Recommendation of Distance Education Colleges to Students based on Alumni Feedback

Harshini G N, Gobi N, A. Rathinavelu

Abstract*: Sentiment Analysis (SA) systems are very common because most people trust it based on the opinions, emotions, attitudes and feelings shared by the users for decision making purposes about the product, service, news analytics etc. Sentiment analysis or opinion mining is used to automatically detect and classify sentiments into positive, negative or neutral opinion on product or service through certain algorithms. The expeditious growth of internet leads to the increase of reviews about product, services, movies, restaurants or vacation destinations and organizations. In order to increase or decrease the market value of the product, spammers may give the fake ratings. Sentiment Analysis system face great difficulties in deploying the algorithms to classify each review as either honest review, posted by the customers after using the products, or spam review, posted by the individual spammer or spammer groups. Another major challenge faced by the sentiment analysis system is that it lacks the accuracy of predicting implicit and explicit features present in the dataset is low, which is the major challenge in opinion mining system. The proposed system deals with text pre-processing which helps in improving the overall performance of the sentiment analysis systems and an effective system is developed to identify the fake reviews present in the dataset. Association Rule Mining along with K-Means clustering is used to achieve higher efficiency in classification of implicit and explicit features. Lexicon method is used for the classification of sentiments into positive and negative polarities. The advantage of proposed system is that, it can identify and remove the fake reviews in the dataset and extraction of both implicit and explicit feature can be identified through Lexicon based Method along with its polarities.

Keywords: **Association Rule Mining, Fake Reviews, K-Means Clustering, Lexicon Method, Opinion Mining**

I. INTRODUCTION

Sentiment analysis or opinion mining is used to detect the sentiment through positive, negative or neutral opinion about the features of the products or services automatically by relying on certain algorithms. It helps in identifying the emotion of a person from a given piece of text written by them. There will be a huge number of sensitive data in digital formats. Thus, the role of a Sentiment Analysis system is to mine the data and extract the user sentiments is considered as an important task focused by many research communities since last decade. Sentiment analysis or sentiment mining or opinion mining is used to mine the information from various sources of data (reviews) and classify those data based upon the sentiment (polarity) such as positive, negative or neutral. Opinion mining is a new discipline which has recently attracted increased attention within fields such as marketing, Recommendation systems and financial market predictions. Sentiment Analysis systems will be used by the customer for buying the new products among the alternatives, however it is also used by the manufacturer to understand the strength and weakness of their products. Manufacturer can also use the SA system to improve their weakness specified by the consumers through reviews. Sentiment classification, feature level identification, opinion summarization are three important classification of sentimental analysis. The three types of opinion mining are document level, sentence level, aspect level.

- Document Level: It is used to obtain overall sentiment value for the entire document.
- Sentence Level: In sentence level, each sentence of the reviews is processed separately to determine the polarity of the review along with its aspects.
- Phrase Level / Aspect Level: The major task in aspect level sentiment analysis is to find the individual features and its opinion.

Aspect Level SA system classifies the reviews based upon the opinion orientation of the aspects namely Explicit and Implicit Reviews. If the reviewer talks about the positive, negative or neutral aspects of the products directly then these kinds of reviews are termed as explicit reviews. For example, the course is good and the lab sessions are knowledgeable. The reviewer directly express about the course. In Implicit Review the reviewer may talk about the aspects of the products, indirectly. Example for Implicit Review is, sometimes he went through the concepts too fast for us to grasp. The reviewer here implicitly pointed negative opinion about the instructor. There are two kinds of opinions with respect to comparison among the products namely direct and comparative opinions Direct opinion gives an opinion about the product features. For example, Excellent introduction to 3D printing. Here the reviews directly provide information about the 3D printing course with positive sentiment.

Revised Manuscript Received on May 21, 2020.

Harshini G N, Department of Computer Science and Engineering, Dr. Mahalingam College of Engineering and Technology, Pollachi, Tamilnadu, India, E-mail: harshini0419@mail.com

Gobi N, Department of Computer Science and Engineering, Dr. Mahalingam College of Engineering and Technology, Pollachi, Tamilnadu, India, Email: gobi@drmcet.ac.in

Dr.A.Rathinavelu, Department of Computer Science and Engineering, Dr. Mahalingam college of engineering and technology, Pollachi, Tamilnadu, India, E-mail: starvce@mcet.in.

In comparative opinions, the opinion is expressed by comparing one product or service with another product. For example, the concept of accounting analytics is easier to understand than 3D printing. Here, the reviewer compares the two courses and suggested accounting analytics is easier to understand. For providing better ranking among the alternative products with respect to aspects, comparative opinion plays a vital role.

II. RELATED STUDY

Sentiment Analysis concepts became common when after the year 2000, the user received reviews of the goods they buy through digital format rather than written format. Before the year 2000, the consumer receives nearly a lot of feedback through written format by filling out a feedback form distributed by either the producer or the third party seller.

The term “Sentiment Analysis” comes into existence with great impact during the year 2003 by J. Yi e. al (2003). Yet Sentiment Analysis is used specifically to evaluate the target

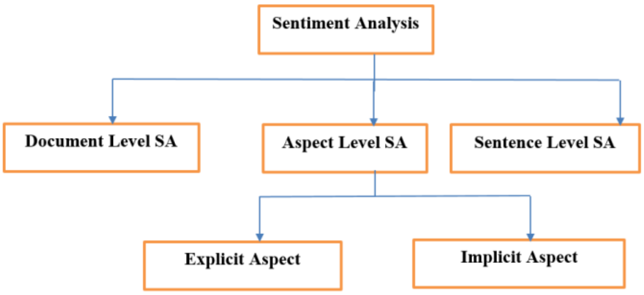

Fig.1.Types of Sentimental Analysis

of opinion (attribute / feature) through the word of opinion (polarity or sentiment) shared by individuals on specific topics of interest.

The need for Sentiment Analysis is to infer the decision-making process because of the large amount of subjective texts accessible over the internet. To handle this enormous subjective data to recent years, several researchers have introduced different approaches to sentiment analysis with the goal of extracting useful information from the evaluation and then communicating it to the consumer for the purpose of making decisions.

Most of the Sentiment Analysis approaches are classified into three levels and they are Document level SA, Sentence level SA, Feature level SA. Document Level Sentiment Analysis is intended to obtain overall sentiment value for the whole text, Turney. P.D. (2002) applied Point Wise Mutual Knowledge to measure an average semantic orientation score of the extracted phases for the description of the orientation of the documents. DLSA's implementation effort requirements are more than two other approaches (Aspect Level SA and Sentence Level SA), because it has to process an entire document at a single trace and evaluate the product polarity based on the individual aspects. Due to the impact of sentence level sentiment analysis, the scale of the document level sentiment analysis is decreasing.

Each sentence of the reviews is separately processed during the Sentence Level Sentiment Analysis to assess the review's polarity along with its aspects. It consumes maximum time during its cycle, but the accuracy of the system's prediction is much better than the Document Level Sentiment Analysis. While Sentence Level Sentiment Analysis performs well for the SA system, the researchers primarily aim to discover a reviewer's feeling of reflecting the overall opinion, but not to decide the features, the customer wants to analyze exactly.

The key challenge in the analysis of aspect level sentiment is recognizing the individual characteristics and their opinion. The Aspect Level Sentiment Analysis has evolved over the past decade into two major branches, model-based method and the statistical approach. Feature Identification or Extraction, Sentiment Prediction are the two main tasks in Feature based Sentiment Mining. During Feature Identification task, the task is to identify features of the reviewed item and to group synonyms of features. The task of predicting sentiments is to decide whether the feeling about the given aspect (feature) reflects the positive, negative or neutral opinion based on the user's comments.

Sentiment Analysis concepts have become popular when the user provided their reviews about the products that they purchase through digital format rather than written format after the year 2000. Before the year 2000, almost many feedbacks are received from the user through written format by filling up a feedback form circulated by either manufacturer or the third-party seller. But Sentiment Analysis are mainly used to determine the opinion target (attribute/feature) and opinion word (polarity or sentiment) expressed by individual about particular topics of interest. Due to large amount of subjective texts available over the internet arise the demand for Sentiment Analysis in order to conclude the decision making process [1]. To handle this vast subjective data during recent years, many researches have proposal various sentiment analysis approaches with an objective of discovering useful information from review and then distribute it to the customer for their decision making purpose.

For the sentiment classification, Stanford NLP library is used which classifies the positive and negative percentage of reviews. POS Tagging is done which extracts the nouns and adverbs from the dataset [2]. POS Tagging is used to identify the structure of the word.

To identify the spam reviews, Group spam behavior indicators and individual spam behaviour are used as an indicator [3]. Some of the indicators used to identify the group spam reviews are Group time window, Group Deviation, Group Content Similarity and Group Size Features. A novel relational model is proposed which identifies fake opinions in the dataset.

Burst pattern detection serves as a fine grained method for capturing intense spamming activity. Spammers aim to strongly influence public opinion and unsuspected users about a particular service or product towards a particular direction, i.e., positive or negative, depending on their end goal [4]. Burst detection has proven to be a valuable resource in spam detection,

An efficient K-means with Association Rule Mining algorithm is developed, to classify both implicit and explicit features [5]. It also identifies whether opinion words present has any matching words inside any of the k clusters. If a match found with any of these clusters, then it determines the implicit feature for its corresponding opinion word otherwise it fires the ARM to predict the implicit feature.

By using co-occurrence association rule mining separate rules for both implicit and explicit features are identified. Opinion words and features are extracted from the explicit features. A co-occurrence matrix is useful in finding the frequency of opinion words [6].

To identify fake review in electronic domain, a system is developed which focuses on user centric and review centric features [7]. These centric features are useful in the identification of fake features in the dataset.

The behavior features of the spammer is considered to detect the spam reviews [8]. Algorithms are designed to implement the formulas used in finding the spam reviews. To identify the similarity between two reviews the algorithms are used. The algorithm achieves higher efficiency than traditional algorithm.For the cross domain sentiment classification of reviews, a new technique is used which consists of Apriori algorithm [9]. For the identification of explicit features from the dataset, an automated system is developed. Data preprocessing, explicit suggestion extraction and visualization is done [10].

III. SYSTEM ARCHITECTURE

The dataset is obtained from [12] kaggle. The dataset consists of input text which expresses the opinions of the students about the courses and teachers. To reduce the size of the dataset, Pre-processing is done, which includes stop word removal, tokenization and parts of speech tagging. From the pre-processed data, Features are extracted using Association Rule Mining. For the sentiment analysis, Lexicon method is used which classifies the sentiments into positive and negative The collected dataset may include irrelevant and fake reviews which is removed by Fake Review Indicators. To identify both implicit and explicit reviews Association Rule Mining with K-Means clustering is used. The overall process of proposed system is shown in figure 2.

IV. PROPOSED SYSTEM

To eliminate the unnecessary data present in the dataset, Preprocessing is done. Preprocessing phase includes Stopword removal, tokenization and POS tagging. Stopwords includes unwanted text like articles, prepositions present in the review sentences.

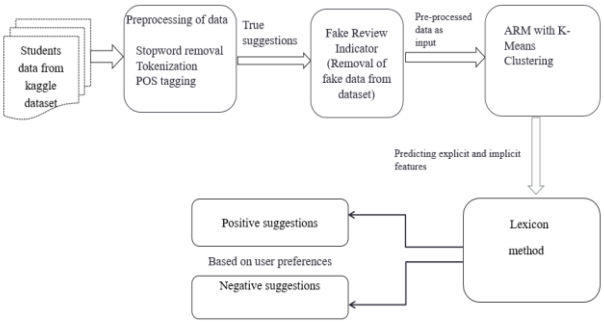

Figure.2. Architecture of Proposed System

Tokenization is a process in which collection of sentences in a dataset are divided into token by removing white space, comma and other symbols etc. As result of Tokenization and Stop Word Removal process, the system has filtered the features which are tokenized into record size chunks like sentences and boundaries are identified for each sentence by marking the special characters or delimiters such as # or % symbols.

The documents are parsed using Stanford parser through which system assign Parts-Of-Speech (POS) tags to every token or word present in the filtered sentence. POS tags are useful to identify the grammatical structure of sentences such as noun, verb, adverb and adjective phrase and their relationship as shown in the Table 1. PoS is widely used in the analyzing the sentences for detecting emotion. Some research findings show that adjective serves as a good indicator to detect the features present in the reviews.

Table- I: POS Tags with Examples

Tag	Description	Example(s)
CC	Coordinating conjunction	And, but,or
CD	Cardinal number	One, two, three
DT	Determiner	A, the
EX	Existential there	There
FW	Foreign word	mea
IN	Preposition or subordinating conjunction	Of, by, in
JJ	Adjective	blue
JJR	Adjective, comparative	greater
JJS	Adjective, superlative	greatest
LS	List item marker	1,I,i
MD	Modal	Can, may
NN	Noun, singular or mass	IPhone
NNS	Noun, plural	IPhones
NNP	Proper noun, singular	IBM
NNPS	Proper noun, plural	carolinas
PDT	Predeterminer	All, both
POS	Possessive ending	‘s
PRP	Personal pronoun	He, she
PRP$	Possessive pronoun	your
RB	Adverb	quickly
RBR	Adverb, comparative	faster
RBS	Adverb, superlative	fastest
RP	Particle	Off, up
SYM	Symbol	+,@,%
TO	To	to
UH	Interjection	Oops, oh ho
VB	Verb, base form	Eat
VBD	Verb, past tense	Ate
VBG	Verb, gerund or present participle	Eating
VBN	Verb, past participle	Eaten
VBP	Verb, non-3rd person singular present	Eat
VBZ	Verb, 3rd person singular present	Eats
WDT	Wh-determiner	Which
WP	Wh-pronoun	What, who

V. ASSOCIATION RULE MINING

It is used for mining frequent item sets for Boolean association rules. It can be used for larger item set property and its implementation is an easy method. Apriori is mainly designed to operate on database which contains any transactions between the cross-domain sentiment classifications. The steps involved in the Apriori algorithm are as follows,

Step 1 :
Create a frequency table of all the items that occurs for transactions.
Step 2:
Find the frequency of the items satisfy the min-support, which is support must be greater than or equal to the threshold.
Step 3:
Take the first item and pair with all other items based on support value.
Step 4:
Similarly, find the frequency for second item set based on min-support and first item Set.
Step 5:
Obtain the string association rules based on the min-confidence value.
Step 6:
Now count the occurrences of each pair in all the item set. Finally, the meaningful rules are learned and generated by using Apriori algorithm.

Classification of Implicit and explicit features by Association Rule Mining is moderate, so in proposed system Association Rule Mining with K-Means Clustering is used for classification of reviews.

VI. FAKE REVIEW IDENTIFICATION

The use of inappropriate and illegal methods to generate large amounts of fake reviews against their target product in order to enhance or demote product is known as Opinion Spamming. These are known as fake or spam reviews and the author who is writing the fake review is known as fake or spam reviewer or spammer. Spam reviews and spammers main impact are:

- Increase in volume of spam may disappoint students.
- Students can be deceived in respect to the grade of the college.

Sentiment analysis system are trusted by most of the people for making decision about the product, service, news analytics etc. based on the opinions, emotion, attitudes and feelings expressed by the users. The researchers used many metrics to effectively identify spam and merged with a lot of steps and techniques, providing a major solution to the spam analysis problems. A score is calculated by means of various metrics through quantitative measures such as review author and review history score assigned to each review under investigation to distinguish false from genuine reviews. There are three major ways to detect the spam reviews present in the review data and they are as follows:

- Spam review detection
- Spam reviewer detection
- Detection of spammer groups.

The proposed work focuses on detecting Spam reviews and the indicators used in identifying spam review are as follows:

A. Review Relevancy Rate

Reviewers may also post advertisement or link in the reviews which is irrelevant and dispensable. To identify and remove the unwanted reviews Review Relevancy Rate is used. Review Relevance Rate refers to the degree of relevance that exists between the content of the review and the subject of the product. The formula used in finding review relevancy rate is

$$RRR(r)=e^{((W(s))\cap W(r)/W(s))}-1 \qquad (1)$$

where W(s) is the set of all segmented words of the product's topic, and W(r) is the set of all segmented words of a review.

B. Content Length

One of the useful indicator used to identify the fake reviews is the review content length. If the review content is short, it shows that the reviewer did not consider it seriously.

$$contentlength = \begin{cases} 1, r.length <= \lambda \\ 0, r.length > \lambda \end{cases} \qquad (2)$$

where r. length denotes the length of the review r, and λ is a threshold to judge the effectiveness of the length of the review content. Based on previous studies where there are less than 6 words in an English review, the system should consider these kinds of reviews as a spam review since the reviewer is less serious about the product.

C. Review content similarity:

One of the most effective strategies to identify the fake in online reviews by analyzing the similarities between reviews by authors. Due to time constraints, spammers prefer to replicate the same written content for multiple products in their reviews, sometimes more advanced spammers, who are actually trying to change the content, seem to use a similar vocabulary each time.

As a consequence, examining the content similarities between the reviews of the author will help determine the validity of the reviewer's intentions to create false reviews. Below is a detailed description of the methodology used in this approach for similarity detection of content.

Content is represented by using the cosine similarity between review text documents. The system inputs a list of (text) reviews and builds the Word Bag model by collecting from them all specific terms (words). Then for each document, a vector is generated which sets the list of documents into a collection of vectors in a model of vector space. In the vector of a text referred to in Figure 4.4, each expression in the Bag of Word sets is represented with a respective meaning.

Term frequency measures how many times a given term j appears in a document i. To treat documents of different sizes, such as two separate reports, the measured frequency is divided by the document's length to standardize.

Using the Eq 3, their cosine similarity can be extracted after constructing the vectors describing the analysis text documents, using a Euclidean dot product formula.

$$a.b = |a||b|cos\theta \qquad (3)$$

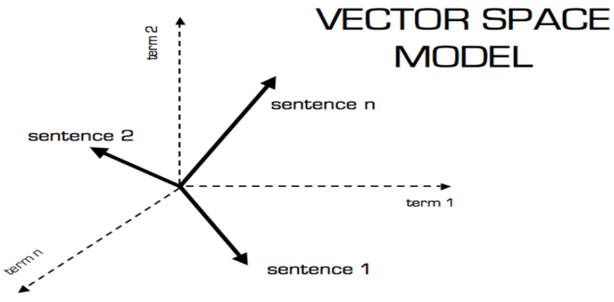

Figure.3. Vector Space Model

A very high similarity score suggests suspicious behaviour, and the user who uses the same content several times is considered to be spam.

D. Calculating Spam Score:

From the combination of several discrete scores related to the review and reputation, reviews overall Spamicity is calculated. The spamicity score can be measured through function *S (r)* and it calculates the scores derived from the Review level characteristics and features. To determine the presence of spamming activity different techniques are used and their resulting scores can contribute to measure the overall level of spam in a review but not all methods and review functions have the same weight to determine the presence of spamming activity. According to its importance, each score is multiplied by an appropriate weight, as represented in Table II.

Table- II: Spamicity Factors

Spamicity affecting factors	Purpose	Weight
CS(a)	Content similarity	1.5
RRR(r)	Review Relevancy Rate	0.25
CL(r)	Content Length	0.25

After applying the above weights and score for the various factors, the Review Spamicity S(r) score is calculated by using the equation 4

$$S(r) = 1.5CS(a) + 0.25RRR(r) + 0.25CL(r) \qquad (4)$$

If a review's overall spam score exceeds a threshold, then a review is considered to be spam. A manual verification of the various scores and their range of values led the system to consider a threshold of 7. Finally, the review can be considered as spam if its score exceeds a threshold value of 7 and it is represented using

Review can be considered as spam or honest = $\begin{cases} Spam > 7 \\ Honest <= 7 \end{cases}$

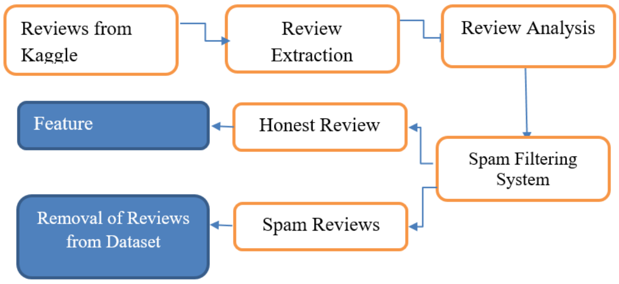

Figure.4. Spam Review Identification

VII. ASSOCIATION RULE MINING WITH K-MEANS CLUSTERING

Association Rule Mining with K-Means clustering is a clustering technique based on prototypes. It is used to identify characteristics that have no predefined class labels but group characteristics using the similarity measures between them. This positions most similar characteristics in one class, and dissimilar features in another. The aim of clustering is to assign each data point to a cluster based on their Euclidean distance. K-Means is used in sentiment analysis process to group the features into groups based on their high similarity. The clustering is based on three observations which follows:

Step 1:

Initially it considers similarity of opinion terms that is useful to direct the clustering.

Step 2:

Identifies the features similar to those in the reviews. It is used for identification of the aspect with same meaning.

Step 3:

It considers the structure of the feature in the comment. POS Tagging of each word is identified. The algorithm automatically constructs the model based on the inputs given which can cluster a similar class of objects to predict the value of the missing attribute.

Initially to create a cluster, a collection of words of opinion are put inside the tuple, where each tuple has a set of words of opinion. K- Means clustering is used to test whether the words of opinion in the sample have any of the words in the clusters that fit. If a match is found, it will evaluate the implicit function of its corresponding opinion word otherwise it will fire the ARM to predict the implied feature. For each tuple a set of association rules are shot, and the mean confidence score for each set of rules is determined. The finally selected words of opinion and the implicit function will be stored for future references in the respective cluster. For the proposed system, the value of K is 4 which yields better results.

VIII. LEXICON METHOD

The lexicon-based approach includes classification of sentiments into positive and negative words. The approach in classifying the text involves constructing classifiers from classified instances of texts or sentences which are basically a supervised process of classification. Lexicon-based approach dictionaries can be generated manually or automatically, using seed words to expand the list of words Many of the lexicon-based research has centered on the use of adjectives as indicators of the semantic orientation of text. In proposed system predefined dictionary which calculates the semantic score for each word is used for the sentiment classification of words. If the score is in positive, it is classified as positive sentence and if the word contains negative score then it is classified as negative Sentence. Lexicon based system are widely used to identify the sentiments sure for each attribute with good accuracy.

IX. RESULTS AND DISCUSSIONS

The expected results are discussed in this section. The evaluation of the proposed system and the association rule mining with K-means clustering is done by considering the major accuracy measures like error rate, precision, recall and F-Measure.

A. Data description:

The dataset is obtained from kaggle. The dataset includes the students reviews about the class, course and staff in distance education system. The dataset contains the student's course Id, course name and reviews. The total reviews present in the kaggle is 12,73, 202.After preprocessing and fake reviews removal, the size is reduced to 10,33,731.

B. Evaluation Metrics:

Several factors are used in identifying accuracy of the proposed system. Though experiments, a comparison is made between the association rule mining algorithm and association rule mining with K-Means clustering. The experimental results show that the proposed system with Association Rule Mining and K-Means clustering performs better in identifying the implicit and explicit features.

Table- IIi: Evaluation Matrics

Factors	Description
True Positive	Number of correct suggestions which are extracted.
False Positive	Number of irrelevant suggestions which are extracted.
True Negative	Number of correct suggestions which are not extracted.
False Negative	Number of irrelevant suggestions which are not extracted.

Precision: Precision is the ratio of number of correctly extracted suggestions to the sum of correctly extracted suggestions and wrongly extracted suggestions. It can be computed as

$$Precision = TP \div (TP + FP) \qquad (5)$$

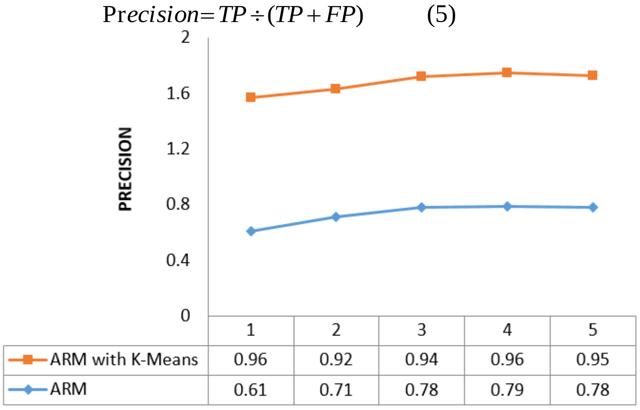

Fig.4. Evaluated Precision

Recall: Recall is the ratio of number of irrelevant suggestions extracted to the sum of irrelevant suggestions and relevant suggestions. Recall can be computed as

$$Recall = TP \div (TP + FN) \qquad (6)$$

F-Measure: The harmonic mean of precision and recall is known as F-Measure. It is evaluated by using

$$F - Measure = 2 * (Precision * Recall) / (Precision + Recall) \qquad (7)$$

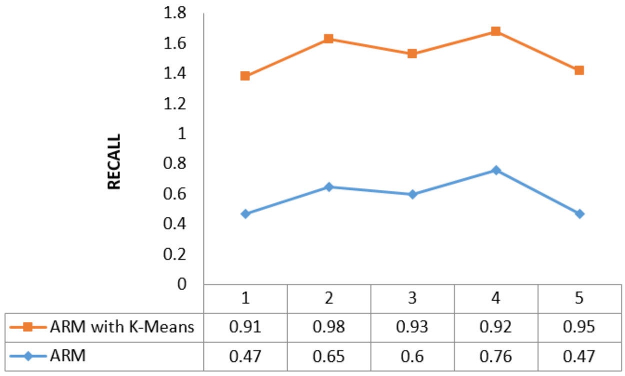

Fig.5.Evaluated Recall

The f score of the proposed system has better values and it can be inferred from the figure 5. that the mean of Precision and Recall values are higher.

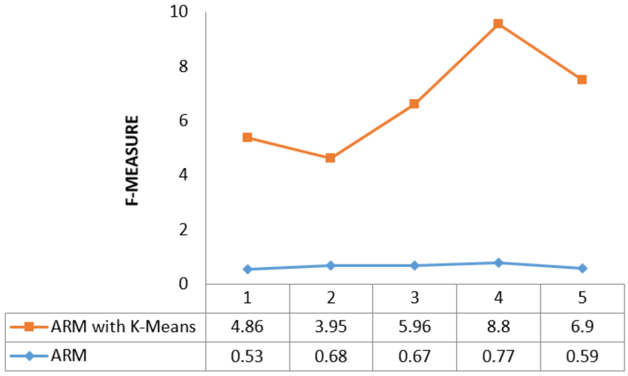

Fig.6. Evaluated F-Measure

X. CONCLUSION

In this paper, Fake review indicators are used to find and remove the fake reviews present in the dataset. Association Rule Mining with K-Means clustering is used in the identification of Implicit and explicit reviews. The systems performance was evaluated using various measures like Precision, Recall and F-Measure are found to have better values than the previous Association Rule Mining. Lexicon based method is used to identify the sentiment sure against each feature for the better decision making by the new students.

REFERENCES

1. Sujata Rani and Parteek Kumar (2017) “A Sentiment Analysis System to Improve Teaching and Learning”, IEEE, vol. 50, issue 5, pp.36-43.
2. Luxchippiriya Balachandran, Abarnah Kirupananda (2017) “Online reviews evaluation system for higher education institution: An aspect based sentiment analysis tool”, International Conference on Software, Knowledge, Information management and applications, pp.1-7.
3. Arun Mukherjee, Bing Liu, Natalie Glance (2012) “Spotting fake review groups in consumer reviews”,21 st International Conference on World Wide Web, pp.191-200.

4. Ioannis Dematis, Eirini Karapistoli, Athena Vakali (2017) "Fake Review Detection via Exploitation of Spam Indicators and Reviewer Behavior Characteristics", International conference on current trends in theory and practice of informatics, vol.10706, pp.581-595.
5. Gang Liu, Wray Buntine, Weiping Fu, Yudan Du (2015) "An association rules text mining algorithm fusion with K-Means improvement", International Conference on Computer Science and Network Technology, pp.781-785.
6. Zhen Hai, Kuiyu Chang, Jung-jae Kim (2011) "Implicit Feature Identification via Co-occurrence Association Rule Mining", International conference on Computational Linguistics and Intelligent Text Processing, vol.6608, pp.393-404.
7. R Barbado, O Araque, CA Iglesias (2019) "A framework for identifying fake review detection in online consumer electronics retailers", Information Processing & Management, vol.56, issue.4, pp.1234-1244.
8. Neha S. Chowdhary, Anala A. Pandit (2018) "Fake Review Detection using Classification", International Journal of Computer Applications, vol. 180, issue .50, pp.0975-8887.
9. Xi-bin Jia, Ya Jin, Ning Li, Xing Su, Barry Cardiff, Bir Bhanu, (2018), "Words alignment based on association rules for cross-domain sentiment classification", Springer (Frontiers of Information Technology & Electronic Engineering), Vol.19, Issue 2, pp 260–272.
10. Sujata Rani and Parteek Kumar (2017) "A Sentiment Analysis System to Improve Teaching and Learning", IEEE, vol. 50, issue 5, pp.36 - 43.Jenifer Jothi Mary, S. Santiago Dr. L. Arockiam (2018) "A Methodological Framework to Identify the Students Opinion using Aspect based Sentiment Analysis", International Journal of Engineering Research & Technology vol.5, issue.02, pp.642-645.
11. Vasileios Kagklis, Anthi Karatrantou, Maria Tantoula, Chris T. Panagiotakopoulos, Vassilios S. Verykios (2015) "A learning analytics methodology for detecting sentiment in student fora: a case study in distance education", European Journal of Open, Distance and e-Learning, vol. 18, no.2, pp.74-94.
12. Xinyue Wang, Xianguo Zhang, Chengzhi Jiang, Haihang Liu (2018)"Identification of Fake Reviews Using Semantic and Behavioral Feature International Conference on Information Management(IEEE), pp.92-97, DOI:10.1109/INFOMAN.2018.8392816.
13. K. C. Ravi Kumar, D. Teja Santosh, B. Vishnu Vardhan (2017) "Extracting Opinion Targets from Product Reviews using Comprehensive Feature Extraction Model in Opinion Mining", Indian Journal of Science and Technology, vol .10(21), pp.1-6.
14. Z. Kamisli Ozturk, Z.İ. Erzurum Cicek and Z. Ergul (2017) "Sentiment Analysis: an Application to Anadolu University", International Conference on Computational and Experimental Science and Engineering, vol.132, issue.03, pp.752-755.
15. Swapna Gottipati, Venky Shankararaman and Jeff Rongsheng Lin (2018), Text analytics approach to extract course improvement suggestions from students Feedback", Springer (Research and Practice in Technology Enhanced Learning) vol.13, issue.6, pp.1-19.

AUTHORS PROFILE

Harshini G N, PG Scholar, Department of Computer Science and Engineering, Dr. Mahalingam College of Engineering and Technology(MCET), Pollachi. She completed her Bachelor of Engineering from Anna university. She is doing her project in Opinion Mining.

Gobi N, Assistant Professor, Department of Computer Science and Engineering, Dr. Mahalingam College of Engineering and Technology. He has teaching experience of more than twelve years. His current areas of research are Opinion Mining and Cloud Computing

Dr.A. Rathinavelu, Principal, Dr. Mahalingam College of Engineering and Technology(MCET), obtained his Ph.D. from NIT, Trichy. He completed his Master of Technology in Edith Cowan University, Perth, Western Australia and Bachelor of Engineering at IRTT, Erode. He has 6 years of Industrial experience in Engineering college. He is working in MCET since 1999 in various capacities. He is a Professor in the department of Computer Science and Engineering. His research area includes Human Computer Interaction, Web Engineering, Speech Processing and Mobile Application Development.