

# Prescriptive Analytics for Publication Venue Recommendation

Tolulope Anthonia Adebayo  
Department of Information Systems

The federal University of Technology, Akure, Nigeria

Bolanle Adefowoke Ojokoh  
Department of Information Systems

The federal University of Technology, Akure, Nigeria

## ABSTRACT

Publication venue recommendation provide answers to one of the major challenges researchers face while seeking to get their research results or findings published in high-valued journals and conferences for easy dissemination and to maximize effects on future research. However, most recommendation systems available use traditional approaches which encounter problems such as cold start, data sparsity, among others. Hence, this study proposes a two-level recommendation model using prescriptive analytics technique (fuzzy logic) algorithm to infer decision on a suitable venue for publication based on key parameters such as the cost of publishing, impact factor of a journal or rank of a conference, and the average duration of review. Experiments were carried out on real world dataset obtained from Digital Bibliography and Library Project (DBLP) and Aminer Digital repositories. Results obtained from the evaluation of the system in terms of Accuracy@N, Precision, and F1-measure shows that the system performed efficiently and provides effective recommendation.

## General Terms

Recommender Systems, Soft computing, Fuzzy logic, publication venue.

## Keywords

Recommender Systems, Soft computing, Fuzzy logic, publication venue.

## 1. INTRODUCTION

The goal of every researcher is to make research results known to the target audience after much work has been put into writing. However, with new publishing venues coming up in recent times, it becomes a problem for researchers in deciding the most suitable venues for their work, while avoiding predatory journals. [1] stated that there are more than 2000 venues in Computer Science alone spanning through different fields, some of which are also interrelated. Consequently, recommender systems have emerged as a good solution for helping researchers deal with this rapid growth. Recommender systems as described by [2] suggest item(s) of interest out of overwhelmingly large options to users which in turn help in making decisions. Some application and service areas these systems have been employed are tourism, electronic-commerce, academics, books, movies, web pages, electronic-learning, news, music and so on. Examples of existing recommender systems are LIBRA (book recommender), Ringo (Music recommender), Profinder, GroupLens (News recommender), Netflix (Movie recommender), PickAFlick, EntreeC (Restaurant recommender) and LaboUr (Labor Management Recommender) [3]. Common recommender systems approach includes collaborative filtering, content-based filtering, and hybrid which are known as traditional approaches [4]. These approaches are limited by cold start problem, data sparsity, over-specialization among others.

However, research has shown the use of prescriptive analytics (PA) as an approach to building an effective recommender system. PA consists of methods such as probabilistic algorithms, machine learning, deep learning, mathematical programming, evolutionary computation, simulation, and logic-based models [5]. The approach broadens the potential of descriptive and predictive analytics by enabling data-driven optimization for decision support and planning [6]. In terms of recommendation, PA uses statistical computations to analyse collection of data containing previous recommendations and give prediction regarding the future based on the learned patterns. This process eliminates the human intervention of rating recommended items, which in turn, resolves problems encountered in traditional recommender systems.

In this work, we apply a prescriptive analytics technique (fuzzy logic) to provide recommendations for researchers in getting suitable venues to submit their manuscripts for publication. The rule-based algorithm uses the profile of users and the research area of the manuscript in providing a first level recommendation, while the fuzzy logic uses the cost of publication, impact factor and the average duration of review to refine and personalize the recommendations. The proposed system is aimed at providing suitable venues for both old and new researchers using the venues of their co-authors and co-affiliated researchers that match the research area of the manuscript and, the researcher's preferred cost of publishing, impact factor or rank of conference and the average duration of review.

The remaining part of this paper is structured as follows: Section 2 presents the review of related work. Section 3 presents the proposed system architecture, and the description of the components that make up the architecture. Section 4 presents the experimental setup and results while the conclusion of the paper and future works is presented in Section 5.

## 2. RELATED WORKS

The speedy growth and complexity of information over the internet in recent times has made recommender systems more and more needed in the society. A recommender system according to [7] is one that suggests items that might be of interest to users when faced with different similar choices. These systems using different techniques and approaches have been successfully implemented in areas such as movies, books, music, and so on [8]. In academic recommendation, several works have been proposed comprising of paper recommendation, academic venue recommendation, reviewers' recommendation, citations' recommendation, collaborators recommendation and Conference Sessions recommendations [9]. [1] proposed a user-based collaborative filtering approach combining the topic and writing-style of papers for venue recommendation. They argued that the writing-style and format of a paper may be used as collaborative filtering features.

However, the accuracy reported for the work was low. A social network-based approach proposed by [10] used the author's network publication history to generate relevant venues. However, it fails to consider new researchers with no publication records in the research domain. A hybrid recommender system for recommending upcoming conferences using the venues from co-authors, co-citers and co-affiliated scholars of target researchers was proposed by [11], the system does not however provide alternative venues apart from the ones that exactly matches the requirements given by the user. [12] proposed Cavnar-Trenkle, Latent Dirichlet Allocation and Latent Dirichlet Allocation + Clustering methods in recommending venues using the topic and abstract of papers, but the system may suffer from text ambiguities. A system was developed for recommending Elsevier journals by [13] using a natural language processing for the feature generation and Okapi BM25 matching algorithm for recommendation. The scope of the work is however limited to Elsevier journals. [9] proposed a system that combined user-based collaborative filtering (CF), item-based CF, a stochastic gradient descent (SGD), and singular value decomposition (SVD) algorithms based on the recent personal article pools and study of users, it however did not consider researchers with no prior publications. [14] proposed a publication recommender system (Pubmender) to suggest suitable PubMed journals based on a paper's abstract. Pretrained word2vec was first used to construct the start-up feature space, a deep convolutional neural network was constructed to achieve a high-level representation of abstracts, and a fully connected softmax model was adopted to recommend the best journals. [15] developed a system that recommends suitable journals or conferences with a priority order based on the abstract of a manuscript, employing a web crawler to continuously update the training set and the learning model and a softmax regression model to provide three class-recommendation. An academic venue recommender system based on a deep learning integrated framework that contains bi-directional LSTM and hierarchical attention network was proposed by [16]. The system uses a combination of features such as paper abstract, title, keywords, field of study and publication records of authors. In this work, we propose a fuzzy logic-based approach in providing relevant recommendations for manuscripts. From the related works reviewed so far, we have not come across any research that proposed a system for recommending venues for researchers using a fuzzy logic approach to provide personalized recommendations that fit researchers' interests and requirements using the cost of publishing, impact factor and the average duration of review.

### 3. PROPOSED SYSTEM

The proposed system is composed of extraction and refinement phases (see Fig 1). The architecture requires target authors' name, email address, institution, and the research area of the

manuscript to be provided via the user interface. The system then searches through the article database to extract the venues of manuscripts whose authors belong to the same institution/department with the target author, possibly the venues where the target author has published in the past, venues where the co-authors of the target author (if any) has published in the past and the venues of manuscripts with the same research area as the new manuscript. The extracted venues are combined to provide the first level of recommendations after which it is passed on to the fuzzy logic component for refinement. To obtain more personalized recommendations, the second phase refines the first set of recommendations provided according to a preferred cost of publishing, impact factor of a journal or rank of a conference and the average duration of review selected by the target author. For instance, a target author might want a free journal, with the cost of '0' with an impact factor of between 2.5-3.0 and average duration of review of 12 weeks. The final recommendations are displayed to the target author via the user interface. The methods used in providing these recommendations are discussed in the remaining parts of this section.

#### 3.1 Extraction Phase

A rule-based approach was used to obtain the venues of co-authors and co-affiliation of target authors, with the impression that people who publish together and belong in the same institution with the target author will tend to publish in the same or related venues. Thus, the process involved in providing the first level of recommendation is discussed in this section.

Given the dataset  $X$  of articles as follows:

$$X = \{ p_1, p_2, p_3, \dots, p_n \}; p_i \in \{ (e_i, r_i, u_i, v_i) \} \quad (1)$$

where  $p_i$  represents an article in  $X$ ,  $n$  represents the total number of articles in  $X$ ,  $e_i$  is the email address,  $r_i$  is the research area,  $u_i$  represents the researcher's institution, and  $v_i$  is the publication venue. Given also that an author  $k$  with a target manuscript  $p_k$  registered with the system by providing email address  $e_k$ , research area  $r_k$ , and institution  $u_k$ , the system searches through the dataset  $X$  and match  $e_k$ ,  $r_k$  and  $u_k$ , of the author with the collection of articles constituting the dataset using rules as follows:

For paper,  $p_i$  in  $X$  and the list of venues  $V$

```

if  $u_k = u_i(p_i)$ 
  if  $r_k = r_i(p_i)$ 
    if  $e_k = e_i(p_i)$ 
      extract venues  $v_i$  into  $V$ 
    endif
  endif
endif

```

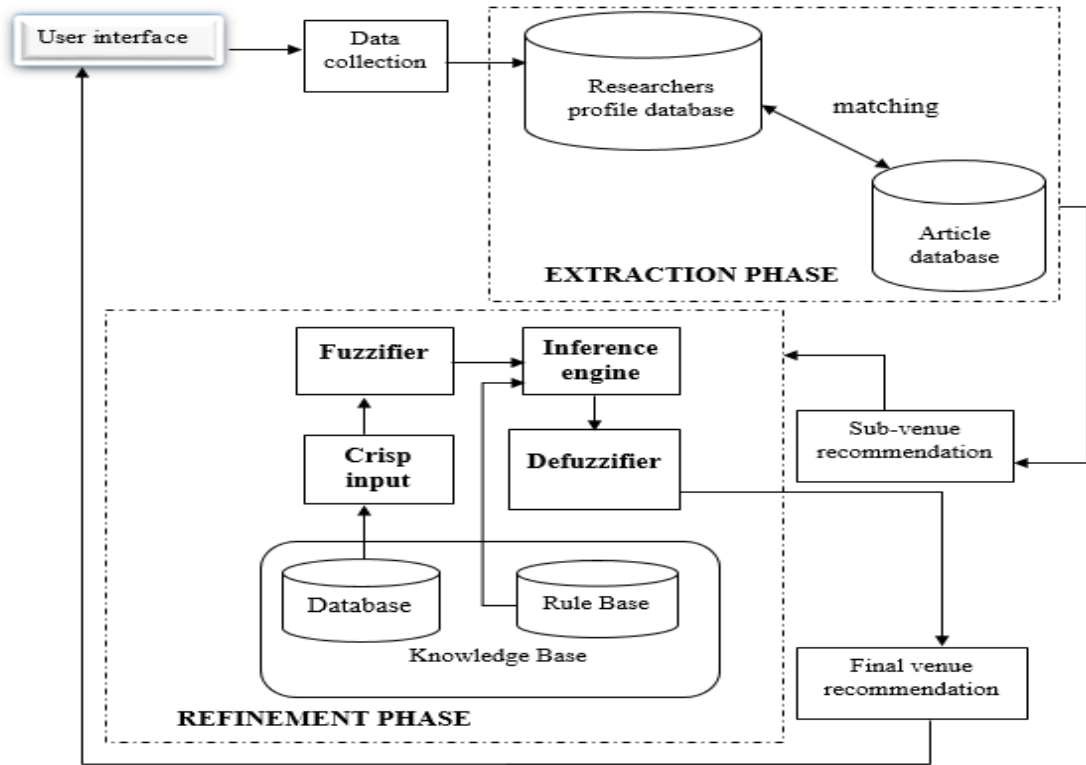


Figure 1: System Architecture

The list of  $V$  obtained from this phase is then passed on to the next phase (fuzzy logic) for refinement.

### 3.2 Refinement Phase

This phase used a fuzzy logic approach to polish the venues obtained from the first phase and the components of the fuzzy logic system are discussed below:

#### 3.2.1 Fuzzy Logic System

The strength of Fuzzy logic lies in providing precise answers to problems that involve the handling of numerous variables [8]. It is adopted in this work to provide more personalized recommendations. The Knowledge Base of the fuzzy logic component consists of the Database and Rule Base. The database stores the venues' cost of publishing, impact factor or rank and average duration of review while the Rule Base consists of a set of IF-THEN statements that help in making decisions. The Fuzzy logic component accepts as input the values of the cost of publishing, impact factor or rank and the average duration of review provided by authors and applies already defined actions to them. Lastly, inputs from the rule base and fuzzification interface is accepted by the Fuzzy Inference Engine, thus applying a pre-defined procedure to produce recommendations. The processes involved are discussed in the following sub-sections:

##### a. Fuzzification

The input values (cost, impact factor, rank, and average duration of review) were fuzzified into fuzzy sets using linguistic values, linguistic variables, and membership functions. The Triangular membership function was adopted in this work because of its generality and simplicity. A fuzzy set ' $K$ ' in the universe of discourse  $U$  as shown in equation (2) contains a set of all input variables pairs arranged in order, and its element represented by  $x_j$ , where  $x_j$  is the input values,  $j = 1, \dots, n$ . The input values were converted using the triangular membership function to define the degree of each input

variables.  $\mu_K(x_j)$  is the membership function which is derived from equation (3).

$$K = \{(x_j, \mu_K(x_j)) \mid x_j \in U, \mu_K(x_j) \in [0,1]\} \quad (2)$$

$$\mu_K(x_j) = \begin{cases} 1 & \text{if } x_j < a \\ \frac{x_j - a}{b - a} & \text{if } a \leq x_j < b \\ \frac{c - x_j}{c - b} & \text{if } b \leq x_j < c \\ 0 & \text{if } c \leq x_j \end{cases} \quad (3)$$

where  $\mu_K(x_j)$  is the membership function of  $x_j$  in  $K$  and  $\mu_K$  is the degree of membership of  $x_j$  in  $K$ , whereas  $a, b, c$  are the parameters of the membership function making up its triangular form. Each of these attributes are described by Linguistic terms (Nil, very low, low, medium, high, very high).

##### b. Rule-base

The fuzzy logic system's rule base is made up of a set of IF-THEN rules in which the IF parts (antecedents) and the THEN parts (consequents) include linguistic variables. A rule is considered fired if any of its precedence parameter such as *very low, low, medium, high, and very high* results to true, else it does not fire.

An instance of the design of the rules is as shown:

IF (cost is *verylow*) AND (impactfactor is *verylow*) AND (review is *verylong*) THEN (class is *class B*)

##### c. Inference Engine

To draw conclusions from a rule base, the fuzzy inference engine uses the inference mechanism on the set of rules in the fuzzy rule base to return a fuzzy output set from a pool of IF-

*THEN* rules. This encompasses matching the input fuzzy set with the basis of the rules, activation of the rules to infer the conclusion of each rule that is fired, and the combination of all conclusions using fuzzy set union to create fuzzy set output. This mapping is the base from which decisions can be made. Fuzzy rule sets typically have quite a lot of antecedents that are joined using fuzzy logical operators, such as AND, NOT and OR. The minimum weight of all the antecedents is used by the AND operator, while operator OR uses maximum value. In this work, we used the aggregator operator AND to calculate the firing power or degree of truth of a rule, by calculating the non-zero minimum values of all the antecedents. The inference engine estimates all the rules in the rule base and converts the consequents of all the fired rules that are weighted into a single fuzzy set using Root Sum Square (RSS) inferential technique given in equation (4).

$$RSS = \sqrt{\sum_{x=1}^n (R_x^2)} \quad (4)$$

where  $R_x$  represent different rules which share the same conclusion,  $X = 1, 2, 3, \dots, n$ . The Mamdani's fuzzy inference system was used for this work.

#### d. Defuzzification

The defuzzification process carried out by the defuzzifier transforms the output of the inference engine into crisp values which is usually vital for accurate analysis and explanation. The defuzzifier accepts as input, the output of the inference engine which is a fuzzy set and uses equation (5) to obtain the defuzzified (crisp) output. There are various methods of defuzzification, but this work employs the centre of gravity technique (CoG) because it is accurate and computationally simple.

$$CoG(Y') = \frac{\sum_{i=1}^n \mu_Y(x_j) x_j}{\sum_{i=1}^n \mu_Y(x_j)} \quad (5)$$

where  $CoG(Y')$  is the crisp output,  $\mu_Y(x_j)$  is the summed-up membership value of  $x_j$  and  $x_j$  is the midpoint of membership function of the output variable.

## 4. EXPERIMENT AND RESULTS

### 4.1 Experimental Set up

Experiments were conducted to verify the system performance and to determine how useful and precise the recommendations provided were. DBLP citation dataset obtained from (<http://arnetminer.org/DBLP> Citation) was used to conduct the experiment. An average of 1200 research papers consisting of 52 conferences and 30 journals from the dataset was used with the addition of the research area, the cost of publishing, impact factor, rank of venues and the average duration of review, which was obtained from

([https://www.academia.edu/18717832/IEEE\\_Transactions\\_and\\_Journals\\_List\\_Review\\_Speed\\_Impact\\_Factors\\_and\\_Open\\_Access\\_Fee](https://www.academia.edu/18717832/IEEE_Transactions_and_Journals_List_Review_Speed_Impact_Factors_and_Open_Access_Fee)). Information about computer science venue ranking was obtained from (CORE 2008, ERA 2010, CORE 2013), which are mined from the Computing Research and Education (CORE) Conference Portal.

### 4.2 Dataset Description

The dataset contains 10 attributes namely: Title, Author(s), Email, Institution, Year of Publication, Venue, Research Area, Impact Factor, Rank, Cost and Average Review duration. The attributes Paper Title, Author(s) and year was however not used in the recommendation process. A sample of the dataset is shown in Fig2.

### 4.3 Evaluation Metrics

The recommendations were assessed with the standard metric venue-accuracy@N, which is defined as the number of correct recommendations divided by the number of all recommendations. We assume that the ground-truth venue where paper 'p' was published is  $Ve_p$ , then, if  $Ve_p$  is among the N venues recommended by the system, we say the recommendation for the paper 'p' is correct, where (N represents 1, 2 and 3). To prove the efficacy of the refinement process, the system was given to 20 researchers to indicate their contentment vis-à-vis the recommendation results and the general performance of the system. Responses were obtained using a questionnaire to define their requirements about the cost of publishing, impact factor/rank and the average duration of review; this was used to calculate the accuracy, precision, recall and F1-measure of the system. The metrics used are described below:

$$Accuracy = \frac{A+D}{A+B+C+D} \quad (6)$$

$$Precision = \frac{A}{A+C} \quad (7)$$

$$Recall = \frac{A}{A+B} \quad (8)$$

$$F1 - measure = \frac{2 * Precision * Recall}{Precision + Recall} \quad (9)$$

where  $A$  is the number of venues that satisfy a certain user out of those recommended by the system,  $B$  is the number of venues that satisfy a certain user, but not recommended by the system,  $C$  is the number of venues that do not satisfy a certain user though recommended by the system, and  $D$  is the number of venues that do not satisfy a certain user and not recommended by the system.

### 4.4 Experimental Results

Using a dataset consisting of 1200 papers obtained from Aminer, we calculated the accuracy @N for the venues at Top 1, Top 2 and Top 3 positions. Fig3 show a sample of the recommendation result.

| S/N | Paper Title  | Author(s)   | Email  | Institution  | Year | Venue  | Research Area       | Impact Factor/Rank | cost | Review Duration |
|-----|--|---|--|--|------|--|---------------------|--------------------|------|-----------------|
| 1   | The influence of global constraints on similarity measures for time-series databases | Vladimir Kurbalija; Miloš Radovanović; Zoltan Geler; Mirjana Ivanović | <a href="mailto:kurb@dm.uns.ac.rs">kurb@dm.uns.ac.rs</a> ; <a href="mailto:radacha@dm.uns.ac.rs">radacha@dm.uns.ac.rs</a> ; <a href="mailto:mira@dm.uns.ac.rs">mira@dm.uns.ac.rs</a>             | Department of Mathematics and Informatics, Faculty of Sciences, University of Novi Sad, Trg D. Obradovića 4, 21000 Novi Sad, Serbia  | 2014 | Knowledge-based Systems  | Databases           | 4.529              | 2400 | 28weeks         |
| 2   | How does high dimensionality affect collaborative filtering?                         | Alexandros Nanopoulos; Miloš Radovanović; Mirjana Ivanović            | <a href="mailto:nanopoulos@ismll.de">nanopoulos@ismll.de</a> ; <a href="mailto:radacha@dm.uns.ac.rs">radacha@dm.uns.ac.rs</a> ; <a href="mailto:mira@dm.uns.ac.rs">mira@dm.uns.ac.rs</a>         | University of Hildesheim, Hildesheim, Germany; University of Novi Sad, Novi Sad, Serbia  | 2009 | Proceedings of the third ACM conference on Recommender systems                 | recommender systems | Rank B             |      |                 |
| 3   | Proving Secrecy is Easy Enough   | Véronique Cortier; Jon Millen; Harald Rueß                            | <a href="mailto:Cortier@lsv.ens-cachan.fr">Cortier@lsv.ens-cachan.fr</a> ; <a href="mailto:millen@csl.sri.com">millen@csl.sri.com</a> ; <a href="mailto:ruess@csl.sri.com">ruess@csl.sri.com</a> | Laboratoire SpCification et VCrification, Ecole Normale Sup.Crietre de Cachan, 61, Avenue du President Wilson, 94230 Cachan, France; SRI International, Computer Science Laboratory, 353 Ravenswood Ave, Menlo Park, CA 94035, USA | 2001 | CSFW'01 Proceedings of the 14th IEEE workshop on Computer Security Foundations | computer security   | Rank A             |      |                 |
| 4   | Telecommunications network developments  | Robyn Bray  | <a href="mailto:robynbray@hotmail.com">robynbray@hotmail.com</a>   | University of Glasgow  | 1986 | Computer Communications  | networking          | 2.613              | 2400 | 16weeks         |
| 5   | Steganographic communication with quantum information                                | Keya Martin   | <a href="mailto:kmartin@itd.nrl.navy.mil">kmartin@itd.nrl.navy.mil</a>   | Center for High Assurance Computer Systems, Naval Research Laboratory, Washington, DC  | 2007 | IHF07 Proceedings of the 9th international conference on Information hiding    | cryptograp hy       | Rank C             |      |                 |

Figure2. Dataset Sample

| Paper Title  | Research Area     | Top 3 Recommended Venues  |
|--|-------------------|---|
| The influence of global constraints on similarity measures for time-series databases   | Databases         | <ol style="list-style-type: none"> <li>The VLDB Journal — The International Journal on Very Large Data Bases</li> <li><i>Knowledge-Based Systems</i></li> <li>PKDD Proceedings of the European conference on Principles and Practice of Knowledge Discovery in Databases</li> </ol>   |
| Telecommunications network developments  | Networking        | <ol style="list-style-type: none"> <li><i>Computer Communications</i></li> <li>SIAM Journal on Computing</li> <li>Networks</li> </ol>   |
| Turning privacy leaks into floods: surreptitious discovery of social network friendships and other sensitive binary attribute vectors. | Computer Security | <ol style="list-style-type: none"> <li>POST Proceedings of the international conference on Principles of Security and Trust</li> <li>NTMS Proceedings of the international conference on New technologies, mobility and security</li> <li><i>Proceedings of the annual ACM workshop on Privacy in the electronic society</i></li> </ol> |

Figure 3. Sample of Result

Table 2: Results showing Accuracy@N

|                          | Top1   | Top2   | Top3   |
|--------------------------|--------|--------|--------|
| Number of correct venues | 950    | 790    | 860    |
| Accuracy                 | 79.17% | 65.83% | 71.67% |

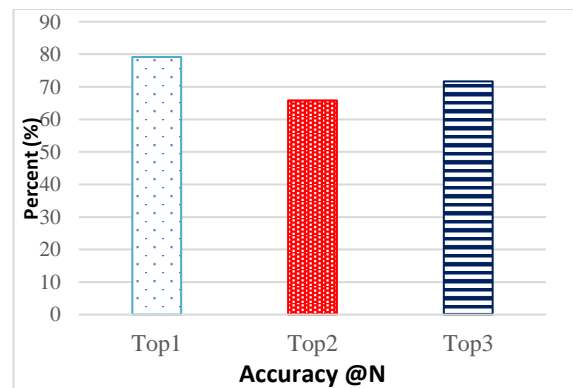


Figure 4: Evaluation of the system for Accuracy @ N

Results of Users' evaluation are displayed in Tables 5 and 6. The Accuracy, Precision, Recall and F1 values are presented in Table 5 while Table 6 show results in percentage. The average Accuracy, Precision, Recall and F1-measure of the proposed system are 96.28%, 84.73%, 70.78% and 73.13% respectively.

**Table 5: The Evaluation of 20 Users**

| User ID | A | B | C | D  |
|---------|---|---|---|----|
| 1       | 2 | 3 | 1 | 76 |
| 2       | 5 | 2 | 0 | 75 |
| 3       | 6 | 2 | 2 | 72 |
| 4       | 6 | 2 | 1 | 73 |
| 5       | 4 | 3 | 2 | 73 |
| 6       | 4 | 2 | 1 | 75 |
| 7       | 4 | 3 | 2 | 73 |
| 8       | 5 | 3 | 1 | 73 |
| 9       | 5 | 2 | 1 | 74 |
| 10      | 6 | 1 | 0 | 75 |
| 11      | 3 | 0 | 0 | 79 |
| 12      | 5 | 2 | 1 | 74 |
| 13      | 4 | 2 | 1 | 75 |
| 14      | 5 | 4 | 1 | 72 |
| 15      | 6 | 4 | 2 | 70 |
| 16      | 6 | 3 | 1 | 72 |
| 17      | 5 | 1 | 0 | 76 |
| 18      | 3 | 0 | 0 | 79 |
| 19      | 4 | 0 | 1 | 77 |
| 20      | 4 | 4 | 0 | 74 |

**Table 6: Accuracy, precision, recall and F1-measure of the system (%)**

| User ID | Accuracy | Precision | Recall | F1-Measure |
|---------|----------|-----------|--------|------------|
| 1       | 95.12    | 66.67     | 40.0   | 50.0       |
| 2       | 97.56    | 100.0     | 71.43  | 83.33      |
| 3       | 95.12    | 75.0      | 75.0   | 75.0       |
| 4       | 96.34    | 85.71     | 75.0   | 79.99      |
| 5       | 93.9     | 66.67     | 57.14  | 61.54      |
| 6       | 96.34    | 80.0      | 66.67  | 72.73      |
| 7       | 93.9     | 66.67     | 57.14  | 61.54      |
| 8       | 95.12    | 83.33     | 62.50  | 71.43      |
| 9       | 96.34    | 83.33     | 71.43  | 76.92      |
| 10      | 98.78    | 100.0     | 85.71  | 92.31      |
| 11      | 100.0    | 100.0     | 100.0  | 100.0      |
| 12      | 96.34    | 83.33     | 71.43  | 76.92      |

|            |              |              |              |              |
|------------|--------------|--------------|--------------|--------------|
| 13         | 96.34        | 80.0         | 66.67        | 72.73        |
| 14         | 93.9         | 83.33        | 55.55        | 66.66        |
| 15         | 92.68        | 75.0         | 60.0         | 66.66        |
| 16         | 95.12        | 85.71        | 66.67        | 74.99        |
| 17         | 98.78        | 100.0        | 83.33        | 90.91        |
| 18         | 100.0        | 100.0        | 100.0        | 100.0        |
| 19         | 98.78        | 80.0         | 100.0        | 88.88        |
| 20         | 95.12        | 100.0        | 50.0         | 66.67        |
| <b>Avg</b> | <b>96.28</b> | <b>84.73</b> | <b>70.78</b> | <b>73.13</b> |

## 4.5 DISCUSSION

Results obtained from the Evaluation of the system in terms of Accuracy@ Top1, Top2 and Top 3 positions as shown in Table 2 show that the highest performance is at Top1 and the least is at Top 2, which means that the number of times the venue where a particular paper was published was predicted correctly mostly at the Top 1 position. Although, as shown in Table 3, when the recommendations appear at the Top 2, Top 3 positions, the venues recommended at these positions are still suitable venues to which the paper could have been submitted considering the target author's preference in terms of cost, impact factor and the average duration of review. From the comparison with other existing works that used this evaluation metric, results show that Accuracy@1 for this work is higher than the existing works while Accuracy @2 and Accuracy@3 is lower in one of the works compared with and this could be due to the number of articles in the dataset used in the work. The proposed system used a dataset consisting of 1200 papers as against 154 used in the other two, and this can be seen to be significantly lower than the number of articles used in this work. Also, the results obtained for the Precision, Recall, Accuracy and F1-Measure show that the users were satisfied with recommendation results, and this in turn demonstrate the impact of the refinement process.

**Table 4 Comparative Analysis of the proposed system with existing works**

| Authors              | Approach       | Top1    | Top2    | Top3    |
|----------------------|----------------|---------|---------|---------|
| Luong <i>et al.</i>  | Content-based  | 44.38 % | 65.63 % | 80.63 % |
|                      | Network-based  | 56.49 % | 79.87 % | 91.56 % |
| Medvet <i>et al.</i> | Cavnar-Trenkle | -       | -       | 26.8%   |
|                      | Two-step-LDA   | -       | -       | 3.4%    |
|                      | LDA+clustering | -       | -       | 16.1%   |
| Current Research     | Fuzzy Logic    | 79.17 % | 65.83 % | 71.67 % |



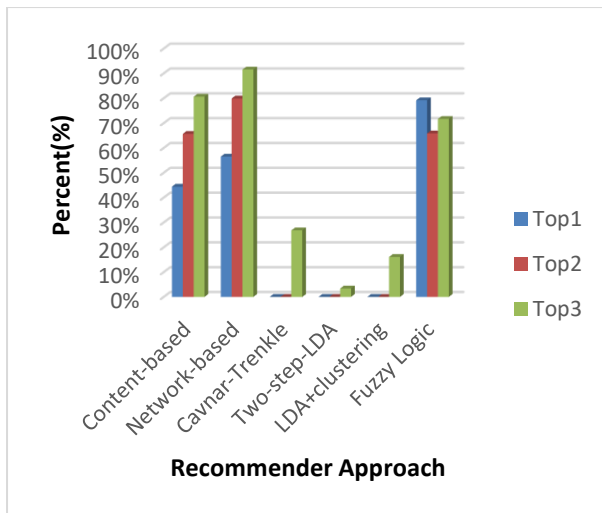


Fig3: Comparison with other works

## 5. CONCLUSION

Collaborative filtering, content-based, social network based and other techniques are widely used especially in academics, for recommending research papers, citations, publication venues and so on. Existing systems in this field are still not personalized enough and suffers from cold-start problem and data sparsity among others. Hence, in this work, a prescriptive analytics technique (fuzzy logic) is adopted to better give recommendations that suit the choice of researchers. The system uses a rule base matching technique to extract the venues of the co-authors and co-affiliated researchers of the target researcher to provide recommendations for the target researcher, which comes from the idea that researchers who publish together or belong to the same institution tend to have similar interests. The fuzzy logic component then refines the recommendations using the preferred cost of publishing, the impact factor of a journal or rank of a conference and the average duration of review, chosen by the target author. The system in turn provides recommendations based on these choices. Experimental tests were conducted using a dataset consisting of 1200 research papers to obtain the accuracy @ Top (1, 2 and 3) positions and the results show that the system performs efficiently. Future works could incorporate deep learning techniques on a larger dataset in providing recommendations which could provide a better performance and higher accuracy. Also, more factors could be considered for personalizing the recommendations provided such as h-index of journals, location, date and venue of conferences, acceptance or rejection rates, scope and guidelines of journals and journal audience (local or international). The system could be made to run on mobile platforms as well as incorporate other disciplines apart from computer science.

## 6. REFERENCES

- [1] Yang, Z. and Davison, B. (2012). Venue recommendation: Submitting your paper with style. In 11th *International Conference on Machine Learning and Applications*, 1, 681-686.
- [2] Cai, H. (2016). Publication venue recommendation in heterogeneous information networks (Doctoral dissertation). University of Illinois at Urbana-Champaign, Urbana, Illinois.
- [3] Isinkaye, F., Folajimi, Y. and Ojokoh B. (2015). Recommendation systems: Principles, Methods and Evaluation. *Egyptian Informatics Journal*, 16, 261-273.
- [4] Bai, X., Wang, M., Lee, I., Yang, Z., Kong, X. and Xia, F. (2019). Scientific Paper Recommendation: A Survey. *IEEE Access*, 7, 9324-9339.
- [5] Lepenioti, K., Bousdekis, A., Apostolou, D. and Mentzas, G. (2020). Prescriptive analytics: Literature review and research challenges. *International Journal of Information Management*, 50, 57-70.
- [6] Frazzetto, D., Nielsen, T., Pedersen, T. and Siksnys, L. (2019). Prescriptive Analytics: A Survey of Emerging Trends and Technologies. *VLDB Journal*, 28(4), 575-595.
- [7] Fayyaz, Z., Ebrahimian, M., Nawara, D., Ibrahim, A. and Kashef, R. (2020). Recommendation systems: Algorithms, challenges, metrics, and business opportunities. *applied sciences*, 10(21), 7748.
- [8] Ojokoh, B., Omisore, M., Samuel, O. and Ogunniyi, T. (2012). A fuzzy logic based personalized recommender system. *International Journal of Computer Science and Information Technology & Security (IJCSITS)*, 2(5), 1008-1015.
- [9] Alhoori, H. and Furuta, R. (2017). Recommendation of Scholarly Venues Based on Dynamic User Interests. *Journal of Informetrics*, 11(2), 553-563.
- [10] Luong, H., Huynh, T., Gauch, S., Do P. and Hoang, K. (2012). Publication venue recommendation using author network's publication history. In *Proceedings on the 4th Asian Conference on Intelligent Information and Database Systems*, 3, 426-435.
- [11] Boukhris, I. and Ayachi, R. (2014). A novel personalized academic venue hybrid recommender. In *IEEE 15th International Symposium on Computational Intelligence and Informatics*, (pp. 465-470).
- [12] Medvet, E., Bartoli, A. and Piccinin, G. (2014). Publication venue recommendation based on paper abstract. In: *Proceedings of the IEEE 26th International Conference on Tools with Artificial Intelligence*, (pp. 10-12).
- [13] Kang, N., Doornenbal, M. and Schijvenaars, B. (2015). Elsevier Journal Finder: Recommending Journals for your Paper. In *Proceedings of the 9th ACM Conference on Recommender Systems* (pp. 261-264).
- [14] Feng, X., Zhang, H., Ren, Y., Shang, P., Zhu, Y., Liang, Y., Guan, R. and Xu, D. (2019). The Deep Learning-Based Recommender System "Pubmender" for Choosing a Biomedical Publication Venue: Development and Validation Study. *International Journal of Medical Internet Research*, 21(5), e12957.
- [15] Wang, D., Liang, Y., Xu, D., Feng, X. and Guan, R. (2018). A content-based recommender system for computer science publications. *Journal of Knowledge-Based Systems*, 157, 1-9. Pradhan, T., Gupta, A. and Pal, S. (2020). HASVRec: A modularized Hierarchical Attention-based Scholarly Venue Recommender system. *Knowledge-Based Systems*, 204, 106181