provides a list of different recommendation techniques

The graph-based filtering technique provides more advantages than other techniques.it consider the hiden edge relationship between the nodes. The main advantage over content-based fil- tering technique is that it increases the recommender system's efficiency and gives better recommendation as it also considers the interest of other similar users and the past history of similar users to give the user recommendation / suggestions.

TABLE I.
Comparision Between Recomender Techniques

Recommender system methods	Content based filtering	collaborative filtering	Graph based filtering
Number of users	Based on single user	Based on many users having similar interest	Combination of content based and collaborative filtering
Advantages	User independence, Transparency	Improve recommendation performance	overcome cold start, sparsity problem
Disadvantages	Limited content analysis, New user	Data sparsity, Scalability, Synonym	Increased complexity, expensive in implemetation

V. CONCLUSION

In this paper, we have reviewed literature on different recommender system approaches. It has been found and inferred from the study of different methods that graph based filtering algorithm has more advantages over other recommender system techniques. Each of these algorithms has its own advantages and disadvantages. Survey on various recommender program focuses primarily on finding ways to improve recommendation efficiency, accuracy, and performance. Some hybrid approaches are used to improve the quality of the recommender system but it does have some disadvantages. Implementation of these was more complicated and expensive.

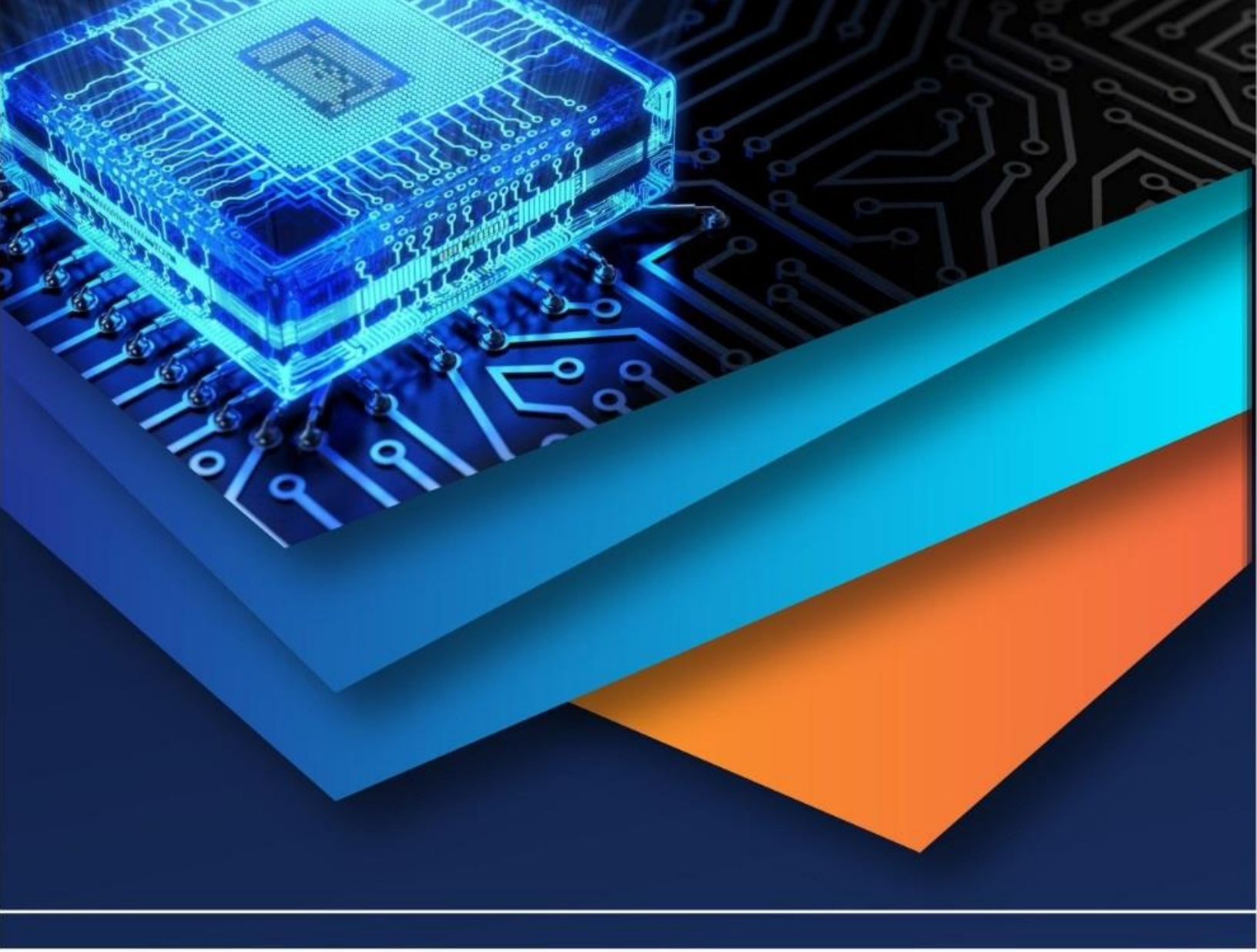
VI. ACKNOWLEDGMENT

I would like to express my gratitude to my guide Praveen. A.N for giving me valuable advice and direction. I would like to thank the department of information technology for their support.

REFERENCES

[1] Haifeng Liu, Xiangjie Kong, Xiaomei Bai, Wei Wang, Teshome Megersa, AND Feng Xia, "Context-Based Collaborative Filtering for Citation Recommendation" School of Software, Dalian University of Technology, Dalian 116620, China.Vol.3, 2015.
[2] Xiaoyan Cai, Junwei Han , Wenjie Li, Renxian Zhang, Shirui Pan, and Libin Yang "A Three-Layered Mutually Reinforced Model for Personalized Citation Recommendation " 2162-237X © 2018 IEEE.
[3] Shakila Shaikh and sheetal Rathi "Recommendation system in E- commerce websites: A Graph Based Approache"2017 IEEE 7th Inter- national Advance Computing Conference
[4] Arif Muhammed A "Content Aware Citation Recommendation System," 978-1-5090-3751-3/16/2016 IEEE .
[5] Bansari Patel and Palak Desai "Methods of Recommender System: AReview" 2017/ICIIECS.
[6] Waleed Waheed, Muhammad Imran "A Hybrid Approach Toward Re- search Paper Recommendation Using Centrality Measures and Author Rankin,"VOLUME 7, 2019.
[7] Marco Gori,Augusto Pucci "Research Paper Recommender Systems: A Random-Walk Based Approach "2006 IEEE.
[8] X. Cai, J. Han, S. Pan, and L. Yang, "Heterogeneous information network embedding based personalized query-focused astronomy ref- erence paper recommendation," Int. J. Comput. Intell. Syst., vol. 11, pp. 591–599, Jan. 2018.
[9] L. Yang, Y. Zheng, X. Cai, H. Dai, D. Mu, and L. Guo, "A LSTM based model for personalized context-aware citation recommendation," IEEE Access, vol. 6, pp. 59618–59627, 2018. doi: 10.1109/ACCESS. 2018.2872730.
[10] Z. Huang and N. Mamoulis. (2017). "Heterogeneous information net- work embedding for meta path based proximity." [Online]. Available: https://arxiv.org/abs/1701.05291

[11] Y.Dong,N.V.Chawla,andA.Swami,"metapath2vec:Scalable representation learning for heterogeneous networks," in Proc. 23rd SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2017, pp. 135–144.
[12] S. Chang, W. Han, J. Tang, G.-J. Qi, C. C. Aggarwal, and T. S. Huang, "Heterogeneous network embedding via deep architectures,"inProc.21st SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2015, pp. 119–128
[13] S. Gupta and V. Varma, "Scientific article recommendation by using distributed representations of text and graph," in Proc. 26th Int. Conf. World Wide Web Companion, 2017, pp. 1267–1268
[14] Z.Wang,C.Chen,andW.Li,"Predictive network representation learning for link prediction,"inProc.40thSIGIRInt.Conf.Knowl.DiscoveryData Min- ing, 2017, pp. 969–972.
[15] Y. Kobayashi, M. Shimbo, and Y. Matsumoto, "Citation recommendation using distributed representation of discourse facets in scientific articles," in Proc. 18th JCDL, 2018, pp. 243–251.
[16] B.Hu,C.Shi,W.X.Zhao,andP.Yu,"Leveraging meta-path based context for top-N recommendation with a neural co-attention model,"inProc.24th SIGKDDInt.Conf.Knowl.Discovery Data Mining,2018,pp.1531–1540.
[17] T. Chen and Y. Sun, "Task-guided and path-augmented heterogeneous network embedding for author identification,"inProc.10th WSDM,2017, pp. 295–304.
[18] Y. Jacob, L. Denoyer, and P. Gallinari, "Learning latent representations of nodes for classifying in heterogeneous social networks," in Proc. 7th WSDM, 2014, pp. 373–382.
[19] T. Fu, W. C. Lee, and Z. Lei, "Hin2vec: Explore meta-paths in hetero- geneous information networks for representation learning," in Proc. 26th CIKM, 2017, pp. 1797–1806.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY