

# **Recommender System based on Learner Knowledge and Opining using Data Mining Techniques in Synchronous E-Learning Environment**

Mohammad Daoud

Department of Computer Science  
and IT, Singhania University,  
Jhunjhunu, Rajasthan- 333515,  
India.

Centre for Education Technology,  
Indian Institute of Technology,  
New Delhi-110016, India.

Asad Ahmad

FIS Global Business Solutions  
Private Ltd., Gurgaon, Haryana  
122016, India

Alok Nikhil Jha

Cluster Innovation Centre,  
University of Delhi, New Delhi

## **ABSTRACT**

Learners are often getting uncertainty by the flow of information and have trouble in selecting the material to learn that satisfies their requirements and interests. It is the fact that the learners 'learning interest, and behaviour changes over the time and subject to subject. It is very important, thus, to know learner preferences and what problem he/she faces during the programme. In this paper, our aim to address a novel framework for an e-learning recommender system that used data mining techniques to find learner preferences and requirements from their opinion. Make a more relevant relationship between learner and his/her preferences. Proposed framework is based on opinion and the knowledge level of learner.

## **Keywords:**

Recommendation System, Structured and unstructured data,  
Learner Opinion, Clustering

## **1. INTRODUCTION**

Learners are often getting uncertainty by the flow of information and have trouble in selecting the material to learn that satisfies their requirements and interests. It is the fact that the learners 'learning interest, and behaviour changes over the time and subject to subject. It is very important, thus, to know learner preferences and what problem he/she faces during the programme. Our aim to address a such a approach for an e-learning recommender system that used data mining techniques to find learner preferences and requirements from their opinion. Make a more relevant relationship between learner and his/her preferences. Our framework is based on opinion and the knowledge level of learner. E-learning is a activity of information exchange that based on the instructions and technology. The instructions means such a system designs that purpose to distribute learning material to learners who are not physically present "on site" (i.e learner are not physically present with instructor in class room). The technology used as a interface between instructor and learner. This mode of learning is appreciated in distance learning mode of education where instructors and learners are in geographically different locations or have interactivity barriers or any other issues which prevents live sessions, many a times this model is preferred by working professionals interested in learning while working and time constraint exists. In this mode we could contain the record of registered learner query and requirements during the session, count the similarly from the learner learning pattern. This rising technology is becoming broadly utilized in institutions and universities around the world. The regular growth in technology,

e-learning learning domain is becoming more intelligent field. In this field researcher used the some interesting concept of data mining. Like as applying the clustering approach to categorize the learner through their query, feedback and online browsing pattern and providing attention individualized or group of learn. E-learning has been internally formed of as pre-packaged text, audio, and/or video. Synchronous learning and asynchronous learning are two type of e-learning. A Synchronous learning where in learners and instructors interact and learning, Instructor teaches Live in classroom and the traditional model of Classroom Learning. The faculty can understand where the class stands, what is the level of Students in understanding the topics discussed, the problem areas, the improvement areas etc. It consists all kinds of Interactivity and involvement and measurement of the cited parameters explicitly &/or implicitly. The other form of learning is Asynchronous learning where students can learn in their comfort zone, depending upon their choice and requirement. Unlike classroom model. In both modes instructor and learner are not physically available together. But in this research we focus on the live problem of synchronous e-learning program between Indian Institute of Technology (IIT) Delhi and Addis Ababa University Ethiopia.

Examples of synchronous learning are video conferencing, and Internet chats. Indian Institute of Technology (IIT) Delhi has commenced an international outreach programmed by learners in Ethiopia. As an initiative in distance edification, IIT-Delhi faculty members have been edifying courses to learners at Addis Ababa University (AAU) in Ethiopia Figure 1. The project commenced at the request of the Ambassador of Ethiopia to India who approached IIT-Delhi to increment the university level of ability [w1][w2]. IIT-Delhi found it an intriguing experiment. These courses are being distributed utilizing a two-way video/audio link utilizing IPLC/IP. Since it is a two-way process, learners additionally get to take part in the process and ask questions when needed. As many faculty members are involved in the project.

The main quandary in distance E- learning system is the instructor have the limited time of discussion with class it is very difficult to find the learner preference (i.e., all things this comes under the study such teaching quality etc.) and problem those he/she faces during the programme. For resolving these issue. We request the learner to give the feedback in structured and unstructured form. In this paper, we extracted the learner opinion from unstructured data and make it similar to structured data and using the extracted learner opinion, we identify the some new entity (questions) that are discussed by the

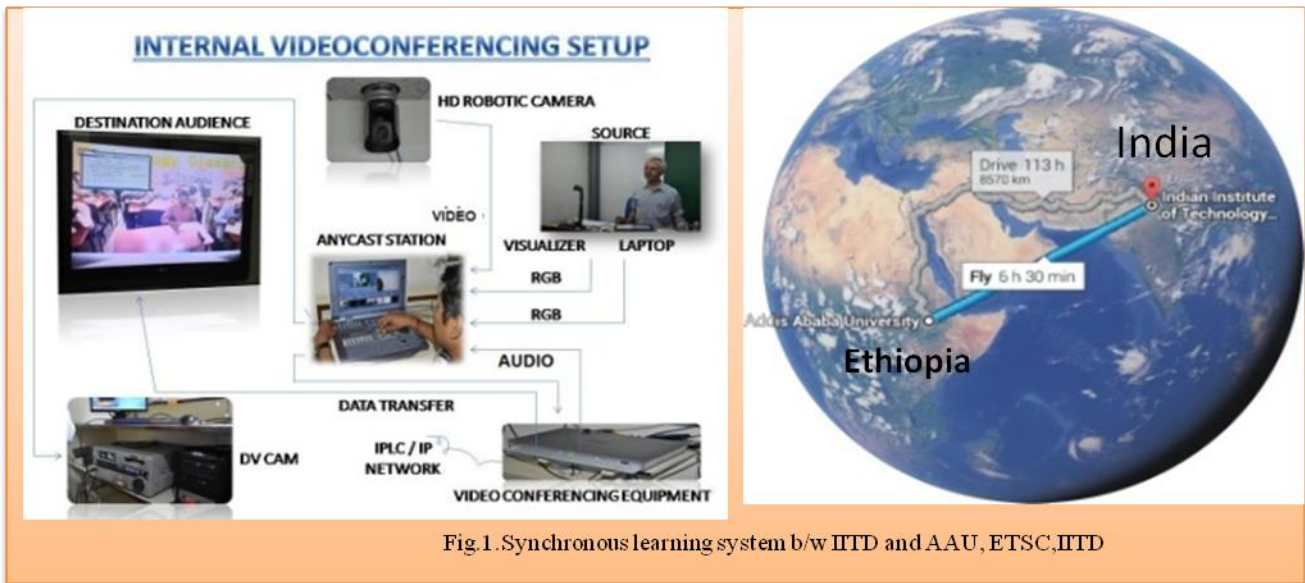


Fig.1. Synchronous learning system b/w IITD and AAU, ETSC, IITD

learner in their feedback. And also used the clustering technique to identify the same type knowledgeable learner, same type opinion holder on any specific entity and also identify those have the similar opinion on all entities. From these clustering, we identified the learner knowledge level, opinion on any specific entity and neighbourhood learner. Using this approach we purposed E-learning framework which identify the learners opinion all those aspect that's come occur during the programme and give recommendations to learners who has/have similar opinion on any specific entity and also recommends same to neighbourhood learners of these learners.

## 2. RELATED WORK

Recommender systems emerged as an independent research area in the mid 90s and they have traditionally been used in the field of electronic commerce[1][2]. Recently it is also used in E-learning. Most of the researcher using the data mining approach for retrieving the user behaviour. Learning is enhanced by information technologies like the network video conferencing class and digital libraries have evolved, they have added new environments for teaching and learning and have given rise to fresh areas for research. Web conferencing is highly used in e-learning systems for users to participate classroom discussions remotely and to conveniently access instructional content online [3] [4]. There is an extensive effort at both private and public universities/ Institute on developing e-learning systems to support distance-learning learners. Live example of distance-learning program between Indian Institute of Technology (IIT) Delhi and Addis Ababa University (AAU). Internal conferencing setup shown in Fig [1]. A recommender system in an E-learning context is a software agent that tries to “intelligently” recommend actions to a learner based on the learners' performance and their opinions. There are several methods to automatically generate recommendations based on a user's browsing patterns or explicit ratings [5]. These systems use the collaborative filtering, which stores all learners' information and then learner K Nearest Neighbours' (KNN) to provide recommendations from the earlier history of the K most similar learners [6]. Recently, [7] [8] used a different approach to recommend documents on the basis of the learner profiles. This approach learns from implicit feedback or past click history. The click history is stored in log data on the server. e-learning system have attempted to recommend learning resources or learning materials based on the rating of good learners, its benefits to learners in their learning process as it helps them to find good quality of learning materials and improve

the learning process [3]. Email group facility and Social networking site like as twitter, Facebook and blogs have the great contribution to development of an e - learning system which allows users to quickly post messages on the web for others to use. These messages are available by specific user those belong to the same group. It has impressively become more and more popular in the last year, and it is likely the most well-known platforms to collect the user feedback available on the web. In this paper, we used the learner opinion on all attribute (such as course infrastructure etc.) of the programme and give the recommendation on it. That improves the all attribute of our programme as well as helps the online learning process.

## 3. ISSUES IN EXISTING SYSTEM AND IMPLEMENTATION

The subsisting e-learning, recommendation system fixate on learner interest in their past click compartment on the web, reviews and personalization, past click compartment designates logs data on that substratum they find the cognition homogeneous attribute between learner from antecedent learning predilection. But the learner learning department changes over the time, it is not always true to give the recommendation from the learner's past learning interest. Personalization gives each recommendation to learners. But if the learners are profoundly and astronomically immense the system becomes more intricate. The key issue in this area is how to find learner's interest, need, knowledge level and problems effectively because the instructor has the limited time to interact the learners. Our purposed framework invites the learner to share their opinion independently in all aspects that relate to the programme and we also approached to the instructor then give the rating to each learner according their class performance. Using these approaches, we find the each learner opinion and its knowledge level during the programme. We follow the some step to recommendation.

1. Collect the learner opinion on all entities.
2. Identify the learner's opinion on each entity and cluster them who have the same opinion on any specific entity (using semantic similarity).
3. Cluster same type learner that's had the similar opinion on all entities (using cosine similarity).

4. Find the neighbourhood learner using the semantic similarity and cosine similarity.
5. Cluster the learners according their class rating (this rating is assigned by the instructor, but still we did not have the class rating so assume the learner class rating).

Using above step system gives the recommendation to learners and its neighbourhood learner.

#### 4. PROPOSED METHODOLOGY

Fig [2] shows the overall system architecture of the proposed e-learning recommender system. The system is designed to detect the learner learning preference and issues during the program and support a learner to give a more correct solution of his preferences and issue. the learners issue and preference are depends on their opinion towards the any entities such as on program facility, services, teaching quality, learning materials etc., because they are independent to share their opinion. In this system we performed our experiments on learner one time feedback during the program. The system provides the recommendation only those points which share by the learner in feedback because our system depends upon learner feedback. Purposed system has the three modules first one is clustering module. Second one is the data mining module that performed the data mining technique on learner feedback and clustering technique. Third one is recommendation module that has the expert and encapsulation module.

#### 4.1. Learner clustering on rating wise

Learners' good performance makes the successes of any type learning program. The instructor gives the rating to learners, according his/her performance in the class. Initially the system assumes all learners are in same cluster Fig [2]. This iterative process to find variation in learner class rating during the programme. There are following parameter to determine learners' knowledge level that's depending on the instructor choice. In this system we are considering some parameters that are given below.

- (i) Questions to the class and Paying close attention to who is answering questions during classroom discussion,
- (ii) Asking learners to comment or expand on one another's answers.
- (iii) And withal check the cognitive level in the assignments and projects.
- (iv) and so on....

We cannot show original class rating in this paper due to any restriction, but for finding the similarity between learners at class level. We consider some rating parameter of each learner that discussed above. We performed our experiments on python using the Pearson correlation coefficient [9][10]. To fix the certain threshold value on Score shown in Fig [3] (rating scale 1 to 5).

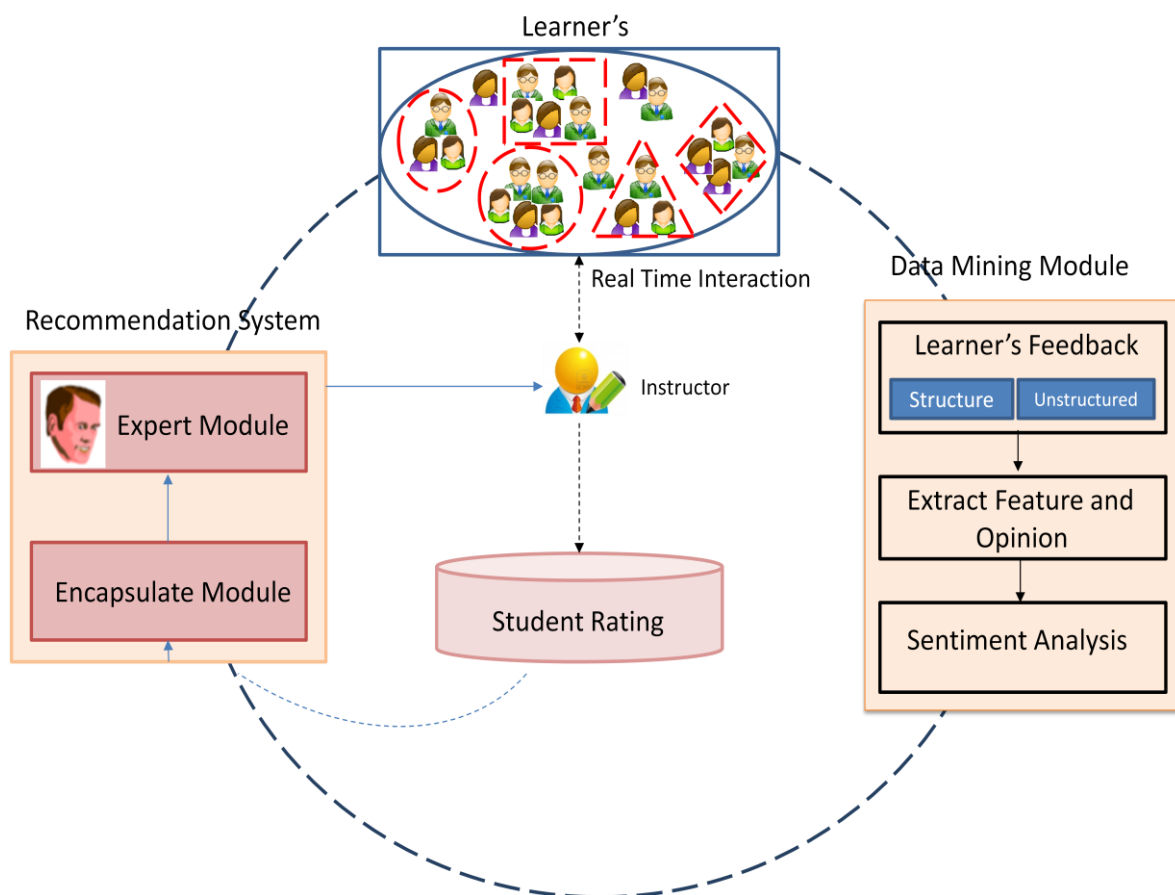


Fig. 2

```

import math
def average(x):
    assert len(x) > 0
    return float(sum(x)) / len(x)
def pearson_def(x,y):
    assert len(x) == len(y)
    n = len(x)
    assert n > 0
    avg_x = average(x)
    avg_y = average(y)
    diffprod = 0
    xdiff2 = 0
    ydiff2 = 0
    for idx in range(n):
        xdiff = x[idx] - avg_x
        ydiff = y[idx] - avg_y
        diffprod += xdiff * ydiff
        xdiff2 += xdiff * xdiff
        ydiff2 += ydiff * ydiff
    return diffprod / math.sqrt(xdiff2 * ydiff2)
S = {"s1": [2,2,5,4,2,2,3,1,5,1,4,1,1,1,1,3,1,1,1],
     "s2": [1,2,1,1,2,2,3,1,1,1,5,1,1,3,1,3,2,1,1],
     .....}
for L1 in S.iterkeys():
    for L2 in S.iterkeys():
        print L1,L2, pearson_def(S[L1],S[L2])
    
```

Stud	Stud	Score
s1	s6	0.39524123
s1	s7	0.60580402
s1	s8	0.8438867
s1	s9	0.67625969
s1	s11	0.70257953
s1	s10	0.31631972
s2	s1	0.401271284
s2	s2	1.0
s2	s3	0.008926634
s2	s4	0.842883165
s2	s5	0.106528743
s2	s6	0.038776407
s2	s7	0.198704835
s2	s8	0.21957936
s2	s9	0.190072951
s2	s11	0.4911957
s2	s10	0.9411264
s3	s1	0.53505518
s3	s2	0.008926634
s3	s3	1.0
.....	.....	.....

Fig. 3

## 4.2. Data mining Module

### 4.2.1 Learner Opinion from structured and unstructured Data

We request the learners to apportion his / her opinion about E-learning program and get prosperously the learner opinion on structured and unstructured data. The opinion poll is divided in structured data on 5-point scale, engenders a more balanced approach because it sanctions learners to have one variable in the neutral position. All entities (question) are predefined shown in table [4.1], the result shown in table [4]. find the learner opinion on some new entity those are missing in structured data and also find the slight difference in learner opinion between structure and unstructured feedback show in table [4][7]. Optically discern the question 1 and 2 opinion poll of learners they accede, vigorously concur of instructor presentation, course material and performance in the class. The question -5 the on synchronous e-learning has the 25 %, 75% opinion to accede, vigorously accede respectively. Another question, opinion is available in Table [4.1]. It is a denouement of successfulness of e-learning. This is a direct opinion about the E-learning program. We have already discussed it has constrained scope in that we got opinion on our point scale we did not get an individual understanding of learners towards e-learning program and what they update? What they are facing the quandary, challenge and acquiring the benefits of E-learning. For making the more precise system, we put the question and get the individual opinion or sentiment in unstructured form towards synchronous E-learning. However, opinions are so paramount in recommendation system that whenever we need to make a decision we opted to aurally perceive others' opinions.

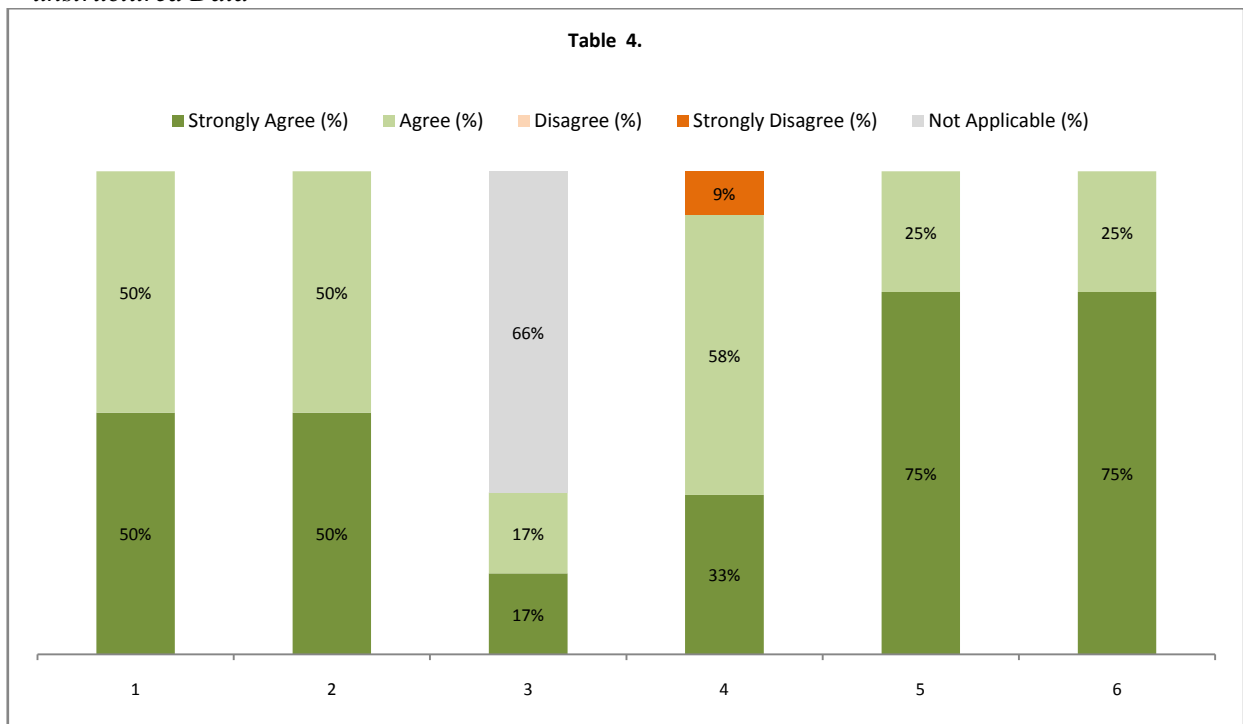


Table 4.1

S.NO	Question
1	The instructor's presentation of the course material is clean and well organized
2	The instructor answers questions effectively and in a timely manner.
3	The instructor's directions and questions for Online examinations are phrased clearly.

4	Our services and video conferencing network quality is good.
5	Synchronous e-learning is helpful to you.
6	Our staff co-operative and helpful

This is not only true for individuals, but additionally true for organizations. The learners express their views and opinion, which is collectively called utilizer engendered content. Extracting the opinion from unstructured text is shown in Fig[5][5.1][5.2][5.3], and categorized in distant group according

to his/her view. We take the learner feedback in both form (structure and unstructured). The main aim behind the learner feedback in unstructured is that to get the more interest of learner about the some important aspect or what they point where they give more impatience or those point are missing in our structured feedback . The result is shown in Fig [7]. we performed our experiments on IBM SPSS[w3] tool. This is Text Analytics for Surveys for retrieving the feature and opinion from learners feedback and also find the semantic similarity between them As a output result, we can obtain a fair understanding of what learners prefer or don't prefer and why. When we understand what learners assume and feel in their own words, we can represent more authentic conclusions about learner's future pattern and utilize that predictive insight to get their requirements more effectively. We can believe on this tool to execute text analysis simpler and more effective for our research because this tool provide the solution on linguistics-based solution particularly this tool created for classifying or coding text reactions and sentiments

**Table. 5. Q.1**

Category	Annotation
Feedback on Instructor	[methodology + like]
	[professor + well-equipped]
	[professor + willing to help]
Feedback on Course Infrastructure	[dvd + available]
	[materials + necessary]
	[network + not interrupted]
	[quality network system + reliable]
	[sound quality + good]
Feedback on Course Timeliness	[semester + fast]
	[time concept + excellent]
Overall Course Feedback	course & [sequentially + delivered]
	[course + delivered]
	[course + sophisticated]
	[course + well-organized]
	[education + available]
	[effort indian government + like]
	[success of the course on indian side + well]

**Table 5.1 Q.2**

Category	Annotation
Feedback on Course Infrastructure	equipment + high-quality
	[assistance + necessary]
	link + not working
	lecture videos + useful
	power + not working
Feedback on Course Timeliness	continuous assessment&include
	lecture notes + timely
	lecture videos + timely
Feedback on Instructor	[assessment + continuous]
	[evaluation of learners + continuous]

Overall Course Feedback	professor + interactive
	professor & knowledge & not vivid
	professor + well-educated
	[course + problem]
	include & senior project
	[special course + possible]
	treatments + thanks

**Table 5.2 Q.3**

Category	Annotation
Self Study	team learning + .
	listening & lecture videos
	Studying
	reading + timely
	teaching learning process + better
	video tutorial from youtube + .
	learning + better
	watch & lecture videos
	reading reference materials
Good Instructors	professor + good
	experience professor + would be good
Practicals/Lab	practical laboratory practices + .

**Table 5.3 Feedback Question**

Question 1- what do you like best about this program?
Question 2- Is there anything else you would like to add?
Question 3- What step could you take to improve your own learning?

In the Table [7] learners' opinion in categorized distinct type of group. It is infeasible for sentiment analysis to ever be 100% precise, but we exert ourselves strenuously get the more preponderant result. 36% learners share opinion about the *instructor* .but in this case all are the positive opinion or sentiments, but in *overall course feedback* ,45% learners share his/her opinion in which 27% have positive, only 9% have negative and neutral opinion . Other results are shown in Table [7] clearly. Those new entities has occurred from the learner opinion(unstructured data) are added to the existing entity (structured data) of the programme. The learners opinion on each entity are shown in Table[4][7] . We used the cosine similarity formula[18] to find the similarity between the

learners according their opinion in all entities. We performed the experiments on python see the Table [7.1]. The learner is

clustered on any certain threshold value. Using this clustering, we find the neighbourhood learner of those learners which have the same opinion on any entity (structure and unstructured data) shown in Table [4][7].

### 4.3. Recommender systems

The system is designed to support a learner to get suitable and relevant learning objects. Our system for the recommendation resources is depending on four modules:

### 4.3.1 Learner rating

this module responsible for System clusters the learner in different group according to their class rating variance. See in Fig [3].



Fig. [6]

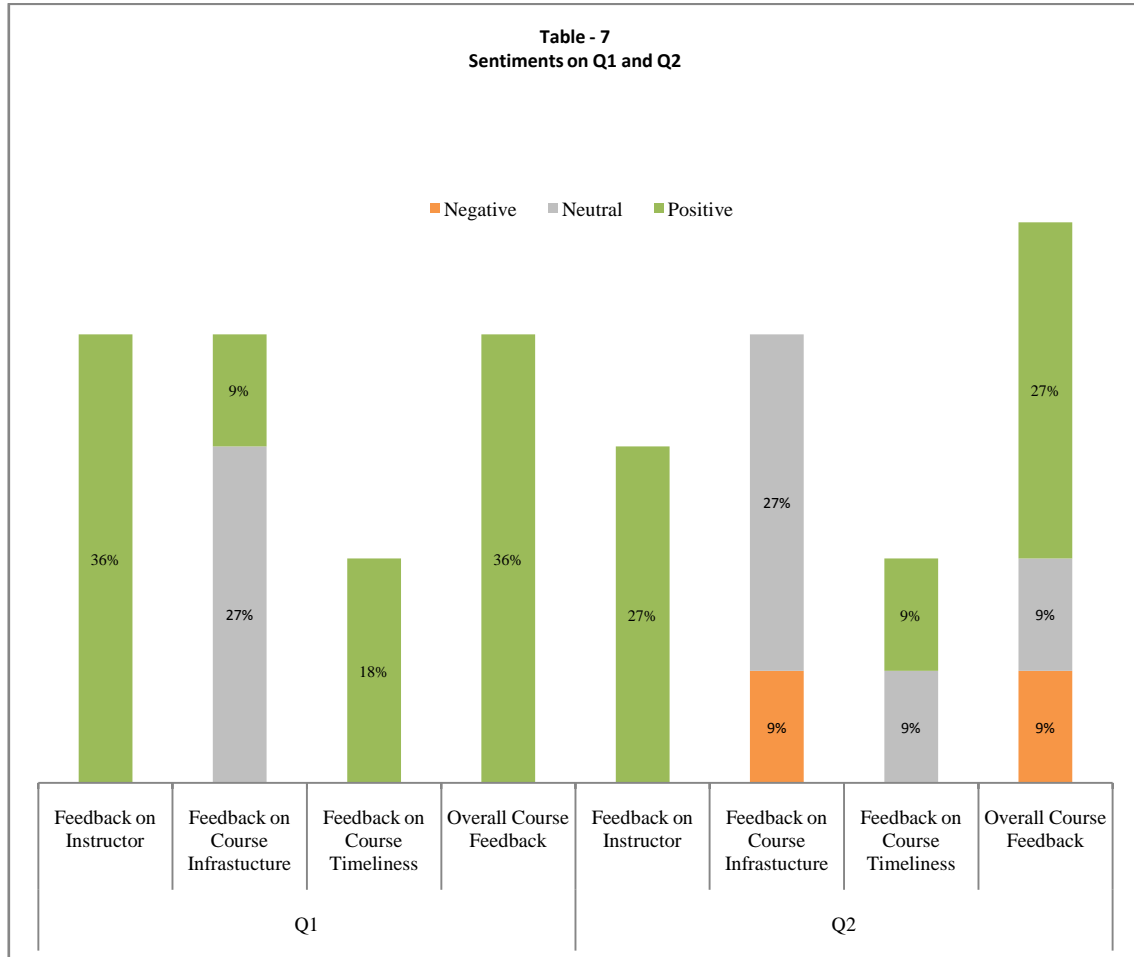


Table 7.1

```

import math

def cosine_similarity(vj,vk):
    sumxx, sumxy, sumyy = 0, 0, 0
    for i in range(len(vj)):
        x = vj[i]; y = vk[i]
        sumxx += x*x
        sumxy += x*y
        sumyy += y*y
    return sumxy/math.sqrt(sumxx*sumyy)

S = {
    "s1": [2,2,0,1,2,2,0,1,0,1,0,0,0,1,0,0,1,1,1],
    "s2": [1,1,0,1,2,2,0,1,0,1,0,1,0,1,1,0,1,1,1],
    "s3": [1,1,0,2,2,2,0,0,0,1,0,0,0,1,0,0,1,1,1],
    .....}

for L1 in S.iterkeys():
    for L2 in S.iterkeys():
        print L1,L2, cosine_similarity(S[L1],S[L2])
    
```

Stud	Stud	Score
s6	s2	0.6676978
s6	s3	0.6676978
s6	s4	0.7001400
s6	s5	0.6288281
s6	s6	1.0
s6	s7	0.4365641
s7	s1	0.5838403
s7	s2	0.4588314
s7	s3	0.3670651
.....	.....	.....

### 4.3.2 Data Mining module

it's the main experimental part of this system in this we take the learners' feedback in structured and unstructured form to identify the learner opinion different entity and also gets the opinion on some new entities during the program See Table [4][7]. in this module we used the clustering on learner in two time first one is used to identify the similarity between learner opinion on any certain entity see Table [4][7] and second one is used to identify the neighbourhood learner of those learner which has same type opinion see Table [7.1].

### 4.3.3 Encapsulation module

See the Fig [3] learners are clustered class rating wise in the same group and mapped with their opinion see Table [4][7]. the system can identify the specific learner or groups of learners' knowledge, requirements and problem. System analysis this data for better recommendation. For example, we assume the learner are categorized in four groups according to their class rating Fig [3]. Ratings group is as sr1, sr2, sr3, sr4 and the learners opinion are mapped with these group ratings.

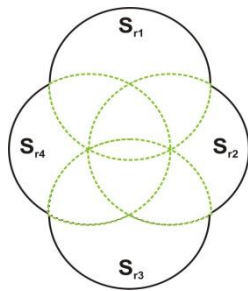


Fig. [8]

From Table [7] 36% learners give the opinion on instructor in this case all are positive, but still we did not know which type of learners have these opinions may be these opinions are given by only sr1 category or sr2 or sr1, sr2, or sr3, sr4 or intersection in between see Fig [8], many combinations can occur. If some learner has the negative opinion on instructor the system identifies the issues and checks that this is happening with the specific group of learner or almost all learners. It retrieves knowledge from learning group and ability to contribute to the recommended.

#### 4.3.4 Expert Module

That pays attention to the specific needs of the learner or learner group, evaluate and diagnostic his problems, and provides the necessary support. Its main role is to get the information of learners from the system and providing appropriate assistance and also inform the instructor about the learners requirements. It works as an interface between learners and the system. During the recommendation the system also gives the recommendation of learner neighbourhood learners. For example See question 3 on Table [5.3], 91% learner share their self study Fig[6]. From those that learner have the lowest rating system can recommend them to follow good learner learning approach. Each module is responsible the recommendation process. System can detect the learners' preference during the program and give the more accurate recommendation. This is the iterative process, learners learning preferences will change over time. They cluster together according their class rating see the Fig [9].

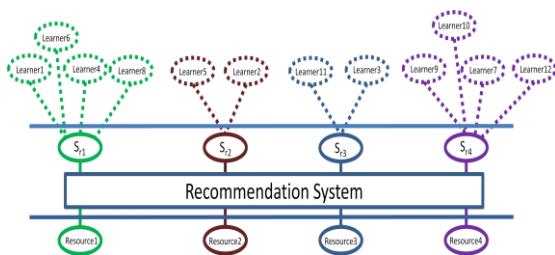


Fig [9]

System must therefore effectively detect changes in learner preferences (dotted line show the variance in rating) and adjust recommendations to suit the learner's current preferences.

## 5. CONCLUSION

In this paper, we purposed e-learning recommender systems framework. Learners share their opinion in structure and unstructured, on programme attribute. From these opinion, we found some new attribute that we are not considering in our existing entities. The clustering technique is used to cluster the learner on their learner knowledge level, cluster the learner on their opinion on programme attribute and cluster those learners who has the same opinion on all attribute. From that we can identify learner neighbourhood learner. During the recommendation the expert provides the same suggestion of targeted learner neighbourhood learners. In this paper we have the feedback data only one course that very small but our

purposed framework scope is not limited we can apply this any large number of learner data. Future work includes strategies to encourage the learners to increase the amount of review share on any social site, in order to make a more robust validation of our heuristic.

## 6. ACKNOWLEDGMENTS

We appreciate the valuable contribution of the team members of the Education technology centre, Indian Institute of technology (IIT) Delhi and Dr. S.K. Naqvi, Additional Director, FTK-Centre for Information Technology, Jamia Millia Islamia (A Central University). We are Very grateful of Addis Ababa University, Ethiopia learners for giving the feedback to this article.

## 7. REFERENCES

- [1] Daoud, M., Naqvi, S. K., & Ahmad, A. (2014). Opinion Observer: Recommendation System on ECommerce Website. *International Journal of Computer Applications*, 105.
- [2] Daoud, M., Naqvi, S. K., & Jha, A. N. Semantic Analysis of Context Aware Recommendation techniques.
- [3] Zhang, Dongsong, et al. "Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness." *Information & Management* 43.1 (2006): 15-27.
- [4] Liu, Tiecheng, and John R. Kender. "Lecture videos for e-learning: Current research and challenges." *Multimedia Software Engineering*, 2004. Proceedings. IEEE Sixth International Symposium on. IEEE, 2004.
- [5] Ricci, Francesco, Lior Rokach, and Bracha Shapira. "Introduction to recommender systems handbook". Springer US, 2011.
- [6] Lu, Jie, et al. "a web-based personalized business partner recommendation system using fuzzy semantic techniques." *Computational Intelligence* 29.1 (2013): 37-69.
- [7] Porcel, Carlos, et al. "A hybrid recommender system for the selective dissemination of research resources in a technology transfer office." *Information Sciences* 184.1 (2012): 1-19.
- [8] Morales-del-Castillo, Jose M., Eduardo Peis, and Enrique Herrera-Viedma. "A filtering and recommender system for e-scholars." *International Journal of Technology Enhanced Learning* 2.3 (2010): 227-240.
- [9] Sarwar, Badrul, et al. "Analysis of recommendation algorithms for e-commerce." *Proceedings of the 2nd ACM conference on Electronic commerce*. ACM, 2000.
- [10] Moldovan, Dan, et al. "Performance issues and error analysis in an open-domain question answering system." *ACM Transactions on Information Systems (TOIS)* 21.2 (2003): 133-154.
- [11] [w1] <http://blogs.thehindu.com/delhi/?p=27504>
- [12] [w2] [http://www.mynews.in/News/ethiopia\\_getting\\_tech\\_aid\\_from\\_iit\\_delhi\\_N25799.html](http://www.mynews.in/News/ethiopia_getting_tech_aid_from_iit_delhi_N25799.html).
- [13] [w3] <http://www-01.ibm.com/software/in/analytics/spss/products/statistics/>