

A Novel Fuzzy Rough Sets Theory Based CF Recommendation System

C. Raja Kumar¹ and VE. Jayanthi²

¹Assistant Professor, Department of IT, ULTRA College of Engineering & Technology for Women, Madurai

²Professor, Department of Electronics and Communication Engineering, PSNA College of Engineering and Technology, Dindigul

E-mail: rkumardgl@gmail.com, jayanthi.ramu@gmail.com

Collaborative Filtering (CF) is one of the popular methodology in recommender systems. It suffers from the data sparsity problem, recommendation inaccuracy and big-error in predictions. In this paper, the efficient advisory tool is implemented for the younger generation to choose their right career based on their knowledge. It acquires the notions of indiscernible relation from Fuzzy Rough Sets Theory (FRST) and propose a novel algorithm named as Fuzzy Rough Set Theory Based Collaborative Filtering Algorithm (FRSTBCF). To evaluate the model, data is prepared using the cross validation method. Based on that, ratings are evaluated by calculating the MAE (mean average error), MSE (means squared error) and RMSE (root means squared error) values. Further the correctness of the model is measured by finding rates like Accuracy, Specificity, Sensitivity, Precision & False Positive Rate. The proposed FRSTBCF algorithm is compared with the traditional algorithms experiment results such as Item Based Collaborative Filtering using the cosine similarity (IBCF-COS), IBCF using the pearson correlation (IBCF-COR), IBCF using the Jaccard similarity (IBCF-JAC) and Singular Value Decomposition approximation (SVD). The proposed algorithm gives better error rate and its precision value is comparatively identical with the existing system.

Keywords: Recommendation System; Collaborative Filtering; Fuzzy Rough Sets Theory; Indiscernibility; prediction

1. INTRODUCTION

Recommendation system (RS) is a stimulating field in this Big data epoch. The CF is an approach widely used in the recommendation systems [1]. It is used to provide the most compatible and suitable suggestions in a decision making process. For example, when a student wants to get an assistance to choose his right career, the RS predict and suggest a career path related courses for further study. That suggestion would be based on his academic performance as well as his co-curricular activities. This kind of a knowledge pedestal which recommends a career path is a boon for the multi disciplinary young generation. As a rule of thumb, the set of functional and technical competence describes the student's caliber. From that, the CF methods executes its thought process for a best recommendation.

A wide variety of industries and businesses needs an application which is capable of providing cogent recommendations. Such kind of recommendation application helps the end users to find the necessary information that they might not have thought of. Some of the domains which require an instance recommen-

ation system are listed below.

- Banking, Financial services and Insurance (BFSI)
- Public Broadcasting
- Online disclosers like publications
- Retail merchandise

User's traditional patterns and consumption patterns are the base to recommend new items to the user. This type of recommendation system is called as collaborative filtering system [1], [2] and [3] which gives a suggestion to the users based on their past behavior and similar users. A number of setbacks like data sparsity, recommendation inaccuracy, and big error in predictions may reduce the users trust in the conventional collaborative filtering system.

To address the big error in prediction issue, the FRST concept indiscernibility is exploited in the proposed model. FRST described in [4], [5], [6] and [7] are used for approximation, feature selection, rule extraction and many more purpose. FRST is derived from the classical Rough Sets model is proposed [8]. The

earlier is based on fuzzy relation the latter is based on Boolean equivalence relation.

This paper is organized as follows. Section 2 summarizes the background and related work. Section 3, introduces the Fuzzy Rough Sets Theory based CF recommendation in detail. In section 4, the proposed Fuzzy Rough Sets Theory based CF recommendation method is evaluated using the stakeholder's functional and technical competence level data set and it is compared with the legacy recommendation methods. The paper is concluded in the section 5.

2. BACKGROUND AND RELATED WORK

2.1 Recommender Systems Basic Concepts

There have been many studies on recommender systems and most of these studies propose a system which recommend an item to the users [2] [3]. The recommender systems are used to discover the items which would be significant for the users. The significant items can be of any type, such as movies, jokes, restaurants, books, news articles and so on. The recommendation methods are broadly classified into collaborative filtering (CF), content based (CB) and hybrid methods [2]. Here the focus is to propose a novel CF method based on FRST to eliminate the big error in prediction.

2.2 Fuzzy Rough Sets Theory Basic Concepts

Fuzzy Rough sets theory (FRST) proposed in [4], [5], [6] and [7] is an extended mathematical tool to represent knowledge. It is used to analysis the vague description of the objects. It influences the uncertain knowledge from the given data set, even without knowing any additional details about the objects. It is a popular scaffold for the applications based on pattern recognition, feature selection, rule induction etc., the fuzzy indiscernible relation is used to determine the similarity degree between two objects. FRST analysis work starts from the information tables. The information tables contain data about objects of interest. The knowledge of these objects is described by their attributes and attributes value. It is often interesting to discover similarity relationship.

Like in the classical rough sets theory, the information system in FRST is defined as $IS = (U, A)$ where U is the universe contains a finite non-empty set of objects, $U = \{O_1, O_2, O_3, \dots, O_n\}$ and A is the non-empty finite set of attributes. Each attribute $a \in A$ is an information function: $f_a: U \rightarrow V_a$ where V_a is the set of allowed values of a . Information systems may include a decision attribute, which contains the decision values of each object.

The formal representation of such information system is, $IS = (U, A \cup \{d\})$ where $d \notin A$ is the decision attribute. It analyzes the information systems that involve uncertainty and indiscernibility. Indiscernibility arises, when the differences between the objects are not able to identify based on their attribute values. FRST express the indiscernibility relation in many different ways. The approaches like fuzzy tolerance, equivalence and T-equivalence relations have been implemented to determine

the indiscernibility relation.

3. FUZZY ROUGHSETS THEORY BASED CF RECOMMENDATION

In this section, the proposed FRST based CF approach is introduced. In which the nearest competences are found based on rated competences of stakeholders. We first introduce the formal definitions of concepts in Section 3.1. The real mechanism and its technical details described in Sections 3.2

3.1 Preliminaries

Assume that in a FRST based CF recommender system, there are a set S of stakeholders and a set C of competence. To identify the similarity between those competences, the FRST indiscernibility relation is calculated. It determines the degree to which those competences are indiscernibility. The indiscernibility relation is the basic concept of rough sets theory which is extended to a fuzzy indiscernibility relation [8].

In literature, the methods like fuzzy tolerance [9], equivalence, T-equivalence (or) T-similarity [10] are commonly approached to find the fuzzy indiscernibility relations. In this paper, the T-equivalence method is considered to find the fuzzy indiscernibility relations. A traditional similarity-based learning method [1] uses the distance similarity matrix to differentiate the competences.

3.2 Mechanism

To calculate the fuzzy indiscernibility relation, five different T_{cos} - Transitive kernel functions are proposed in [10] [11] [12]. It is employed to determine the degree of similarity among two objects. Here the fuzzy T-equivalence relation with the equation gaussian kernel function is considered to estimate the indiscernibility relation. From that, further process will takes place to obtain the recommendation.

Let's start on with a data table IS whose rows represent a set of stakeholders in an organization and columns represent a set of functional and technical competence that describe those stakeholders skill levels. Formally, the set $S = \{S_1, S_2, S_3, \dots, S_n\}$ of stakeholders and the set $C = \{C_1, C_2, C_3, \dots, C_n\}$ of functional and technical competences.

The set of competences rated by the stakeholders are examined and finds the competences similarities with the targeted competence. In the conventional recommendation systems, similarities between the competences are identified by using the cosine similarity measures (or) pearson similarity measurements. But in this paper, fuzzy indiscernibility relation computes the similarity. Once similar competences are found then rating for the new competence is predicted by taking the weighted average of the stakeholders rating on these similar competences.

Users	Java	C	CPP	Android	ionic	Spring	Python	HTML	CSS	JavaScript	XML	Scala	Hadoop	REST	Hibernate	jQuery	Struts	JSP	Servlets	EJB	Colarp	ASP.NET	PHP	Wadeen	Liferay	Drupal
1 Michele	2.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5
2 Bob	3.0	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5
3 James	2.5	3.0	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5
4 Juler	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5
5 Tom	4.0	2.0	3.0	2.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
6 Lakshika	3.0	4.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
7 Chan	4.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5
8 Abinaya	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5
9 Anshavalli	1.0	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5
10 Durga	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
11 Karthik	3.0	4.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
12 Kavitha	2.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5
13 Kohila	5.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
14 Parveen Nisha	2.0	5.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
15 Priya Sangeetha	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5
16 Rukshana Farveen	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
17 Sowmya	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
18 Sudharsan	2.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0	10.5	11.0	11.5	12.0	12.5	13.0	13.5
19 Mitthun Pranava	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
20 Uma	4.0	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5
21 Beulah crescenta j	2.5	3.0	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5
22 Karthiga p	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5
23 Maheswari m	4.0	2.0	3.0	2.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
24 Karthikey B	3.0	4.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
25 Parvathi k	4.0	4.0	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5
26 Rajesh d	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5

Figure 1 Sample Dataset.

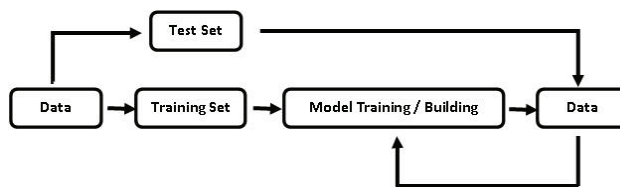


Figure 2 Experiment Plan.

3.3 Prediction

The constructed FRST based CF recommendation system helps to predict the probable rating for any stakeholder S to his functional and technical competences C by using the traditional weighted sum technique [3]. The system get all the competences similar to the target competences, and from those similar competences, it pick the competences which the active stakeholder has rated. The system weights the stakeholder’s rating for each of these competences by the similarity between that and the target competences. At the end, the system scale the prediction by the sum of similarities to get a reasonable value for the predicted rating.

4. FRSTIBCF MODEL EXPERIMENTAL RESULT

Recommended system needs dataset and pre-processing steps. Making a dataset, the real time data are collected and the dataset is preprocessed to segregate. In this experiment 90% of the data set is consider as a training set and the remaining 10% is consider as a test data for training and testing the proposed FRSTIBCF model and existing models of the recommended system.

4.1 Data Set and Pre-processing Step

To evaluate the proposed model, the sample data set is shown in the Figure 1 and it is used. It contains the collection of student’s functional and technical competence level in a rating matrix fashion (rating score from 1-5). The dimensionality of the data set is 40 rows and 26 columns. Rows represent the students and the column represents the competence levels in a particular technical course. An element e_{ij} from the data set represents the knowledge level in a technical course j by the student i. Not yet studied/registered technical courses competence levels are set as 0 by default. The dataset is divided into training set and test set to conduct the experiment and it is illustrated in Figure 2. For dividing the dataset, the popular approaches like split, bootstrapping & cross validation are used [13] [14][15].

4.2 Results and discussion

The dataset preparation approaches are split, bootstrapping and cross validation to segregates the dataset into training set and test set. The training set is used to train the proposed and existing models like Item Based Collaborative Filtering using the cosine similarity (IBCF-COS), IBCF using the pearson correlation (IBCF-COR), IBCF using the Jaccard similarity (IBCF-JAC) [18] and Singular Value Decomposition approximation (SVD)

Table 1 Rating Evaluation.

No	Dataset preparation approach	Algorithm	RMSE	MSE	MAE
1.	Split	FRSTIBCF	1.329298	1.767033	0.8569364
		IBCF-COS	1.812712	3.285924	1.2613502
		IBCF-COR	2.713602	7.363636	2.0303030
		IBCF-JAC	2.440599	5.956522	1.7173913
		SVD	1.624380	2.638610	1.4001650
2.	Bootstrapping	FRSTIBCF	1.389010	1.929348	0.7644928
		IBCF-COS	2.001946	4.007787	1.2442144
		IBCF-COR	2.796046	7.817871	2.1816406
		IBCF-JAC	2.565801	6.583333	1.8115942
		SVD	1.740973	3.030986	1.3560051
3.	Cross Validation	FRSTIBCF	1.749764	3.061673	1.123689
		IBCF-COS	2.089390	4.365551	1.347056
		IBCF-COR	2.502524	6.262626	1.868687
		IBCF-JAC	1.763149	3.108696	1.147826
		SVD	1.751814	3.068852	1.320612

Top n List Evaluation	Algorithm	Precision	Recall	TPR	FPR	Specificity	Sensitivity
Evaluation@1	FRSTIBCF	0.2314685	0.06667468	0.06667468	0.02986517	0.97018349	0.06667468
	IBCF-COS	0.09696970	0.02424242	0.02424242	0.03618479	0.96444954	0.02424242
	IBCF-COR	0.20202020	0.08530544	0.08530544	0.02619083	0.97362386	0.08530544
	IBCF-JAC	0.00000000	0.00000000	0.00000000	0.02275908	0.97706421	0.00000000
	SVD	0.1190476	0.04096521	0.04096521	0.04234816	0.95756880	0.04096521
Evaluation@5	FRSTIBCF	0.1406993	0.25556357	0.25556357	0.16554769	0.83600915	0.25556357
	IBCF-COS	0.09748252	0.21907969	0.21907969	0.17871184	0.82454130	0.21907969
	IBCF-COR	0.13468013	0.24305756	0.24305756	0.13848919	0.86123853	0.24305756
	IBCF-JAC	0.12357143	0.2662177	0.2662177	0.09987045	0.89908258	0.2662177
	SVD	0.1095238	0.32807439	0.32807439	0.21505981	0.78555045	0.32807439
Evaluation@10	FRSTIBCF	0.1077399	0.31953664	0.31953664	0.31044280	0.69151371	0.31953664
	IBCF-COS	0.10572067	0.35101010	0.35101010	0.33140444	0.67201833	0.35101010
	IBCF-COR	0.10538721	0.32399391	0.32399391	0.28649765	0.71330269	0.32399391
	IBCF-JAC	0.11142857	0.3440516	0.3440516	0.19559653	0.80160553	0.3440516
	SVD	0.1142857	0.53396665	0.53396665	0.42645265	0.57339449	0.53396665

[19]. After training the test data is given to the proposed and existing models. Root Mean Square Error (RMSE), Mean Square Error (MSE) and Mean Absolute Error (MAE) values are calculated for predicting the rating value [16] and it is tabulated in Table 1.

The proposed model FRSTIBCF has much lower error rate and smaller MAE value is observed from the table 1 and it indicated that the proposed model is the best fit model than the other existing models. The cross validation approach is the most accurate one as compared to the remaining two approaches [17]. Here, the cross validation approach is considered for segregate the dataset and validate the recommendation evaluation experiment.

In order to evaluate the proposed FRSTIBCF model’s recommendation with the existing models recommendation, the recommendation’s positive ratings are compared. To know more

about the model’s sensitivity and precision, the experiments are conducted with different numbers of top n recommended items [20]. For example, the models are evaluated while recommending only one course, 5 numbers of course and 10 numbers of course.

The performances of the proposed and existing models are validated by the metrics like precision, recall, True Positive Rate (TPR) and False Positive Rate (FPR) and the outputs are listed out in Table 2. The measured precision represents the percentage of recommended courses that have been registered and the measured sensitivity/recall represents the percentage of registered courses that have been recommended. Specificity measure shows that, the ability of the proposed model to identifies the negative cases.

ROC curve of FPR Vs TPR and Precision Vs Recall are shown in figure 3 and figure 4 respectively. From the figure 3 and figure

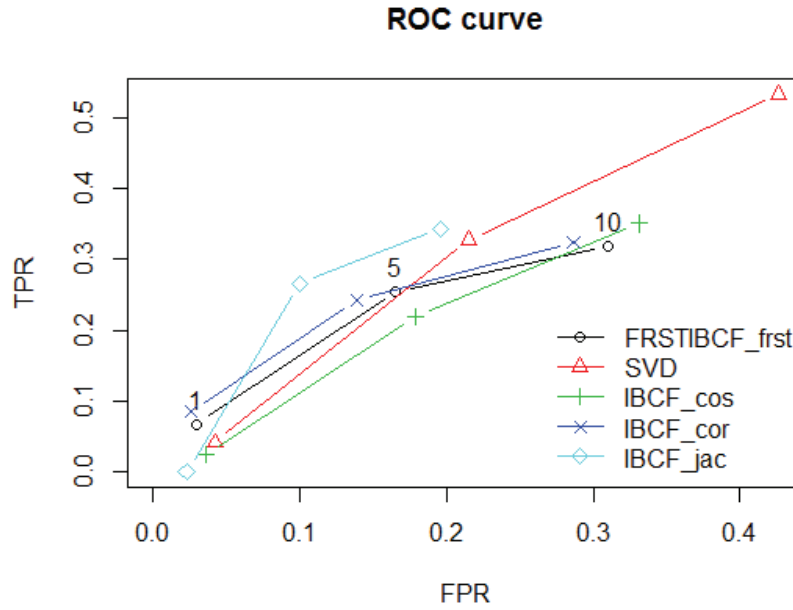


Figure 3 TPR Vs FPR.

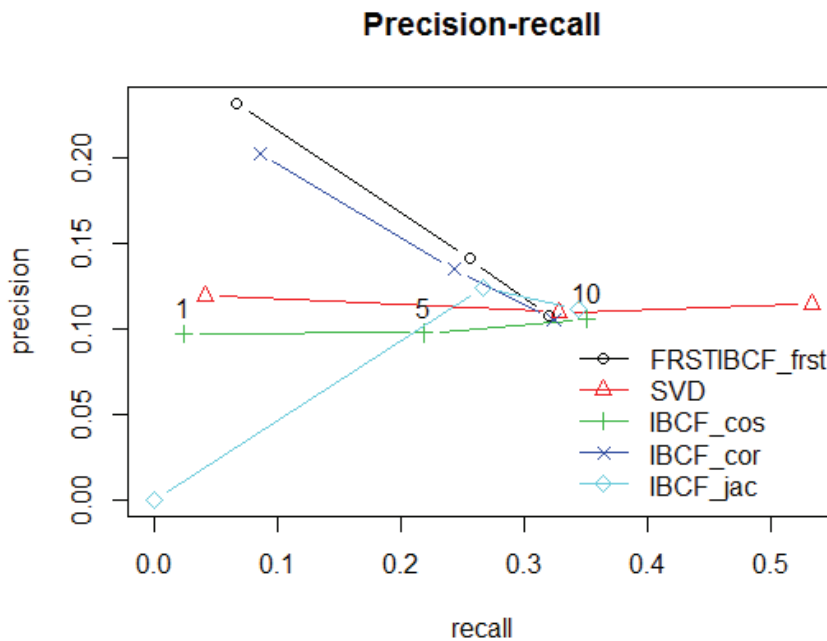


Figure 4 Precision Vs Recall.

4 The proposed model works much better than the other models, especially if the goal is high recall and the plotted ROC curves shown in Figure 3 is the base to measure the models performance. Usually the Area under the ROC curve (AUC) tells how well the model separates the positive cases from the negative cases. Figure 3 clearly shows that, comparatively the proposed model FRSTIBCF performances well in the lower range.

The proposed model FRSTIBCF has one of the numeric parameter k , which is for configuring the number of courses to recommend to the stakeholders by the system. Based on the k value, the relationship between FPR and TPR and recall Vs

precision are views in figure 5 and figure 6 respectively. When the choice of getting output from the system k is less then the performance of the system is high and it is observed.

5. CONCLUSION AND FUTURE WORKS

In this paper, a novel FRST is proposed to find the similarity measure in the collaborative filtering recommendation system and the model is named as FRSTIBCF. The proposed model is experimented by the student data of their carrier selection pro-

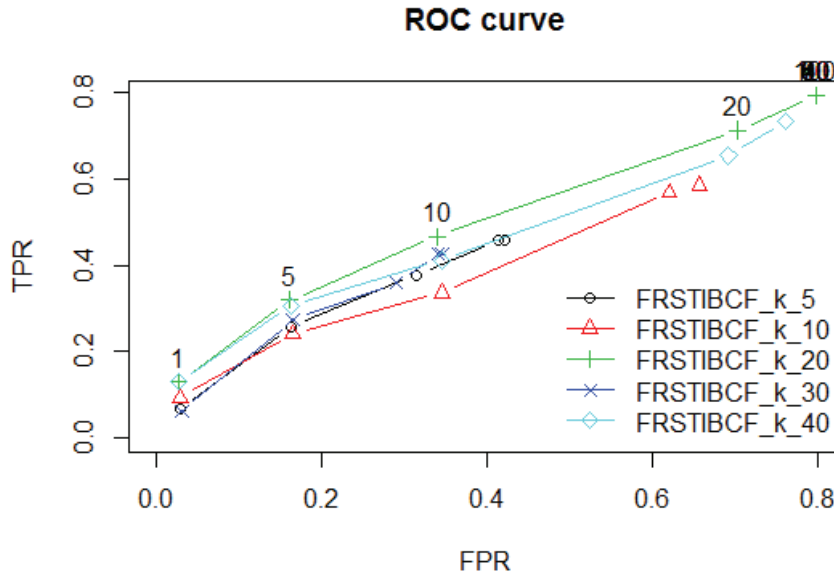


Figure 5 TPR Vs FPR.

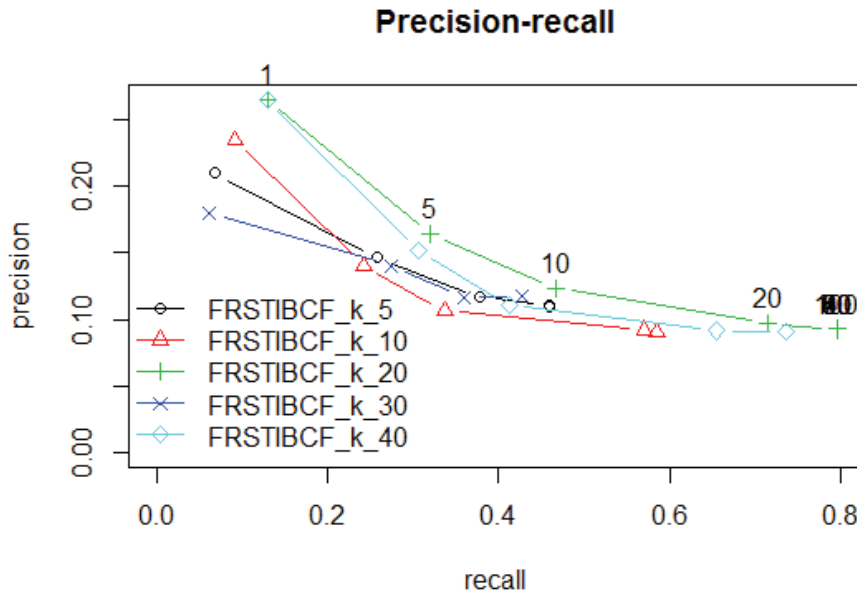


Figure 6 Precision Vs Recall.

cess. For predicting the rating value, Root Mean Square Error (RMSE), Mean Square Error (MSE) and Mean Absolute Error (MAE) values are calculated. Based on the rating the proposed model performance is evaluated using precision, recall, sensitivity, specificity, TPR and FPR. The recommendation system can be extended by using traditional prediction technique.

REFERENCES

1. Yi Cai, Ho-fung Leung, Qing Li, Huaqing Min, Jie Tang and Juanzi Li, "Typicality-based Collaborative Filtering Recommendation" in IEEE Transactions on Knowledge and Data Engineering, Volume:26 , Issue: 3, 2013, pp. 766 - 779
2. G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," Knowledge and Data Engineering, IEEE Transactions on, vol. 17, no. 6, pp. 734–749, 2005.
3. B. Sarwar, G. Karypis, J. Konstan, and J. Reidl, "Item-based collaborative filtering recommendation algorithms," in WWW '01. New York, NY, USA: ACM, 2001, pp. 285–295..Pawlak, Z. (1982) 'Rough Sets', International Journal of Computer and Information Sciences, Vol. 11, pp. 341–356.
4. Z. Pawlak, Rough sets and fuzzy sets, Fuzzy Sets and Systems 17, 1985, 99-102
5. A. M. Radzikowska and E. E. Kerre, "A Comparative Study of Fuzzy Rough Sets", Fuzzy Sets and Systems, vol. 126, p. 137 - 156 (2002).
6. D. Dubois and H. Prade, "Rough Fuzzy Sets and Fuzzy Rough

- Sets”, *International Journal of General Systems*, vol. 17, p. 91 - 209 (1990).
7. D. Dubois and H. Prade, “Putting fuzzy sets and rough sets together”, in: R. Slowinski (Ed.), *Intelligent Decision Support*, Kluwer Academic, Dordrecht, 1992, pp. 203 - 232.
 8. Z. Pawlak, *Rough Sets: Theoretical Aspects of Reasoning About Data*. Kluwer Academic Publishers, 1991.
 9. R. Jensen and Q. Shen, “New Approaches to Fuzzy-Rough Feature Selection”, *IEEE Trans. on Fuzzy Systems*, vol. 19, no. 4, p. 824 - 838 (2009).
 10. Q. Hu, D. Yu, W. Pedrycz, and D. Chen, “Kernelized Fuzzy Rough Sets and Their Applications”, *IEEE Trans. Knowledge Data Eng.*, vol. 23, no. 11, p. 1649 - 1471 (2011).
 11. M. Genton, “Classes of Kernels for Machine Learning: a Statistics Perspective”, *J. Machine Learning Research*, vol. 2, p. 299 - 312 (2001).
 12. H. Naessens, H. De Meyer, and B. De Baets, “Algorithms for the Computation of T-Transitive Closures”, *IEEE Trans. on Fuzzy Systems*, vol. 10, No. 4, p. 541 - 551 (2002).
 13. Breiman, L., *Bagging Predictors*, *Machine Learning*, 24(2), pp.123-140, 1996.
 14. Refaeilzadeh, P., Tang, L., and Liu, H., *Cross-validation*, in *Encyclopedia of Database Systems*, pp. 532–538, Springer US, 2009.
 15. Picard, R. R. and Cook, R. D., *Cross-validation of regression models*, *Journal of the American Statistical Association*, 79 , 575–583, 1984.
 16. I. C. G. Alejandro Bellogín Kouki, Pablo Castells Azpilicueta. *Recommender system performance evaluation and prediction: An information retrieval perspective*. 2012.
 17. C. Schaffer, *Selecting a classification method by cross-validation*, *Machine Learning* 13 (1993) 135–143.
 18. Michael Hahsler (2016). *recommenderlab: A Framework for Developing and Testing Recommendation Algorithms*
 19. Paolo Cremonesi, Yehuda Koren, Roberto Turrin: *Performance of recommender algorithms on top-n recommendation tasks*. *RecSys 2010*: 39-46. <https://doi.org/10.1145/1864708.1864721>
 20. Daniel Valcarce, Alejandro Bellogín, Javier Parapar, Pablo Castells: *On the robustness and discriminative power of information retrieval metrics for top-N recommendation*. *RecSys 2018*: 260-268. <https://doi.org/10.1145/3240323.3240347>