Research Article

A Novel Approach To Solve Cold Start Problem In Sentiment Associated Content-Based Recommender System: K-L Divergence Method

Raj Kumar1*, Dr. Shaveta Bhatia2

1: Research Scholar and Assistant Professor, Department of Computer Applications, Faculty of Computer Applications, Manav Rachna International Institute of Research and Studies (Deemed to be University), Faridabad. Email: raj.k1033@gmail.com
2: Professor and Head, Department of Computer Applications, Faculty of Computer Applications, Manav Rachna International Institute of Research and Studies (Deemed to be University), Faridabad. Email: shaveta.fca@mriu.edu.in

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 16 April 2021

Abstract:
E-learning has become a prominent part of education nowadays where recommendation system is an integral part of it. Recommender systems provide suitable course recommendations to an interested learner. But recommendation systems also have certain limitations. Cold-start problem is one such problem. In the case of a new course, generating the recommendations is very tedious task due to non-availability of past data related to that course. In such scenario, K-L divergence method is used to recommend the list of tentative students for newly launched course. Further, the overall sentiment of a student is used to boost the initial recommendations.

Keywords: e-Learning**,** K-L Divergence, Recommender System, Cold-Start, Sentiment Score

1. Introduction

The exponential growth of e-learning resources has resulted into mass availability of large number of courses online. If a user registers with a specific e-learning portal and shares his liking and disliking, then depending on the similarity between the user attributes and the e-learning course, recommendations can be easily made. This is the most fundamental principle on which recommender system works. Raj Kumar and Bhatia [1] have enhanced the recommendations by incorporating the sentiment scores associated with the course. One of the major challenges in above approach is when an organization launches a new course. In this scenario, the organization does not have the sentiment associated with the course. Due to non-availability of required information recommendations for the course cannot be provided. This is known as the problem of "*cold-start*" in recommender systems. The solution to cold-start problem in recommender system can generate lot of business opportunities for an organization which has launched new courses. The e-learning service provider is interested in finding the suitable list of learners to whom they can share the course details.

In e-commerce industry, the recommendations about various products are made to customer based on the past interaction. But, when a new user visits many e-commerce websites for various products. Here the problem of cold-start comes into picture for making better recommendations [2].The various researchers have given various ways to generate recommendations. Few of the recommendation methods are Content-based, Collaborative filtering, Demographic filtering, Knowledge-based recommender system and Hybrid-recommender systems [3]. To make recommendations about the tweets on Twitter, matrix factorization is an effective method to give recommendations. Neural network based meta-learning strategy has been used to generate solutions to cold-start problem [4].Big data techniques can be used to solve the problem of cold-start in the recommender systems of these days [5].In collaborative filtering method, a novel matrix completion strategy has been used to solve the problem of cold-start. The matrix completion approach utilizes the similarity information between the user and the item and then generates the recommendations [6].Few researchers have used mathematical model of bipartite approach to solve the cold start problem [7]. Few researchers have used deep learning to provide recommendations to solve the problem of cold start in recommender systems [8]. There is one variant of cold-start problem which is faced by e-commerce website. Due

to change in user preferences over a period of time, the e-commerce website faces the problem of continuous cold-start. The continuous cold-start problem occur when some new product is launched or the user liking changes [9].
In this paper, we are focusing on the following research question:

R Q 1: How to recommend the newly launched courses to potential learners i.e. solving the cold-start problem in recommender system?

This paper is organized into 5 sections. After a brief introduction about e-Learning recommender system and their associated limitation of cold start in section 1, a detailed literature survey about the cold-start problem in recommender systems and similarity measuring techniques are given in section 2, section 3 contains the proposed solution to the cold-start problem in recommender system where the suggested K-L divergence technique is explained in detail, detailed simulation based investigation for the proposed recommender system is carried out in section 4. Finally, discussion about the findings and the conclusion is given in section 5.

2. Related Work

Ke Yin Cao, Yu Liu1, and Hua Xin Zhang [10] proposed a new method to overcome cold-start problem. The researchers have worked on community detection algorithm. The bipartite approach was used to identify the similarity between the user and the item. Louvain algorithm has been used to identify the community detection. Pearson correlation coefficient was used to calculate the single mode network. The researchers run the test cases on multiple datasets and eventually proved that the new approach had effectively improved the cold start problem.

Saman Forouzandeh, Saman Forouzandeh, Shuxiang Xu and Soran Forouzandeh [11] proposed a Cuckoo algorithm on facebook data to overcome cold-start problem. The researchers have tried to work on the cold start cases where the past behaviors of user are not available on social media. When a user is using social media, then depending on the past behavior of user, it is really very easy to recommend the content and information to that particular user. But this task becomes challenging if we don't have past data. At this stage, data mining techniques can play a crucial role to suggest recommendations. The Cuckoo algorithm was proposed as a solution to cold start. The algorithm makes use clustering techniques and association rules.

Maksims Volkovs, Guang Wei Yu, Tomi Poutanen[12] proposed content based neighborhood method for finding a solution to cold-start problem. A model was created to benchmark cold start problems. Models were created for offline as well as online mode. Here the model produced 80% successful results in online mode. The XING (European version of LinkedIn) platform was used to benchmark the results. The data from user-job interactions from the career related social sites was used. The purpose of this data was to create recommendations and then test these recommendations against the benchmark standards. After conducting various test results on different users, researchers were of the opinion that the inclination of users keep on changing towards jobs with time. That's why creating temporal modeling is better for such users. The temporal model can incorporate user's inclination and hence it can provide better recommendations.

Maksims Volkovs, Guang Wei Yu, Tomi Poutanen [13] implemented collaborative filtering based method to provide a solution to the cold-start problem. In this paper, the researcher has worked on neural network based latent model known as dropoutnet for providing the solution for cold start problem. The latent model is used because of its ability to scale and better performance. Collaborative filtering approach has been used to generate recommendations in this paper. The collaborative filtering works on two principles. One is neighbor-based and the other is model based methodologies. Due to lack of data sparsity, the concept of cold start arises.

Youssouf ElAllioui [14] proposed a new method for cold-start problem in recommender systems using collaborative filtering. In this paper, collaborative filtering is used to generate recommendations based on the concept of neighborhood method. The researcher has included the demographic information of the user to find similarity to the neighborhood. The dataset has been taken from GroupLens, Online Communities, digital libraries, etc. In this paper, each user was classified in a group and ratings prediction was used to result ratings for items.

Li Li and Xiao-jia Tang [15] proposed social choice theory for generating recommendations. In this paper, collaborating technique has been used for providing recommendations. The central idea in collaborating filtering is based determining and locating the likeminded users. But, if there is a new user, then in that case making recommendations is difficult due to lack of insufficient information about the liking and disliking of that user. In this paper, social choice theory has been used for providing recommendations to new users. Social choice theory takes into consideration the individual liking and judgments to reach to a collective decision.

Blerina Lika, Kostas Kolomvatsos, Stathes Hadjiefthymiades [16] provided solution for the cold start problem in recommender systems. In this paper, the researcher has focused on content based as well as collaborative filtering approach for generating recommendations. But, it is difficult to provide recommendation in case of new user i.e. cold start problem. The information about the inclination of a user related to a particular item is not available then classification algorithm alongwith similarity technique is used to make recommendations.

Vala Ali Rohani, Zarinah Mohd Kasirun, Sameer Kumar, and Shahaboddin Shamshirband [17] proposed a new method for Academic Social Networks cold-start problem. In this paper, the researchers have used extended content based algorithm using social networking. The proposed model used the ratings of colleagues and friends in addition to the inclination of a user. An academic social network model was created and named as MyExpert in Malaysia. Random, content based and collaborative approach wasused on data to derive results using MyExpert. The empirical results represented significantly accurate results.

3. Proposed Method:

The proposed below mentioned block diagram of recommender system with sentiment score as booster is implemented as content based recommender system. The recommendations are calculated based on similarity between student attributes and course attributes.

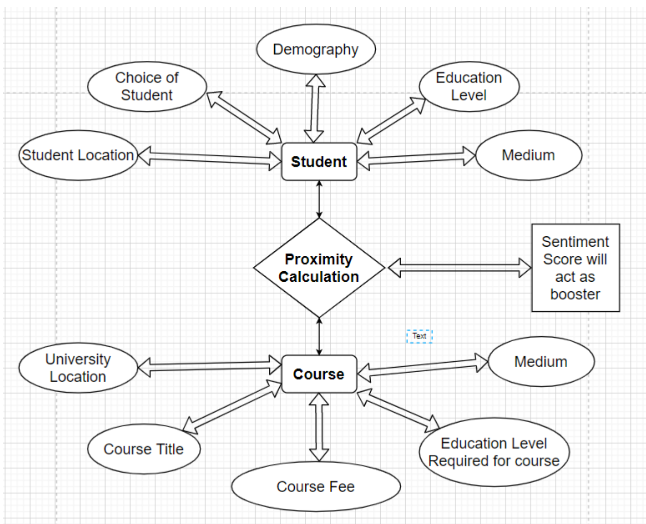

Figure 1: Block Diagram of Recommender System with Sentiment Score as booster (Source: Self)

The student's attributes are: (i) Student Location means geographical location, (ii) Choice of Student means course title, (iii) Demography means financial condition, (iv) Education Level means academics achievements till date, and (v) Medium means language in which student is willing to learn. The course's attributes are: (i) University Location means geographical location, (ii) Course Title means course being offered, (iii) Course Fee means free or paid course, (iv) Education Level means pre-requisite academic requirements for a course, and (v) Medium means language in which course is offered

3.1 Assumptions for implementation

To implement the proposed model, the following assumptions have been made:

The scores of students attribute range from 1 to 9, the score of course attributes range from 1 to 9, the total number of courses: 10,and the total number of students: 50. Let's assume that the five attributes of student are x1, x2, x3, x4 and x5 and the attributes of the course are y1, y2, y3, y4 and y5.

$$S = KL\ (Course\ ||\ Student) \tag{1}$$

At initial stage, the recommendations will be based on the mapping between the student attributes and course attributes which is done by using K-L Divergence between the Student and the Course matrix. The overall sentiment score will be incorporated to generate more relevant recommendations.
The formula to include sentiment scores for final recommendations is given below:

$$\overline{S} = S * \text{sentiment score} \tag{2}$$

3.2 Efficient Algorithm with K-L Divergence Technique:

The student data and the course data will be in the form as given in Table 1. The attribute values (xij) are positive integers.

Table 1: Student and Course Data Format

Student / Course	Attribute 1	Attribute 2	…	Attribute M
1	X11	X12	…	X1M
2	X21	X22	…	X2M
3	X31	X32	…	X3M
4	X41	X42	…	X4M
.	.	.	…	.
.	.	.	…	.
N	XN1	XN2	…	XNM

Step I:

SMatrix, CMatrix and AMatrix are student data, course data and Attributes weights respectively.

SMatrix = [Xij] n X m : N = Number of students
CMatrix = [Ykj] p X m : P = Number of Courses
AMatrix = [Z1j] 1 X m with condition Sum(AMatrix) =1: M= Number of attributes which will be :equal

Step II:

Normalize SMatrix and CMatrixasNSMatrix and NCMatrix with a condition:

Sum(NSMatrix[a,:] = 1 ɏi&Sum(NCMatrix[k,:] = 1 ɏ k

i.e. attribute value will be 1 student wise and coursewise

Step III:

Compute the effectiveness ofnormalized attribute scores i.e.

ESmatrix = NSmatrix(i,1) XAmatrix
ECmatrix = NSmatrix(k,1) XAmatrix

Step IV:

Compute KL Divergence between student and course attributes.

KL (Course || Student) = abs (ECmatrix.log(ECmatrix / ESmatrix)

Choose the minimum value as the best case.

If there is only one course:
    Recommend the course and exit.
Else:
    Go to step V.

Step V:

In order to obtain the most suitable student for a new course, compute the overall sentiment score of a student: (SSenti) as: SSenti= [Sij]n X 1

Now, apply the sentiment score with the distance computed in step IV as given below:

If KL (Course || Student) = 0:
Select the course with maximum positive sentiment.
Else:
Final-score = K-L (Course || Student) * Csenti

Hence, Final-score gives the efficient similarity score between the student and the course. Now, ignore the courses which have negative similarity score as it indicates high level of negative sentiment for the course and choose the minimum positive value as the best case for course recommendation.

4. Result and Discussion:

Table 2: Initial values for 10 courses

	Location of Student	Course Choice	Demography	Academic Qualification	Medium of Education
C1	6.000	2.000	3.000	9.000	3.000
C2	5.000	4.000	4.000	9.000	7.000
C3	9.000	8.000	1.000	1.000	6.000
C4	6.000	8.000	9.000	7.000	5.000
C5	8.000	1.000	2.000	3.000	7.000
C6	5.000	4.000	1.000	4.000	6.000
C7	4.000	5.000	4.000	5.000	2.000
C8	8.000	4.000	2.000	9.000	2.000
C9	1.000	6.000	5.000	4.000	9.000
C10	2.000	6.000	4.000	9.000	2.000

Table 3: Initial values for 50 students

	Location of Student	Course Choice	Demography	Academic Qualification	Medium of Education
S1	1.000	5.000	2.000	5.000	8.000
S2	6.000	6.000	7.000	6.000	4.000
S3	8.000	4.000	5.000	5.000	8.000
S4	7.000	8.000	2.000	7.000	4.000
S5	2.000	7.000	4.000	5.000	7.000
S6	4.000	9.000	6.000	9.000	4.000
S7	5.000	5.000	2.000	2.000	8.000
S8	9.000	3.000	7.000	1.000	7.000
S9	2.000	1.000	3.000	1.000	4.000
S10	8.000	6.000	9.000	1.000	2.000
S11	6.000	8.000	3.000	4.000	8.000
S12	4.000	4.000	7.000	5.000	9.000
S13	2.000	1.000	2.000	4.000	3.000

S14	4.000	3.000	3.000	7.000	7.000
S15	5.000	2.000	1.000	6.000	4.000
S16	2.000	3.000	6.000	7.000	8.000
S17	6.000	4.000	7.000	9.000	7.000
S18	3.000	5.000	5.000	9.000	2.000
S19	4.000	5.000	4.000	2.000	8.000
S20	6.000	8.000	6.000	2.000	9.000
S21	3.000	5.000	6.000	7.000	5.000
S22	3.000	9.000	7.000	1.000	8.000
S23	6.000	6.000	6.000	5.000	6.000
S24	3.000	9.000	9.000	5.000	2.000
S25	8.000	3.000	2.000	8.000	2.000
S26	9.000	7.000	7.000	5.000	4.000
S27	7.000	3.000	3.000	4.000	7.000
S28	4.000	7.000	2.000	7.000	8.000
S29	6.000	7.000	6.000	7.000	8.000
S30	1.000	1.000	5.000	5.000	3.000
S31	9.000	3.000	5.000	4.000	5.000
S32	8.000	3.000	6.000	2.000	1.000
S33	8.000	7.000	7.000	6.000	2.000
S34	3.000	8.000	4.000	3.000	2.000
S35	6.000	4.000	6.000	1.000	7.000
S36	1.000	8.000	4.000	7.000	5.000
S37	4.000	7.000	8.000	3.000	2.000
S38	3.000	1.000	8.000	4.000	5.000
S39	2.000	6.000	3.000	7.000	2.000
S40	2.000	4.000	6.000	4.000	1.000
S41	4.000	9.000	6.000	7.000	8.000
S42	1.000	1.000	5.000	4.000	6.000
S43	6.000	5.000	8.000	7.000	9.000
S44	5.000	4.000	3.000	7.000	7.000
S45	7.000	5.000	3.000	4.000	6.000
S46	7.000	7.000	2.000	1.000	8.000
S47	6.000	3.000	9.000	3.000	8.000
S48	1.000	8.000	6.000	4.000	9.000
S49	1.000	5.000	5.000	3.000	1.000
S50	3.000	1.000	6.000	2.000	8.000

Table 4: Normalized course attribute values

	Location of Student	Course Choice	Demography	Academic Qualification	Medium of Education
C1	0.261	0.087	0.130	0.391	0.130
C2	0.172	0.138	0.138	0.310	0.241

C3	0.360	0.320	0.040	0.040	0.240
C4	0.171	0.229	0.257	0.200	0.143
C5	0.381	0.048	0.095	0.143	0.333
C6	0.250	0.200	0.050	0.200	0.300
C7	0.200	0.250	0.200	0.250	0.100
C8	0.320	0.160	0.080	0.360	0.080
C9	0.040	0.240	0.200	0.160	0.360
C10	0.087	0.261	0.174	0.391	0.087

Table 5: Normalized student attribute values

	Location of Student	Course Choice	Demography	Academic Qualification	Medium of Education
S1	0.048	0.238	0.095	0.238	0.381
S2	0.207	0.207	0.241	0.207	0.138
S3	0.267	0.133	0.167	0.167	0.267
S4	0.250	0.286	0.071	0.250	0.143
S5	0.080	0.280	0.160	0.200	0.280
S6	0.125	0.281	0.188	0.281	0.125
S7	0.227	0.227	0.091	0.091	0.364
S8	0.333	0.111	0.259	0.037	0.259
S9	0.182	0.091	0.273	0.091	0.364
S10	0.308	0.231	0.346	0.038	0.077
S11	0.207	0.276	0.103	0.138	0.276
S12	0.138	0.138	0.241	0.172	0.310
S13	0.167	0.083	0.167	0.333	0.250
S14	0.167	0.125	0.125	0.292	0.292
S15	0.278	0.111	0.056	0.333	0.222
S16	0.077	0.115	0.231	0.269	0.308
S17	0.182	0.121	0.212	0.273	0.212
S18	0.125	0.208	0.208	0.375	0.083
S19	0.174	0.217	0.174	0.087	0.348
S20	0.194	0.258	0.194	0.065	0.290
S21	0.115	0.192	0.231	0.269	0.192
S22	0.107	0.321	0.250	0.036	0.286
S23	0.207	0.207	0.207	0.172	0.207
S24	0.107	0.321	0.321	0.179	0.071
S25	0.348	0.130	0.087	0.348	0.087
S26	0.281	0.219	0.219	0.156	0.125
S27	0.292	0.125	0.125	0.167	0.292
S28	0.143	0.250	0.071	0.250	0.286
S29	0.176	0.206	0.176	0.206	0.235
S30	0.067	0.067	0.333	0.333	0.200
S31	0.346	0.115	0.192	0.154	0.192

S32	0.400	0.150	0.300	0.100	0.050
S33	0.267	0.233	0.233	0.200	0.067
S34	0.150	0.400	0.200	0.150	0.100
S35	0.250	0.167	0.250	0.042	0.292
S36	0.040	0.320	0.160	0.280	0.200
S37	0.167	0.292	0.333	0.125	0.083
S38	0.143	0.048	0.381	0.190	0.238
S39	0.100	0.300	0.150	0.350	0.100
S40	0.118	0.235	0.353	0.235	0.059
S41	0.118	0.265	0.176	0.206	0.235
S42	0.059	0.059	0.294	0.235	0.353
S43	0.171	0.143	0.229	0.200	0.257
S44	0.192	0.154	0.115	0.269	0.269
S45	0.280	0.200	0.120	0.160	0.240
S46	0.280	0.280	0.080	0.040	0.320
S47	0.207	0.103	0.310	0.103	0.276
S48	0.036	0.286	0.214	0.143	0.321
S49	0.067	0.333	0.333	0.200	0.067
S50	0.150	0.050	0.300	0.100	0.400

Table 6: Uniformly generated attribute weights

	Location of Student	Course Choice	Demography	Academic Qualification	Medium of Education
Weightage	0.2	0.25	0.15	0.18	0.22

Table 7: Effective attribute values for course

	Location of Student	Course Choice	Demography	Academic Qualification	Medium of Education
C1	0.052	0.022	0.020	0.070	0.029
C2	0.034	0.034	0.021	0.056	0.053
C3	0.072	0.080	0.006	0.007	0.053
C4	0.034	0.057	0.039	0.036	0.031
C5	0.076	0.012	0.014	0.026	0.073
C6	0.050	0.050	0.008	0.036	0.066
C7	0.040	0.063	0.030	0.045	0.022
C8	0.064	0.040	0.012	0.065	0.018
C9	0.008	0.060	0.030	0.029	0.079
C10	0.017	0.065	0.026	0.070	0.019

Table 8: Effective attribute values for students

	Location of Student	Course Choice	Demography	Academic Qualification	Medium of Education
S1	0.010	0.060	0.014	0.043	0.084
S2	0.041	0.052	0.036	0.037	0.030

S3	0.053	0.033	0.025	0.030	0.059
S4	0.050	0.071	0.011	0.045	0.031
S5	0.016	0.070	0.024	0.036	0.062
S6	0.025	0.070	0.028	0.051	0.028
S7	0.045	0.057	0.014	0.016	0.080
S8	0.067	0.028	0.039	0.007	0.057
S9	0.036	0.023	0.041	0.016	0.080
S10	0.062	0.058	0.052	0.007	0.017
S11	0.041	0.069	0.016	0.025	0.061
S12	0.028	0.034	0.036	0.031	0.068
S13	0.033	0.021	0.025	0.060	0.055
S14	0.033	0.031	0.019	0.053	0.064
S15	0.056	0.028	0.008	0.060	0.049
S16	0.015	0.029	0.035	0.048	0.068
S17	0.036	0.030	0.032	0.049	0.047
S18	0.025	0.052	0.031	0.068	0.018
S19	0.035	0.054	0.026	0.016	0.077
S20	0.039	0.065	0.029	0.012	0.064
S21	0.023	0.048	0.035	0.048	0.042
S22	0.021	0.080	0.038	0.006	0.063
S23	0.041	0.052	0.031	0.031	0.046
S24	0.021	0.080	0.048	0.032	0.016
S25	0.070	0.033	0.013	0.063	0.019
S26	0.056	0.055	0.033	0.028	0.028
S27	0.058	0.031	0.019	0.030	0.064
S28	0.029	0.063	0.011	0.045	0.063
S29	0.035	0.051	0.026	0.037	0.052
S30	0.013	0.017	0.050	0.060	0.044
S31	0.069	0.029	0.029	0.028	0.042
S32	0.080	0.038	0.045	0.018	0.011
S33	0.053	0.058	0.035	0.036	0.015
S34	0.030	0.100	0.030	0.027	0.022
S35	0.050	0.042	0.038	0.008	0.064
S36	0.008	0.080	0.024	0.050	0.044
S37	0.033	0.073	0.050	0.023	0.018
S38	0.029	0.012	0.057	0.034	0.052
S39	0.020	0.075	0.023	0.063	0.022
S40	0.024	0.059	0.053	0.042	0.013
S41	0.024	0.066	0.026	0.037	0.052
S42	0.012	0.015	0.044	0.042	0.078
S43	0.034	0.036	0.034	0.036	0.057
S44	0.038	0.038	0.017	0.048	0.059
S45	0.056	0.050	0.018	0.029	0.053

S46	0.056	0.070	0.012	0.007	0.070
S47	0.041	0.026	0.047	0.019	0.061
S48	0.007	0.071	0.032	0.026	0.071
S49	0.013	0.083	0.050	0.036	0.015
S50	0.030	0.013	0.045	0.018	0.088

Table 9: Transpose of the Similarity score values with K-L divergence method between course and student

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
S1	0.047	0.018	0.072	0.006	0.066	0.001	0.015	0.070	0.005	0.033
S2	0.044	0.026	0.089	0.004	0.096	0.051	0.014	0.048	0.074	0.049
S3	0.037	0.011	0.074	0.015	0.042	0.025	0.032	0.051	0.059	0.083
S4	0.022	0.004	0.050	0.015	0.081	0.018	0.013	0.014	0.059	0.020
S5	0.057	0.000	0.068	0.012	0.092	0.018	0.005	0.066	0.013	0.029
S6	0.038	0.014	0.087	0.010	0.128	0.044	0.005	0.039	0.055	0.014
S7	0.052	0.005	0.035	0.006	0.029	0.007	0.016	0.060	0.013	0.072
S8	0.059	0.031	0.067	0.020	0.033	0.034	0.044	0.069	0.072	0.127
S9	0.067	0.016	0.093	0.025	0.041	0.031	0.059	0.101	0.036	0.116
S10	0.073	0.059	0.066	0.001	0.094	0.064	0.014	0.060	0.098	0.080
S11	0.048	0.000	0.040	0.014	0.058	0.003	0.001	0.047	0.023	0.045
S12	0.056	0.012	0.098	0.016	0.064	0.034	0.041	0.083	0.037	0.077
S13	0.019	0.002	0.126	0.026	0.065	0.041	0.039	0.047	0.065	0.057
S14	0.023	0.005	0.100	0.017	0.056	0.024	0.031	0.048	0.045	0.055
S15	0.001	0.006	0.089	0.018	0.044	0.022	0.022	0.017	0.077	0.052
S16	0.044	0.005	0.127	0.021	0.077	0.041	0.044	0.079	0.037	0.060
S17	0.031	0.013	0.114	0.019	0.075	0.047	0.033	0.052	0.069	0.059
S18	0.023	0.016	0.125	0.004	0.136	0.063	0.006	0.032	0.082	0.016
S19	0.066	0.005	0.054	0.002	0.048	0.009	0.024	0.079	0.015	0.074
S20	0.071	0.013	0.046	0.009	0.060	0.014	0.013	0.073	0.023	0.066
S21	0.042	0.015	0.109	0.006	0.104	0.049	0.019	0.058	0.054	0.039
S22	0.098	0.019	0.047	0.019	0.090	0.020	0.006	0.095	0.007	0.052
S23	0.048	0.018	0.075	0.002	0.076	0.035	0.016	0.054	0.055	0.057
S24	0.071	0.040	0.087	0.015	0.166	0.067	0.004	0.063	0.060	0.022
S25	0.004	0.010	0.090	0.015	0.062	0.040	0.011	0.001	0.124	0.044
S26	0.043	0.030	0.068	0.000	0.080	0.043	0.009	0.040	0.083	0.058
S27	0.031	0.004	0.066	0.016	0.029	0.015	0.032	0.045	0.056	0.087
S28	0.032	0.011	0.062	0.009	0.067	0.006	0.003	0.042	0.021	0.028
S29	0.043	0.009	0.078	0.002	0.075	0.029	0.016	0.053	0.045	0.050
S30	0.040	0.022	0.183	0.033	0.110	0.080	0.051	0.078	0.074	0.058
S31	0.032	0.023	0.075	0.019	0.042	0.036	0.031	0.039	0.092	0.093
S32	0.044	0.055	0.076	0.014	0.070	0.066	0.022	0.037	0.140	0.093
S33	0.039	0.034	0.078	0.001	0.102	0.055	0.003	0.033	0.096	0.044
S34	0.065	0.024	0.044	0.032	0.141	0.033	0.025	0.043	0.039	0.008

S35	0.070	0.024	0.062	0.010	0.045	0.027	0.035	0.080	0.045	0.103
S36	0.047	0.002	0.080	0.018	0.132	0.029	0.008	0.052	0.021	0.004
S37	0.075	0.046	0.079	0.010	0.138	0.065	0.002	0.065	0.068	0.041
S38	0.059	0.034	0.151	0.036	0.081	0.075	0.064	0.096	0.074	0.099
S39	0.027	0.007	0.092	0.014	0.145	0.043	0.012	0.028	0.055	0.002
S40	0.058	0.043	0.120	0.002	0.159	0.084	0.011	0.062	0.083	0.034
S41	0.052	0.007	0.073	0.009	0.094	0.026	0.006	0.058	0.029	0.032
S42	0.055	0.010	0.147	0.033	0.069	0.048	0.065	0.104	0.036	0.084
S43	0.046	0.014	0.098	0.015	0.069	0.039	0.034	0.067	0.052	0.069
S44	0.024	0.003	0.086	0.011	0.056	0.021	0.023	0.042	0.047	0.052
S45	0.035	0.007	0.054	0.001	0.046	0.015	0.013	0.038	0.053	0.064
S46	0.056	0.000	0.015	0.016	0.030	0.012	0.002	0.049	0.017	0.068
S47	0.065	0.029	0.099	0.024	0.057	0.047	0.051	0.089	0.058	0.106
S48	0.079	0.004	0.069	0.012	0.095	0.019	0.012	0.093	0.000	0.038
S49	0.074	0.040	0.094	0.016	0.187	0.071	0.006	0.067	0.056	0.015
S50	0.070	0.015	0.110	0.034	0.038	0.035	0.074	0.116	0.033	0.128

Table 10: Uniformly generated sentiment score for 50 students

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
Sentiment Score	0.23	0.98	0.06	-0.04	0.60	-0.54	0.00	0.80	0.15	0.69
	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20
Sentiment Score	0.48	0.17	-0.51	0.33	-0.83	0.25	0.32	0.46	0.78	0.96
	S21	S22	S23	S24	S25	S26	S27	S28	S29	S30
Sentiment Score	0.54	0.16	0.86	0.16	-0.97	-0.76	0.73	-0.03	0.69	-0.58
	S31	S32	S33	S34	S35	S36	S37	S38	S39	S40
Sentiment Score	0.10	0.26	-0.94	0.23	-0.28	-0.90	-0.02	-0.61	-0.75	-0.59
	S41	S42	S43	S44	S45	S46	S47	S48	S49	S50
Sentiment Score	-0.71	-0.62	-0.91	0.27	-0.44	0.08	0.39	0.00	0.07	-0.11

Table 11: Final attribute values after applying sentiment score on K-L similarity score

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10

S1	0.011	0.017	0.004	0.000	0.040	-0.001	0.000	0.056	0.001	0.023
S2	0.010	0.026	0.005	0.000	0.058	-0.028	0.000	0.039	0.011	0.034
S3	0.008	0.010	0.004	-0.001	0.025	-0.013	0.000	0.041	0.009	0.057
S4	0.005	0.004	0.003	-0.001	0.049	-0.010	0.000	0.012	0.009	0.014
S5	0.013	0.000	0.004	0.000	0.055	-0.010	0.000	0.053	0.002	0.020
S6	0.009	0.014	0.005	0.000	0.077	-0.024	0.000	0.032	0.008	0.010
S7	0.012	0.005	0.002	0.000	0.018	-0.004	0.000	0.048	0.002	0.050
S8	0.013	0.031	0.004	-0.001	0.020	-0.018	0.000	0.056	0.011	0.088
S9	0.015	0.016	0.005	-0.001	0.024	-0.017	0.000	0.081	0.005	0.080
S10	0.016	0.058	0.004	0.000	0.057	-0.035	0.000	0.048	0.015	0.055
S11	0.011	0.000	0.002	-0.001	0.035	-0.002	0.000	0.038	0.003	0.031
S12	0.013	0.011	0.005	-0.001	0.038	-0.019	0.000	0.067	0.006	0.053
S13	0.004	0.002	0.007	-0.001	0.039	-0.022	0.000	0.038	0.010	0.039
S14	0.005	0.005	0.006	-0.001	0.033	-0.013	0.000	0.039	0.007	0.038
S15	0.000	0.005	0.005	-0.001	0.026	-0.012	0.000	0.014	0.011	0.036
S16	0.010	0.004	0.007	-0.001	0.046	-0.022	0.000	0.064	0.006	0.042
S17	0.007	0.013	0.006	-0.001	0.045	-0.025	0.000	0.042	0.010	0.041
S18	0.005	0.016	0.007	0.000	0.082	-0.034	0.000	0.026	0.012	0.011
S19	0.015	0.005	0.003	0.000	0.029	-0.005	0.000	0.063	0.002	0.051
S20	0.016	0.013	0.003	0.000	0.036	-0.008	0.000	0.059	0.003	0.046
S21	0.009	0.014	0.006	0.000	0.063	-0.027	0.000	0.047	0.008	0.027
S22	0.022	0.019	0.003	-0.001	0.055	-0.011	0.000	0.076	0.001	0.036
S23	0.011	0.017	0.004	0.000	0.046	-0.019	0.000	0.043	0.008	0.039
S24	0.016	0.039	0.005	-0.001	0.100	-0.036	0.000	0.051	0.009	0.015
S25	0.001	0.010	0.005	-0.001	0.038	-0.022	0.000	0.001	0.018	0.030
S26	0.010	0.029	0.004	0.000	0.048	-0.024	0.000	0.032	0.012	0.040
S27	0.007	0.004	0.004	-0.001	0.018	-0.008	0.000	0.036	0.008	0.060
S28	0.007	0.011	0.003	0.000	0.041	-0.003	0.000	0.033	0.003	0.020
S29	0.010	0.009	0.004	0.000	0.045	-0.016	0.000	0.043	0.007	0.034
S30	0.009	0.022	0.010	-0.001	0.066	-0.043	0.000	0.062	0.011	0.040
S31	0.007	0.023	0.004	-0.001	0.025	-0.020	0.000	0.032	0.014	0.064
S32	0.010	0.054	0.004	-0.001	0.042	-0.036	0.000	0.030	0.021	0.064
S33	0.009	0.033	0.004	0.000	0.062	-0.030	0.000	0.026	0.014	0.030
S34	0.015	0.024	0.002	-0.001	0.085	-0.018	0.000	0.035	0.006	0.006
S35	0.016	0.023	0.003	0.000	0.027	-0.015	0.000	0.065	0.007	0.071
S36	0.011	0.001	0.004	-0.001	0.080	-0.016	0.000	0.042	0.003	0.003
S37	0.017	0.045	0.004	0.000	0.083	-0.035	0.000	0.052	0.010	0.028
S38	0.013	0.033	0.008	-0.001	0.049	-0.041	0.000	0.077	0.011	0.068
S39	0.006	0.007	0.005	-0.001	0.087	-0.024	0.000	0.022	0.008	0.002
S40	0.013	0.042	0.007	0.000	0.096	-0.046	0.000	0.050	0.012	0.023
S41	0.012	0.007	0.004	0.000	0.057	-0.014	0.000	0.047	0.004	0.022
S42	0.012	0.009	0.008	-0.001	0.042	-0.026	0.000	0.084	0.005	0.058
S43	0.010	0.014	0.005	-0.001	0.041	-0.021	0.000	0.054	0.008	0.048

S44	0.006	0.003	0.005	0.000	0.034	-0.012	0.000	0.034	0.007	0.036
S45	0.008	0.007	0.003	0.000	0.028	-0.008	0.000	0.031	0.008	0.044
S46	0.013	0.000	0.001	-0.001	0.018	-0.006	0.000	0.040	0.003	0.047
S47	0.015	0.028	0.005	-0.001	0.035	-0.026	0.000	0.072	0.009	0.073
S48	0.018	0.004	0.004	-0.001	0.057	-0.010	0.000	0.074	0.000	0.026
S49	0.017	0.039	0.005	-0.001	0.113	-0.039	0.000	0.054	0.008	0.010
S50	0.016	0.015	0.006	-0.001	0.023	-0.019	0.000	0.093	0.005	0.088

4.1 Analysis for Cold Start Problem:

Cold start problem arises due to lack of data. In e-learning recommender system, cold start will arise when an organization launches a new course. In this case, there will be no previous data available for the newly launched course. So, no sentiment score is available for the course. Here, the organization will be interested in identifying potential students to whom the newly launched course may be recommended. This problem can be resolved by finding the tentative students for the course. Thus, K L divergence between the courses to students will be helpful to identify the tentative students.

In order to the cold start problem an efficient algorithm is proposed in section 3 and same is implemented in section 4. Now, from table 9 we notice that for course C1, only the student 15 has the minimum similarity score. Thus, S15 may be the tentative learner for course C1. Now, for course C2, the table 9 is showing two students S5 and S46 best suitable match. Now, in order to find the only tentative student, sentiment score is working as a booster. As the similarity score for C2||S5 and C2||S46 is zero. Whereas the overall sentiment scores of S5 and S46 is 0.60 and 0.70 respectively in table 10. So, the student S5 is the best suitable student for the course C2 as it has the higher positive sentiments.

5. Conclusion

The solution of cold start problem in recommender systems is of great importance. First, it helps in identifying the potential students for a course. Second it can help the e-learning course provider to reach the potential students. In this paper, the course and student data has been used in the proposed model to generate the potential student list. K-L divergence method has been used to generate the list based on the principle of closeness between the course and the student. Sentiment score of student has been used to boost the initial list of potential students.

References:

1. Raj Kumar and S. Bhatia, “Content-Based Recommender System with Sentiment Score”, International Journal of Information Technology (BJIT), Second Review Submitted
2. R. Venkatesan, A. Sabari, “ISSUES IN VARIOUS RECOMMENDER SYSTEM IN E-COMMERCE – A SURVEY”, Journal of Critical Reviews ISSN-2394-5125 Vol 7, Issue 7, 2020, DOI: http://dx.doi.org/10.31838/jcr.07.07.109
3. Mehdi Hosseinzadehet. al.,“A Systematic Study on the Recommender Systems in the E-Commerce”, IEEE Open Access Volume 8 2020 DOI: 10.1109/ACCESS.2020.3002803
4. ManasiVartak, ArvindThiagarajan, Conrado Miranda, JeshuaBratman and Hugo Larochelle, “A Meta-Learning Perspective on Cold-Start Recommendations for Items”, 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA
5. Mohammed FadhelAljunid&Manjaiah D.H., “A SURVEY ON RECOMMENDATION SYSTEMS FOR SOCIAL MEDIA USING BIG DATA ANALYTICS”, International Journal of Latest Trends in Engineering and Technology Special Issue SACAIM 2017, pp. 048-058 e-ISSN:2278-621X

6. ImanBarjasteh ,RanaForsati , Dennis Ross, Abdol-HosseinEsfahanian, HayderRadha, “Cold-Start Recommendation with Provable Guarantees: A Decoupled Approach, TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING”, VOL. 28, Issue 6, pp. 1462-1474, June 2016
7. Jin-Hu Liu,Tao Zhou,Zi-Ke Zhang,Zimo Yang,Chuang Liu, and Wei-Min Li, “Promoting Cold-Start Items in Recommender Systems”, Open Access, Published: December 5, 2014 https://doi.org/10.1371/journal.pone.0113457
8. ZeynepBatmaz, Ali Yurekli, Alper Bilge and CihanKaleli, “A review on deep learning for recommender systems: challenges and remedies”, Artificial Intelligence Review, https://doi.org/10.1007/s10462-018-9654-y
9. Lucas Bernardi, JaapKamps, Julia Kiseleva and Melanie J.I. Mueller, “The Continuous Cold Start Problem in e-Commerce Recommender Systems”, CBRecSys 2015, September 20, 2015, Vienna, Austria.
10. KeYin Cao, Yu Liu1, and HuaXin Zhang, “Improving the Cold Start Problem in Music RecommenderSystems”, Journal of Physics: Conference Series 1651 (2020) 012067, doi:10.1088/1742-6596/1651/1/012067
11. SamanForouzandeh, SamanForouzandeh, ShuxiangXu and SoranForouzandeh, “Addressing the Cold-start Problem Using Data Mining Techniques and Improving Recommender systems by Cuckoo Algorithm: a Case Study of Facebook”, 1521-9615 (c) 2018 IEEE
12. MaksimsVolkovs, Guang Wei Yu, TomiPoutanen, “Content-based Neighbor Models for Cold Startin Recommender Systems”, RecSys Challenge ’17, August 27, 2017, Como, Italyc 2017 Association for Computing Machinery.ACM ISBN 978-1-4503-5391-5/17/08. . . $15.00 https://doi.org/10.1145/3124791.3124792
13. MaksimsVolkovs, Guang Wei Yu, TomiPoutanen, “DropoutNet: Addressing Cold Start in Recommender Systems”, 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.
14. Youssouf El Allioui, “A novel approach to solve the new user cold-start problem in recommender systems using collaborative filtering”, International Journal of Scientific & Engineering Research Volume 8, Issue 11, November-2017 ISSN 2229-5518
15. Li Li and Xiao-jia Tang, “A Solution to the Cold-Start Problemin Recommender Systems Based on Social Choice Theory”, © Springer International Publishing Switzerland 2016K. Lavangnananda et al. (eds.), Intelligent and Evolutionary Systems,Proceedings in Adaptation, Learning and Optimization 5,DOI: 10.1007/978-3-319-27000-5_22
16. BlerinaLika, Kostas Kolomvatsos, StathesHadjiefthymiades, “Facing the cold start problem in recommender systems, Expert Systems with Applications” 41 (2014) 2065–2073, www.elsevier.com/locate/eswa
17. Vala Ali Rohani, ZarinahMohdKasirun, Sameer Kumar, and ShahaboddinShamshirband, “An Effective Recommender Algorithm for Cold-Start Problem inAcademic Social Networks”, Hindawi Publishing CorporationMathematical Problems in Engineering, Volume 2014, Article ID 123726, 11 pages, http://dx.doi.org/10.1155/2014/123726