



LEEDS  
BECKETT  
UNIVERSITY

---

Citation:

Abbasi, II and Abbas, MA and Hammad, S and Jilani, MT and Ahmed, S and Nisa, SU (2020) A Hybrid Approach for the Recommendation of Scholarly Journals. In: 2020 International Conference on Information Science and Communication Technology (ICISCT). IEEE, pp. 1-6. ISBN 978-1-7281-6899-9 DOI: <https://doi.org/10.1109/ICISCT49550.2020.9080032>

Link to Leeds Beckett Repository record:

<http://eprints.leedsbeckett.ac.uk/id/eprint/7282/>

Document Version:

Book Section

---

© 2020 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

The aim of the Leeds Beckett Repository is to provide open access to our research, as required by funder policies and permitted by publishers and copyright law.

The Leeds Beckett repository holds a wide range of publications, each of which has been checked for copyright and the relevant embargo period has been applied by the Research Services team.

We operate on a standard take-down policy. If you are the author or publisher of an output and you would like it removed from the repository, please [contact us](#) and we will investigate on a case-by-case basis.

Each thesis in the repository has been cleared where necessary by the author for third party copyright. If you would like a thesis to be removed from the repository or believe there is an issue with copyright, please contact us on [openaccess@leedsbeckett.ac.uk](mailto:openaccess@leedsbeckett.ac.uk) and we will investigate on a case-by-case basis.

# A Hybrid Approach for the Recommendation of Scholarly Journals

Iqra Iftikhar Abbasi  
*University Institute of Information  
 Technology*  
*Pir Mehr Ali Shah Arid Agriculture  
 University Rawalpindi*  
 Rawalpindi, Pakistan  
 iqraabbasi901@gmail.com

Muhammad Azeem Abbas  
*University Institute of Information  
 Technology*  
*Pir Mehr Ali Shah Arid Agriculture  
 University Rawalpindi*  
 Rawalpindi, Pakistan  
 azeem.abbas@uaar.edu.pk

Shiza Hammad  
*University Institute of Information  
 Technology*  
*Pir Mehr Ali Shah Arid Agriculture  
 University Rawalpindi*  
 Rawalpindi, Pakistan  
 shizahammad122@gmail.com

Muhammad Taha Jilani  
*Karachi Institute of Economics and  
 Technology*  
 Karachi, Pakistan  
 mtaha.jilani@gmail.com

Shabbir Ahmed  
*University Institute of Information  
 Technology*  
*Pir Mehr Ali Shah Arid Agriculture  
 University Rawalpindi*  
 Rawalpindi, Pakistan  
 azeem.abbas@uaar.edu.pk

Saba un Nisa  
*University Institute of Information  
 Technology*  
*Pir Mehr Ali Shah Arid Agriculture  
 University Rawalpindi*  
 Rawalpindi, Pakistan  
 azeem.abbas@uaar.edu.pk

**Abstract**—The increasing number of scholarly journals have made it difficult for authors to select the most suitable journal that publishes their research. Existing search systems that recommend journals for manuscript submission are either based on author's profile, bibliographic data or the co-publication network. These approaches are not useful for beginner researchers who have no publication records or for those who are interested in new research domains. The present work proposes a hybrid approach that combines clustering and document similarity for the recommendation of scholarly venues. The proposal was evaluated both objectively and subjectively using domain experts. The results of mean average precision (0.84) and normalized discounted cumulative gain (0.89) shows positive recommendations made by the proposed approach.

**Keywords**—*recommendation system, journal recommendations, clustering.*

## I. INTRODUCTION

In today's diverse and multifaceted research environment, finding the relevant scholarly journal for publishing article among thousands of available journals is challenging. The increasing number of scholarly journals publishing academic work makes it difficult for authors to identify scholarly journals that publish research results close to their field [2].

In the age of proliferated publication venues, it is extremely necessary for researchers to be aware of well-known and most relevant scholarly journal for publication. Because, wrong selection of venue may cause long delays in publication, which can sabotage the academic career of a researcher and impose a negative impact on the academic performance. For instance, the research article is rejected after a long review process with the notion that the work does not match the scope of the journal. One of the main reasons for rejections is that the submitted articles are not close to the scope of the scholarly journal as many scholarly journals have a broad range of topics and many research articles involve multiple academic disciplines [10]. Therefore, researchers are needed to submit their articles to the relevant and high quality academic venues. The impact of publication journals is an important consideration for authors when they are seeking to

publish their manuscripts [9]. The initial step in finding a scholarly journal to publish is identifying the type of manuscript because different scholarly journals are publishing different types of articles. It is essential for the researchers to choose a journal that publishes relevant to their research results.

Journals are the venues where authors can publish their manuscripts for the use of community. Choosing the right scholarly journal to submit a paper is a primary concern for researchers and they need to know about several characteristics of scholarly Journals that are difficult to obtain [11] including aim and scope, acceptance rate, publication matrices, impact, aim and scope and article processing time [16]. On the basis of aim and scope journals differ from each other as all journals have their own aim and scope that defines their areas of publications. Therefore, the present work uses aim and scope of the journals from multiple categories of research domains to build recommendations. However, the present work limits the scope by using the scholarly journals from computer science domain only that cover the topics of artificial intelligence (AI) and robotics, Computer and engineering by Thomson Reuters.

The well-known methods to find the publication venues are publication history, personal venue rating [2], bibliographic data and co-publication network [12,14,18]. Participation or publication history of authors refers to the scholarly activities like published and cited papers by the target author. Personal venue rating (PVR) measures the cited references and the year and time in which they were added to a researcher's library are used for pertaining a researcher's interest to make the venue recommendations [2]. Bibliographic data refers to the citation relationship between papers from co-citers, co-authors and co-affiliated. It includes publication history of co-authors, cited papers and the research interest of the users working in the same organization [15]. These approaches are not useful for beginner researchers who have no publication records or for those who are interested in other (new) research domains.

The current research study identifies the gaps in finding the publication venues and proposing a hybrid approach which

combines with clustering and document similarity that manages multiple publishers for the recommendation of scholarly journals. In this paper, we make the following contributions: the detailed analysis of the existing approaches for recommendation of scholarly journals is presented, the current research aims to facilitate the beginner researchers with appropriate scholarly journals for the submission of their research results and the proposed system considers the researchers current field of research interest in terms of abstract of their current article to generate relevant ranked journal list even if the researcher has no publication history or he/she wants to enter in a new research area. Additionally, our proposal is not publisher specific as we are managing multiple publishers. For the evaluation purpose both objective and subjective evaluation was performed. Extensive experiment by using aim and scope of the scholarly journals was conducted and in comparison, to relevant approaches the proposed approach performs better in terms of recall, precision and f-measure. Further, results are evaluated by subjective evaluation as well that is based on expert's opinion.

The rest of the paper is structured as follows; Section 2 provides a detailed review of the existing venue recommendation approaches. In Section 3, we describe the proposed hybrid approach for the recommendation of scholarly journals. An evaluation of the recommendation model is provided in Section 4 with the discussion of achieved results and conclude the paper in Section 5 with the suggestions for future work.

## II. LITERATURE REVIEW

In previous studies, most of the research attempts focused on researcher's interests for the recommendation of scholarly journals. Recommendations were generated through analyzing researcher's behavior. Recommendation systems are introduced for a large number of real life domains. However, the current study reviews the most relevant to the proposed approach. The following literature review will contribute to our work.

Recommender systems (RS) are used for dealing with information overload. Early attempts in scholarly venue recommendation systems were based on researcher's profile [2]. They used CiteULike dataset to conduct the experiment. To recommend scholarly venues to researchers they compared authors with similar interests in terms of their personal venue ratings (PVRs). A hybrid recommender system [6] is proposed in the field of computer science for upcoming conferences. Their proposal was based on academic venues from co-citers, co-authors and co-affiliated researchers.

CF (Collaborative filtering) is a technique used in recommender systems. CF examines the relationships between the users and the recommended items to the target users according to the ratings provided by their neighbors [14]. They make Clusters of the users based on their social networks by the use of the number of papers a researcher had published in a scholarly venue to derive the researchers rating for that specific venue. They used co-author publication history. Their work recommends the upcoming events or special issues in which researcher might be interested.

Research is a primary task for scholarly communities nowadays and the information is growing exponentially on the web, the information overload represents a great challenge to the academic world. Tejada et al. [15] present a quality-based fuzzy linguistic recommender system (REFORE) for

researchers by using some bibliometrics measures. Their proposed system takes the measured quality as the main factor for the re-ranking of the top-N recommendations list in order to point researchers to the best and latest papers in their domain. A framework [17] is provided for scholarly venue recommendation as both papers' inter-relationships and inter-similarity. In this work, they introduce the memory-based CF (collaborative filtering) for academic venue recommendation and extract stylometric features to measure the similarity between papers in terms of their writing styles.

Choosing a relevant scholarly journal to submit a paper is one of the most important step for authors when they are looking to publish their research results. Kang et al. [10] introduced the Elsevier journal finder, a scholarly journal recommender system that works for all major scientific domains under 2900 peer reviewed Elsevier journals. Their recommendations are limited to Elsevier journals only.

Recently, some online services have started to recommend relevant scholarly journals using keyword, title and abstract matching. These services are EndNote manuscript matcher, Elsevier journal finder, JANE, Springer journal selector, Cofactor Journal Selector, Deans Journal Selector and IEEE publication recommender. Journal/Author Name Estimator, Cofactor Journal Selector and Journal guide recommends only in biomedical domain [9]. These services are proposed by the publishers and limited to only searching their own pool of publications, which assumes that an author begins their decision process by first choosing a publisher. An algorithm [1] is developed using python library that builds a recommendation system based on the content of articles and make suggestions using author assigned keywords. They perform testing on 15K posters of the Neuroscience Conference 2015. Their proposed system only works for the conferences in the society of Neuroscience. Choosing where to submit research results among thousands of available journals is challenging. Likis [11] describes the important aspect of writing for publication is choosing a suitable scholarly journal. Publication is important for the researchers because the future of profession depends upon it. Moreover, finding a suitable home among thousands of available scholarly journals for manuscript is now considerable.

Recommender systems help authors to mine relevant information in big scholarly data. It is necessary for authors to know high quality academic venues of their research interest. PAVE (personalized academic venue recommendation) [18] is proposed that is based on co-publication networks. Their proposal consider co-author relation and author-venue relation using DBLP data set. To evaluate the quality of recommended venues precision, recall and F1-measure is used.

One of the important task of recommender system is to suggest relevant venues to researchers within a city. A user model [3] based on users' reviews from Yelp by using support vector machine classifier is introduced. Alshareef et al. [4] aim to identify the suitable venues, and researchers within a citation network by using IEEE dataset. They integrate the authors' similarities, the topical similarity, and the venues' similarities among a citation network of a given article. A continuous predictive social media analysis [5] is used to build venue recommendations. They integrate deductive and inductive techniques to analyze the social network activities and to build high quality link predictions among scholarly community. They focused on location based social networks by the use of smartphones.

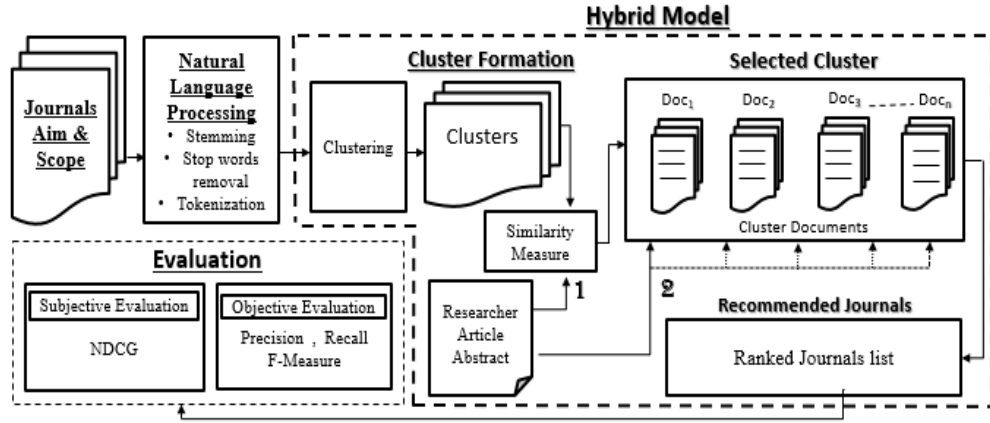


Fig. 1. Overall Architecture of the Proposed Approach

In literature, most of the research to find the academic venues have used the researchers profile that refers to the participation or publication history of authors and citation analysis to build recommendations. These approaches are not useful for students and beginner researchers (who have no scholarly publication history). Moreover, using only the publication venues in which an author has published work previously undermines the process of recommendation, as the author may be interested in other research domains in which he or she has not published any research paper. This research study identifies the gaps in finding the publication venues and proposing a tailored scholarly journal recommendation system using aim and scope of the journals, which will recommend the relevant scholarly journals to the researchers for their current field of research.

### III. SCHOLARLY JOURNAL RECOMMENDATION SYSTEM

The complete architecture of the proposed hybrid approach is shown in Figure 1 and it is divided into three main components: preprocessing, hybrid model and evaluation. We conducted the experiment by using aim and scope of the journals from computer science discipline covering the topics of artificial intelligence (AI) and robotics, Computer and engineering from Thomson Reuters journal list. A dataset of 236 scholarly journal's aim and scope was used for the experimentation. To formulate the corpus aim and scope are directly retrieved from the journal's website and are stored in the text files. These text files refers to the documents having aim and scope of the journals.

#### A. Preprocessing

In the next step preprocessing is performed by using natural language processing (NLP) to clean the dataset. NLP includes different steps for manipulating and interpreting human language, however in the proposed work stemming, stop words removal and tokenization is used. One of the important step in natural language processing (NLP) is stemming which is performed to reduce words to its base form. Stop words are removed to extract the meaningful terms from the text. In the next step tokenization is performed on the text. The proposed approach uses word tokenization to mine unique words and to remove unwanted words from the sentences by tokenizing the text into a sequence of tokens using the library Nltk. Scikit learn library is used in python for the preprocessing.

#### B. Hybrid Model

In this section, we discussed the details of the proposed hybrid recommendation strategy which combines clustering and document similarity approaches. In Figure 1 step1 and step2 represents clustering and document similarity that merges to form hybrid model.

*a) Clustering:* Clustering methods aims to divide the dataset into sub sets (clusters) that reduces the computational load. In the proposed system a popular clustering method k-means clustering is used. To implement k-means clustering text is preprocessed and tokenized vocabulary is generated. In the next step term-frequency inverse document frequency (tf.idf) vectorizer is applied on vocabulary and number of clusters are formulated. A total number of 10 clusters are generated, the cluster-id's and their respective number of journals of each cluster are shown in Table 1 and journal names can be seen in appendix A.

Unique terms from each cluster are extracted and a prediction function is used in python which predicts that query falls in which cluster. Our proposed system is based on the mixed hybrid approach in which initially the system produces the clusters of the scholarly journals on the basis of similarity and then using a prediction function in python which predicts that the input document falls in which cluster. Document refers to the abstract of the researcher's article. Once the cluster is suggested the journals within the cluster are displayed. In the next step, document similarity is done where the system computes the similarity between the input document and the scholarly journals of the cluster one by one by using document similarity algorithms such as Cosine similarity and Latent semantic analysis (LSA). These algorithms generates recommendation of the most similar scholarly journals to the researcher's article.

*b) Document Similarity:* To compute the document similarity and for the analysis of the journals aim and scope semantics are needed. However, to deal with semantics latent semantic analysis (LSA) is used. To implement LSA we created a dictionary from the data then convert to bag-of-words corpus which result in a term-document matrix(occurrence of terms in a document) and save the dictionary. Tf-Idf vectorizer is used, a simple transformation that takes documents represented as bag-of-words counts. After the creation of dictionary LSA model is generated. Finally, the similarity between the input document and every journal inside the cluster is computed. In result, documents

and their similarity score is represented. To do so, a python library genism is used. To implement cosine similarity, input document and journals of the selected clusters are passed to a vector and tf.idf vectorizer is applied. In the last part, cosine score is calculated and documents are ranked with respect to their cosine score. For this purpose, an open source python library sklearn is used .For instance, there are ten number of clusters and the value of k is 10. When a new document is passed ,the system will first compute the similarity of the document with all the cluster centroids and the prediction function predicts that the query falls in which cluster on the basis of similarity. Once the cluster is selected, then the system will compute the similarity of the document with the journals inside the cluster using similarity measure algorithms and on the basis of similarity score the ranked list of top-5 scholarly journals will be displayed.

TABLE I. CLUSTERS AND THEIR RESPECTIVE JOURNALS

Cluster #	No. of journals	Journal Abbreviation
1	31	ACM TOG , AEI, AAI , CAD,COMP, CCDS,EAAI, Eng Comput, FAOC, GENET PROGRAM EVOL M , TAC, TCIAIG,IET CTA, IJBIC, IJHPCA, IJSR, JHEURISTICS, J INTELL FUZZY SYST, J LOG ALGEBR METHODS, JWE, J WEB SEMANT, MACH VISION APPL,MULTIMEDIA SYST,Nat Mach Intell, Neural Comput, RE, SW, SI, Syst Control Lett, J LOGIC LANG INFORM,T I MEAS CONT.
2	27	CSUR, TOCE,TOMM, TWEB, APPL SOFT COMPUT, CLUSTER COMPUT,COMPUT INFORM,CONCURR COMP PRACT E,DC, FUTURE GENER COMP SY, IEEE T CLOUD COMPUT, IEEE T SERV COMPUT, ICAE,INT J UNCONV COMPUT,IJWGS,IJWSR,J MULT VALUED LOG S,J PARALLEL DISTR COM,JVLC, MEMET COMPUT,NAT COMPUT,NEURAL COMPUT APPL, PARALLEL COMPUT,SCI PROGRAMMING-NETH, SC, SWARM EVOL COMPUT, Theor. Comput. Sci
3	17	TACO,TOMS,TOSEM,ASE,EMPIR SOFTW ENG,IEEE Softw.,TSE, IET Softw., IST, IJSEKE,JSS, JSA,SCP,STVR, SPE,SoSyM
4	18	TRETS, AI C,APPL ONTOL,COMPUT GRAPH-UK,Geoinform, HCIS,IEEE MUL,IEEE TIFS,IET IS, IF,Inf Syst Front,IV, INT J INF SEC,IJ WMIP,IJWSIW, J CRYPTOL, UAIS,UMUAI
5	24	TODS,TKDD,TOS,APPL INTELL,Big Data Res,COMPUT INTELL, DKE,DATA MIN KNOWL DISC,Eur J Control,IEEE Cloud Comput.,IEEE TLT,ITPDS,IDA,IJAR,IJCS,IJDWM,IJAIT,IS A T,KBS,NCNS,NGC,PE,SADM,VLDB J
6	17	ACM TRANSACTIONS ON ALGORITHMS, ACM TA,AES,COMB.,CAGD, CSE,CONST.,FI,IC,IJFCS,JCA, JFP, JMIV,J ACM, LMCS,NN,SIAM JC,TCS
7	13	GM,IEEE TVCG,IJC,JRTIP, JSEP,MIA,MSSP,NEURO.,PAA,PRL,SIAM JIS,SC,VC
8	31	ACM JETCS,ACM TCS,ACM TDAES,ACM TECS,ACM TPLS,AL,AAMAS,AR,CSR,CSSE,CEE,CAS,D EDS,HCI,IEEE DT,IEEE ESL,IEEE MICRO,IEEE TC,IEEE TM,IET CDT,Autosoft.,IJGS,IJRN,JI,RS,JSC,MECH., MM,MTA,OCAM,RTS,RAS
9	23	ACM TIT,AB,CACM,CAVW,CLSS,CI,DSS,DAES,D ISP.,DPD,EC,ES,IBM JRD,IEEE CGA,IEEE TNSM,ISR,JDM,JGC, SR, STSMSI,SR, TC, VR
10	35	ACM TAR, ACM TCHI, AMC,ARC,CMTCL,CGF,CJ, CSI, CVIU,COMP

In order to measure the quality of our generated recommendation, evaluation is done which categories in to two steps Objective evaluation that includes Recall, Precision and F-measure and Subjective evaluation based on expert's opinion using NDCG (Normalized Discounted Cumulative Gain).

#### IV. EVALUATION

To measure the recommendation performance of the proposed system we used both subjective and objective evaluation. For the objective evaluation, the performance of the algorithms is evaluated by using multiple performance benchmarks. The performance benchmarks used for the current system were recall, precision and f-measure. Precision is derived by dividing the number of relevant journals by the total number of retrieved journals. For each researchers article, the top five scholarly journals were ranked constituted the precision at (P@5).

$$Precision(P) = \frac{tp}{tp + fp} = \frac{\text{number of relevant scholarly journals}}{\text{number of retrived scholarly journals}}$$

Recall is a measure of number of retrieved scholarly journals that are relevant by the total number of relevant scholarly journals.

$$Recall(R) = \frac{tp}{tp + fn} = \frac{\text{number of retrived scholarly journals}}{\text{total number of relevant scholarly journals}}$$

In f-measure we computed the harmonic mean between the calculated precision and recall.

$$F \text{ Measure}(f) = \frac{2 * Precision * Recall}{Precision + Recall}$$

For the subjective evaluation normalized discounted cumulative gain (NDCG) is used that is based on expert's opinion. Expert opinion from four experts were taken. Experts are provided by the ranked scholarly journal list generated by the respective algorithm. In addition, they are provided by the researchers article abstract and recommended journals aim and scope. They have evaluated the results by reading the abstract and top 5 journals aim & scope and provided the results by ranking (0,1,2) .if the abstract is most similar to any journal's aim and scope it gets rank(2) and (0) rank for the least similar journal by the expert. This is the example of one ranked list provided by an expert [ J14(2) ,J2(0), J3(1), J24(2) ,J5(2) ] where J14 is journal#14 and it is most similar to the query so it gets rank (2) ,similarly journal#2 is least similar so it gets rank(0) and so on. To compute NDCG, firstly cumulative gain (CG) is computed by adding the ranks provided by experts for recommended journal list. After the computation of (CG), discounted cumulative gain (DCG) is measured.

DCG (Discounted cumulative gain) measures the extent to which a scholarly journals ranking is relevant to an expert's ideal ranking, as shown below:

$$DCG_n = \sum_{i=1}^n \frac{rel_i}{\log_2(i+1)}$$

Relevance of the *i*th journal (*rel<sub>i</sub>*) is the relevance given by an expert to the journal at the *n*th position. We measured

the NDCG (normalized discounted cumulative gain), which ranges from zero to one, with one is the ideal ranking as shown below in equation. As recommendation lists vary in length, we used normalized discounted cumulative gain. Ideal discounted cumulative gain (IDCG<sub>n</sub>) is the maximum possible ideal discounted cumulative gain (DCG) at *n*th position.

$$NDCG_n = \frac{DCG_n}{IDCG_n}$$

a) *Recommendation-Evaluation Using LSA*: In the present work, hybrid approach mainly concerns to build relevant journal recommendation for researchers. However, evaluation of the proposed approach for its accuracy was conducted. Table 2 describes the comparison of results before and after clustering by using latent semantic analysis (LSA). However, to evaluate the quality of recommended journal list subjective evaluation under normalized discounted cumulative gain (NDCG) is applied as a ranking measure that is based on experts opinion.

TABLE II. EVALUATION OF RECOMMENDED JOURNALS USING LSA

Test Docs.	Evaluation Parameters						
	Without Clustering			With Clustering			
	P	R	F-I	P	R	F-I	NDCG
Abs <sub>1</sub>	0.298	0.5	0.37	0.498	0.6	0.54	0.57
Abs <sub>2</sub>	0.248	0.5	0.33	0.654	0.604	0.628	0.65
Abs <sub>3</sub>	0.34	0.6	0.43	0.454	0.536	0.491	0.53
Abs <sub>4</sub>	0.298	0.5	0.37	0.288	0.402	0.335	0.37

Table 2 shows the evaluation of four test documents (i.e. four abstracts of manuscripts) with and without clustering. A decision can be made about recommended scholarly journal list that a low percent 29.5 % of mean average precision (MAP) and 0.33 for f-measure is achieved before applying clustering. Moreover, clustering approach was used to analyze significant improvement in recommended scholarly journal list and it reduces the computational load. This integrated hybrid model showed improvement in results and is evaluated by the experts opinion as well under NDCG. A total number of 4 expert's opinion has taken as a relevance assigned by an expert for a scholarly journal at respective position.

b) *Recommendation-Evaluation Using Clustering And Cosine Similarity*: In the present section, clustering merges with cosine similarity to build the recommendations. Table 3 shows the improvement in recommended scholarly journals by using cosine similarity. Multiple performance measures are applied for the evaluation purpose under objective and subjective evaluation.

Table 3 shows the evaluation of four test documents by the performance benchmarks. These results are based on the recommended scholarly journals ranked list using cosine similarity algorithm. The precision of input document1 is 0.91. Similarly recall is 0.83 and f-measure is 0.868. The performance for the document2, document3, document4 is clearly shown in Table 3. Subjective evaluation results are also displayed under normalized discounted cumulative gain (NDCG) score by using expert's opinion. A decision can be

made about recommended scholarly journal list that a high percent 84.45% of mean average precision (MAP) and average of 0.79 for f-measure is achieved that shows the relevant scholarly journals for the respective documents. Documents refers to the abstract of the researcher's article. Similarly, for subjective evaluation using the expert's opinion a mean of 0.89 NDCG score can be seen for the recommended ranked list.

## V. CONCLUSION

The rapid increase in the number of publication venues represents a great challenge in the academic world. Researchers are needed to be aware of potential scholarly journals to publish their manuscripts. The current research deliberates a system for authors to identify scholarly journals that publish research results close to their field. The proposed system considers the researchers current field of research interest to generate relevant ranked journal list and it performs good to build recommendation for new researchers or for those who are interested to change their research domains. Additionally, our proposal is not publisher specific as we are managing multiple publishers.

TABLE III. EVALUATION OF RECOMMENDATION USING COSINE SIMILARITY

Test Documents	Evaluation Parameters			
	Precision(P)	Recall(R)	F-measure	NDCG
Abs <sub>1</sub>	0.91	0.83	0.868	0.93
Abs <sub>2</sub>	0.754	0.69	0.720	0.85
Abs <sub>3</sub>	0.844	0.70	0.765	0.88
Abs <sub>4</sub>	0.87	0.802	0.834	0.90

In this research, a hybrid model is introduced that merges with clustering and document similarity to build relevant recommendations. A significant increase perceived in the recommendation of scholarly journals by using aim and scope. Out of all the algorithms clustering combines with cosine similarity provides the good results with respect to multiple evaluation parameters recall, precision and f-measure. Overall accuracy by computing mean average precision (MAP) is 84.45% and average of 0.79 for f-measure is achieved. In future, the proposed hybrid model can be used for other disciplines i.e. Mathematics, Neuroscience, Medical etc. as well to recommend relevant scholarly journals for publication and it can be extended to work for other venue such as workshops, conferences. In this paper, we only consider aim and scope of the journals to build recommendation. In addition, we are planning to exploit other characteristics of scholarly journals as future work.

## REFERENCES

- [1] Achakulvisut, T., Acuna, D. E., Ruangrong, T., & Kording, K. (2016). Science Concierge: A fast content-based recommendation system for scientific publications. *PLoS One*, 11(7), e0158423.
- [2] Alhoori, H., & Furuta, R. (2017). Recommendation of scholarly venues based on dynamic user interests. *Journal of Informetrics*, 11(2), 553–563. <https://doi.org/https://doi.org/10.1016/j.joi.2017.03.006>
- [3] Aliannejadi, M., Mele, I., & Crestani, F. (2016). User model enrichment for venue recommendation. In *Asia Information Retrieval Symposium* (pp. 212–223).

- [4] Alshareef, A. M., Alhamid, M. F., & El Saddik, A. (2018). Recommending Scientific Collaboration Based on Topical, Authors and Venues Similarities. In 2018 IEEE International Conference on Information Reuse and Integration (IRI) (pp. 55–61).
- [5] Balduini, M., Bozzon, A., Valle, E. Della, Huang, Y., & Houben, G. (2014). Recommending Venues Using Continuous Predictive Social Media Analytics. *IEEE Internet Computing*, 18(5), 28–35. <https://doi.org/10.1109/MIC.2014.84>
- [6] Boukhris, I., & Ayachi, R. (2014). A novel personalized academic venue hybrid recommender. In 2014 IEEE 15th International Symposium on Computational Intelligence and Informatics (CINTI) (pp. 465–470). <https://doi.org/10.1109/CINTI.2014.7028720>
- [7] Chen, Z., Xia, F., Jiang, H., Liu, H., & Zhang, J. (2015). AVER: random walk based academic venue recommendation. In Proceedings of the 24th International Conference on World Wide Web (pp. 579–584).
- [8] Davis, N. L., Galliher, J. M., Spano, M. S., Main, D. S., Brannigan, M., & Pace, W. D. (2008). Evaluating conflicts of interest in research presented in CME venues. *Journal of Continuing Education in the Health Professions*, 28(4), 220–227.
- [9] Forrester, A., Bjrk, B.-C., & Tenopir, C. (n.d.). New web services that help authors choose journals. *Learned Publishing*, 30(4), 281–287. <https://doi.org/10.1002/leap.1112>
- [10] Kang, N., Doornenbal, M. A., & Schijvenaars, R. J. A. (2015). Elsevier Journal Finder: Recommending Journals for Your Paper. In Proceedings of the 9th ACM Conference on Recommender Systems (pp. 261–264). New York, NY, USA: ACM. <https://doi.org/10.1145/2792838.2799663>
- [11] Likis, F. E. (2018). Where Should You Publish? Choosing a Journal for Your Manuscript. *Journal of Midwifery & Women's Health*, 63(1), 7–9.
- [12] Luong, H., Huynh, T., Gauch, S., & Hoang, K. (2012). Exploiting social networks for publication venue recommendations, (May 2017).
- [13] Medvet, E., Bartoli, A., & Piccinin, G. (2014). Publication Venue Recommendation based on Paper Abstract, 1005–1011. <https://doi.org/10.1109/ICTAI.2014.152>
- [14] Pham, M. C., Cao, Y., Klamma, R., & Jarke, M. (2011). A clustering approach for collaborative filtering recommendation using social network analysis. *J. UCS*, 17(4), 583–604.
- [15] Tejada-Lorente, A., Porcel, C., Bernab-Moreno, J., & Herrera-Viedma, E. (2015). REFORE: A recommender system for researchers based on bibliometrics. *Applied Soft Computing*, 30, 778–791. <https://doi.org/https://doi.org/10.1016/j.asoc.2015.02.024>
- [16] Wang, X., Ounis, I., & Macdonald, C. (2019). Comparison of Sentiment Analysis and User Ratings in Venue Recommendation. In European Conference on Information Retrieval (pp. 215–228).
- [17] Yang, Z., & Davison, B. D. (2012). Venue Recommendation : Submitting your Paper with Style. <https://doi.org/10.1109/ICMLA.2012.127>
- [18] Yu, S., Liu, J., Yang, Z., Chen, Z., Jiang, H., Tolba, A., & Xia, F. (2018). PAVE: Personalized Academic Venue recommendation Exploiting co-publication networks. *Journal of Network and Computer Applications*, 104, 38–47.