



INTELLIGENT DATA MINING OF FEEDBACK FROM STUDENTS FOR IMPROVING COLLEGE INFRA

P. Asha, M. Jahnavi and Dega Vasanthi

Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, India

E-Mail: ashapandian225@gmail.com

ABSTRACT

Extracting data intelligently from the social networks has attracted every field with greater interest. All education communities also used social media data extraction to simultaneously improve infra outcomes using student generated sentiments. With the excessive growth of social network (i.e., reviews, forum discussions, social networks) on web, individuals and organizations are mostly using public views and opinions for their decision making. Potential users also want to know the views and opinions of previous users before they purchase a product or use a service. These online social media techniques are providing greater opportunities for users to discuss about their experience and share opinions. So we preferred these techniques for analyzing the feedback given by students for improving college infra and hostel related reviews for hostel improvement and to come up with a decision where the development is in need. This project helps to increase the improved quality of education system. With the help of this, we can open the door for the education system to optimize service and quality in college infra. The proposed system gives opinion related to college info and improvement of student requirements.

Keywords: graph based method, self-arranging fluffy neural.

INTRODUCTION

Education gained the great importance from the past few decades. For the great achievements of the education system, we should update the views and opinions of students with the help of getting feedback from students regularly. Universities may follow the best methodologies but still face the problem to satisfy all the needs of students even though they receive the feedback from students. This is because the institution cannot come to a conclusion to decide which part of the curriculum need to be improved. If universities or colleges can identify the reasons for the low performance of college they will be able to improve the college infra by getting the accurate result from feedback and this helps them to take proactive actions. Management can come up with better ideas and strategies to improve the performance. Eventually this will help in improving college infra where students get satisfied with their needs.

Analyzing the problem and predicting it can be done using mining techniques which gives good outcomes in the existing systems. In this project we are using natural language processing algorithm for the separation of positive and negative feedback given by students. And also we use decision making algorithm for automatic review given to a college database based on the feedback. Here in our project, we also use Graph based method for the easy understand of the feedback.

LITERATURE SURVEY

Sanjiv *et al.* [1] has developed a method from stock message boards to extract small investor sentiment. They used 5 various algorithms to perform well analysis of human sentiments. To improve the quality of the sentiment cross-sectional aggregation and time series methods are used. Also they used pattern recognition, statistical methods and phase log analysis. It is an important and collection of sentiment in stocks results

better than the normal stocks. Hence all these algorithms results in developing the third party news, management announcements and press releases.

To describe datasets they were classified as a sequence collection of correlated objects. Interrelations represented as homogenous networks with a single link type and object type or heterogeneous networks with multiple link and object types. Some of the examples of homogenous networks are social networks or WWW and few examples for heterogeneous networks are domains on medical for describing patient diseases, contacts and treatments or in other academics domains. A technique called as link mining is used for building models of raw data. Group detection, collective classification, object ranking, sub-graph discovery and link prediction are the majorly solved tasks in link mining. Likewise in network analysis, it has been used in particular areas such as hypertext data mining, social network and web analysis. It was an emerging technology and also mentioned about other emerging techniques [2].

Extraction of knowledge from the social media brought some interest from a medical database to improve healthcare and also decrease costs. They proposed a two different step analysis framework which mainly concentrates on positive as well as negative opinions and forum posts on cancer treatment. SOM is used to analyze the frequency data from reviews or forum posts. Later, they introduced a network based approach for modeling users and forum interactions. This helped to identify influential user opinion. This approach expanded research into various fields such as hospitals, social media, pharmaceutical industry or for future treatments [3]. Newman *et al.* [4] showed some interest in algorithms which are used for finding the communities in network-in the groups, some connections are dense, but between some connections are sparser. Based on reviews, the progress is shown. Here they started describing traditional methods of



community detection that are the Kernighan–Lin algorithm, hierarchical clustering and spectral bisection which are based upon the similarities. However, no method is similar to another method, which existing research is connected, like web data, the internet, social and biological networks. They also described so many algorithms which do well with related data, including algorithms based on short loops count, on voltage variance and on edge between scores networks.

Erwan *et al.* [5] believed that based on the original system's information, the real systems should be represented as networks which consist of more information. In the previous community detection, they act as a active field of research and receives attention. Currently stability is introduced as a measure which is used for quality of partition. In this way, the work decides an advancement equation that uses a Markov procedure perspective of systems to permit multi-scale group location. Numerous heuristics and contrast of a calculation, the streamlining steadiness are displayed as covering groups. Looks into demonstrate that the technique indicates exact multi-scale arrange examination. Now a day's education is an important part of our life and plays a major role in development of our nation. Education institutions are the foundation of education for an individual. Also another important thing is to predict the student's performance in education. Educational performance depends on different factors like personal, psychological and social activities. In order to measure this performance a mining tool is used. In this paper, the data of a student from an institution is considering different mining techniques which are used to predict the student's performance based on current data. This study helps us to understand the data mining techniques well [6]. Creators recognized that real issues for the advanced education group are examined and assess it from the understudy's criticism for building up the instruction framework. The standard techniques for investigation and assessment are not bolstered to look through the data which is escaped the understudy criticism sets. In this they demonstrated the examination of input utilizing k-implies calculation for good and viable basic leadership by group. Thus these are in charge of observing and getting the enhanced rendition of both educating and learning procedures and encounters with understudies [7].

Evaluation is making the judgment about the value of something. This acts as a performance measure and effects on decision making of individuals. This paper gives information how to evaluate a person considering various factors for the evaluation of performance. The main objective of this project is to predict productivity, potential & quality of the faculty by evaluating different factors which will allow higher level officials to know different types of faculty and to take decisions for differentiating each faculty based on their performance. The analysis depends on various different factors such as analyzing the student's feedback, institutional feedback, financial support, research and satisfaction. The data mining technique is used for collecting information from the database which helps in extracting faculty performance

when evaluated across several factors [8]. RuiMin *et al.* [9] says the importance of web-based distance learning education system of assessment system which is characterized based on adaptability and positive performance. For an integrated education system, the evaluation system collaborate with other subsystems while teaching and learning process, this helps in the improvement of the educational system with its performance.

In this modern education system, the task is on to find out different ways to make more efficient and effective. The data is collected from database in large amount, but that is not used. To get the efficient data we need to use more powerful tools. For this we use a powerful tool called Data Mining for analyzing and predicting. This can be applied in marketing, advertising, fraud detection, prediction and loan assessment. The applicable amount of work is done, but there are many fields which are not used and unified approach is also not used in these researches [10].

Another crossover learning calculation depends on hereditary calculation to plan a fluffy neural system, named as self-arranging fluffy neural system on hereditary algorithm(SOFNNGA), to apply Takagi-Surgeon(TS) type fluffy models is proposed. Another technique is included based geometric developing model and the E-culmination of fluffy tenets, which is utilized to produce the beginning structure. This calculation depends on hereditary calculations, back proliferation and parameters [11]. It consists of two steps: i) Firstly, the linear parameter can be adjusted accordingly, and ii) The centers and widths are modified for all the membership functions. These similarities show the performance of their algorithm.

These ethics have emphasis on gaining skills rather than change in behavior but they are still not developed. A measurement tool is used for evaluating the student's abilities to identify & settle ethical problems. In order to do this we have designed & proved an analytic method, which consists of 5 components. They are identification and constructing the dilemma, the usage of information such as facts or concepts to settle problems but not involved in analysis, appearance taken and put forwarded for formal decision. We evaluated the capabilities of nearly 120 student's including new joiners and graduate persons. This evaluation consists of 3 different tests. These are then recorded as scores used for further evaluation of the student's level. Then these results of test are discussed. Here the main objective is to introduce a web-based tool similar to CSM's cogito system for improvement used by faculty to identify & settle ethical problems [12].

Hameed *et al.* [13] deals with student assessment. According to learning objectives, the evaluating performance level is based on individual students. A best evaluation supports, certifies and helps to improve individual student achievement and also protects the fair evaluation of students. Hence, it should be reviewed regularly and should improve the suitable, and beneficial to deduce scores of students.



Odai Falah *et al.* [14] provides an important observant which dominates the business world; recently it is known as Customer Relationship Management (CRM). In developed nations now this management became the new competitive weapon. Most of the organizations are focusing on customers to make them as loyal brands, and also satisfying their personal needs using CRM systems Globalization. Their new belief is on Telecom industries. These services help in complex issues of modern technologies, and also continuous innovations. The main goal is to provide a software application that allows them to rise in the customer's interactions and also in productivity, organization, the internet, distribution channels and customer contact centers. The main objective of this paper is to increase the marketing strategies and their relationships that the Umniah Mobile Company which is located in Jordan follows today's competitive world, and to analyze its results by better understanding the typical contingencies prevalent in the business environment.

Dorina *et al.* [15] presents the outcome from data mining research, performed at famous Bulgarian universities, with the main aim to tell the high potential of data mining applications for management and to contribute the enrolment drive and to attract the most desirable students. For the better performance of students the research is focused on the performance of data mining model, and they also go through the student's personal life, previous-university and current university performance. The dataset is characterized based on the performance into three consecutive years. Many familiar mining algorithms [17-19], including a neural network a rule learner, a Nearest Neighbor classifier and a decision tree classifier are applied on the dataset. The performance of these algorithms is analyzed and compared and resulted.

The process of collecting various data from databases or data warehouses and evaluating it is called as data mining [16]. Educational Data Mining is an evolving technique, which includes various methods such as predicting student performance, visualization and analysis of data, providing feedback, social network analysis and so on. Improving the quality of education is one of the greatest challenges so as to improve student's performance. Thus, it is very important to set new plans and designs for a better improvement of the current processes. This model helps to predict the student's future learning outcomes using data sets of senior students. There by to design a student's data base using J48 algorithm which proved to be an efficient algorithm in terms of accuracy identified by a comparative study of data mining classification algorithms [20].

Overview of existing system

- Some web based methods has been used for the analysis of student reviews.
- It only gives the positive and negative reviews separately, but it does not give the decision what to be improved.

Disadvantages of existing system

- It is not efficient.
- In the case of Starbucks brand, nearly 80% of all comments found were neutral in nature without mentioning any emotion.
- They are statements of fact or information, which do not express either positivity or negativity.
- All these reviews do not come up with a decision to perform an improvement action.
- For brands, both positive & negative reviews are important and as a result here the automated sentiment analysis fails.

PROPOSED SYSTEM

The proposed system collects all the feedback and divides the reviews accordingly. Natural language processing and data mining techniques will be used for mining user feedback. At first using the Stanford nlp toolkit and clustering exploratory analysis will be employed to gain a relationship between student opinions. Then a network will be build using the sub graph method. The module obtained will give information broker users, most discussed improvement of college infrastructure will be identified which can be used to improve care (Figure-1).

Advantages

- It accurately classifies the data and Quantifies performance.
- To come up with good accuracy levels of between 70% and 80% this automated sentiment analysis which compares very accurately as a human.
- To know more about the applications of sentiment analysis in the education related information.

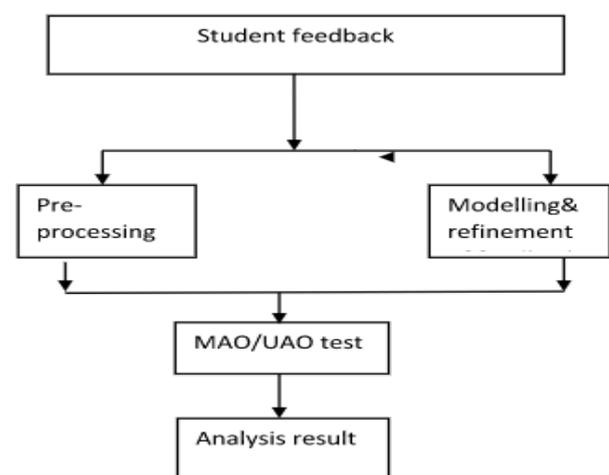


Figure-1. Architecture Diagram.



RESULTS AND DISCUSSIONS

First the person has to enter their email id and password to give feedback. As soon as this is finished it is directed to the above page where there is an option to give their review, name and their register number and then

clicks on post button. The review may be any type of words and these can be analyzed by the natural language processing algorithm which divides the positive and negative reviews.

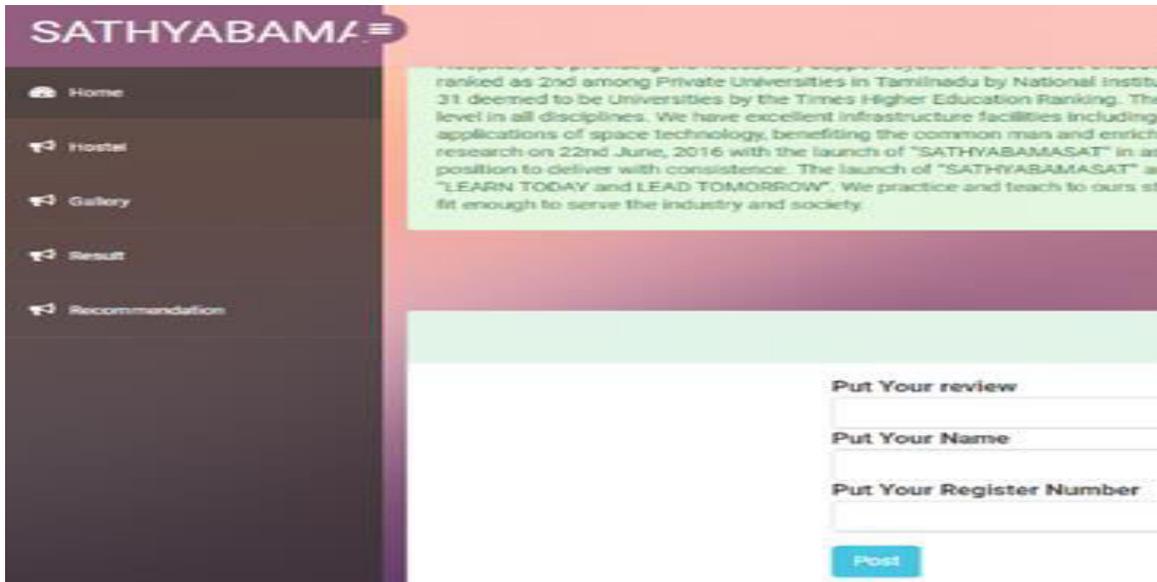


Figure-2. Main Review.

User Name	Reg No	REview	Sentiment
lokeshwari	111413205036	good	Positive
lokeshwari	111413205036	good	Positive
vasanthi	3411102	good	Positive
vasanthi	3411102	good	Positive
Mohan	8985766763	food is very bad	Negative
vinoth	8012386342	teaching very well	Positive
benin	7092166412	food good	Positive

Figure-3. Review details.

As we click on post button the entire details and reviews can be stored in the next page as above. The table consists of name; register number, the review entered and

the type of review whether it is positive or negative (Figure-3).

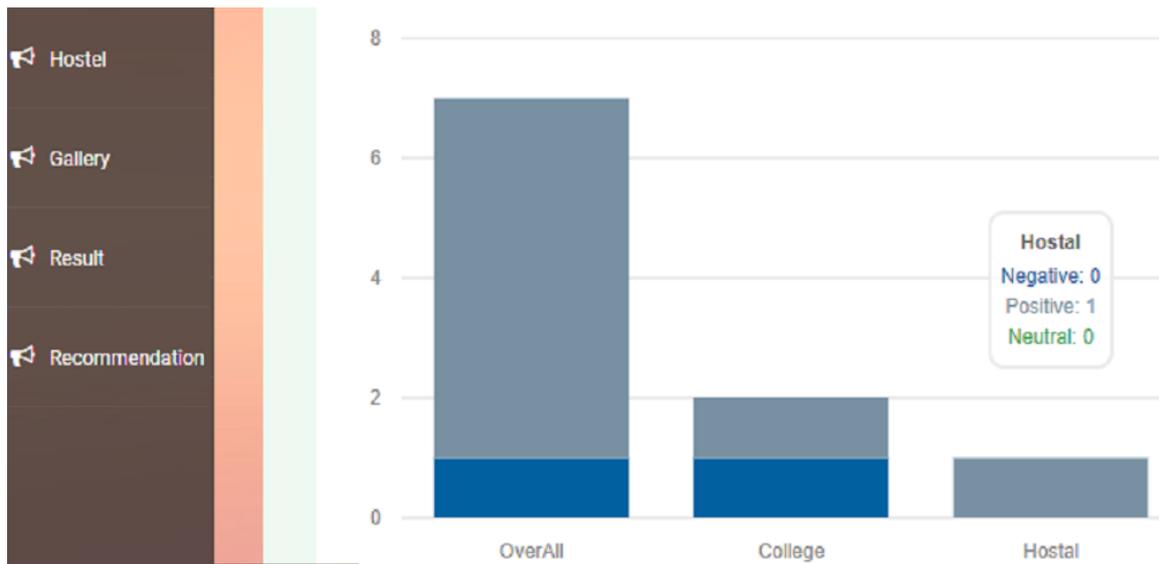


Figure-4. Positive and negative reviews.

The entire reviews are analyzed and pictured as graph which depicts the level of

positivity and negativity for hostel, college and overall institution (Figure-4).

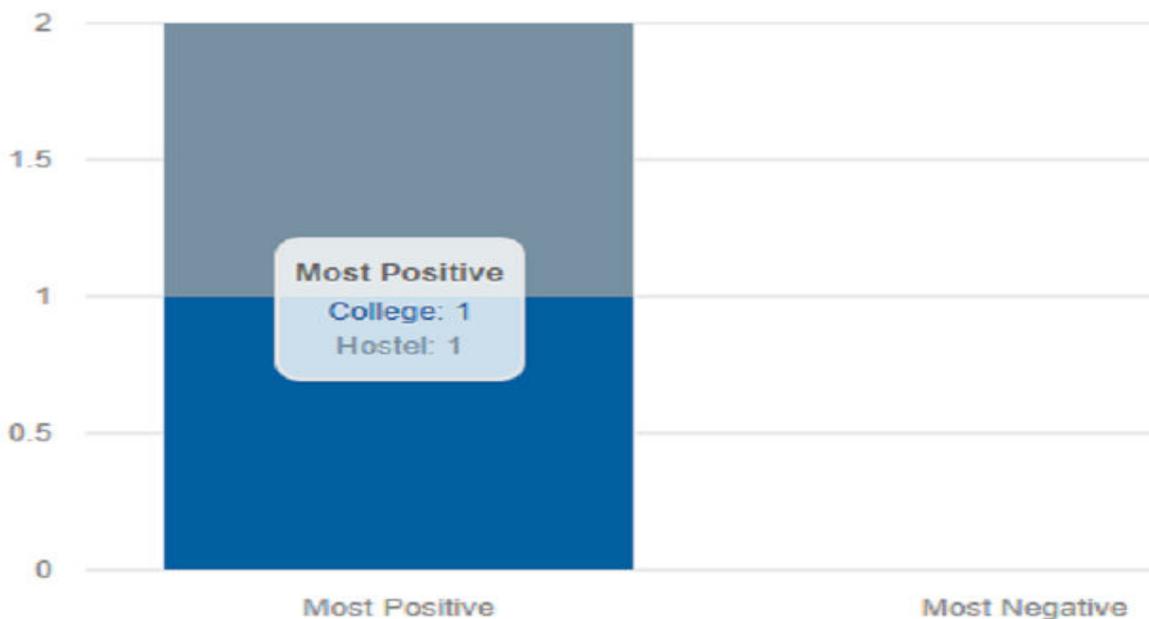


Figure-5. Aggregated results.

This is the final page which gives the decision for the management where the improvement is needed which is calculated through the negative reviews and also where the facilities are good, this is calculated from positive reviews (Figure-5).

CONCLUSIONS

Here we collected all the thoughts of students using feedback form. All the reviews are evaluated using positive & negative terms, and we can list all the resulted side effects. Social media can play a major role for the education sector in a way to provide product and service

optimization, cost reduction and quality education system. These obtained results could be used as feedback loop for student info. In future studies the post can be categorized on basis of their rankings or likes of post.

REFERENCES

[1] S. R. Das and M. Y. Chen. 2007. Yahoo! for Amazon: Sentiment extraction from small talk on the Web. *Manag. Sci.*, 53: 1375-1388.



- [2] L. Getoor and C. Diehl. 2005. Link mining: a survey. SIGKDD Explor. Newsl. 7: 3-12.
- [3] A. Akay, A. Dragomir and B. E. Erlandsson. 2015. Network-Based Modeling and Intelligent Data Mining of Social Media for Improving Care. in IEEE Journal of Biomedical and Health Informatics. 19(1): 210-218.
- [4] M. E. J. Newman. 2004. Detecting community structure in networks. Eur. Phys. J. 38: 321-330.
- [5] E. Le Martelot and C. Hankin. 2011. Multi-scale community detection using stability as optimization criterion in a greedy algorithm. Proceedings of the 2011 International. Conf. erence on Knowledge Discovery and Information Retrieval (KDIR 2011), Paris, France: Sci Te Press. pp. 216-225.
- [6] Singh, Randhir. 2013. An Empirical Study of Applications of Data Mining Techniques for Predicting Student Performance in Higher Education. (2013), IJCSMC. 2(2): 53-57.
- [7] 2015. Identifying Hidden Patterns in Students. Feedback through Cluster Analysis. Anwar Muhammad Abaidullah, Naseer Ahmed, and Edriss Ali, International Journal of Computer Theory and Engineering. 7(1).
- [8] Faculty Performance Evaluation Using Data Mining. 2014. I Nirmala G, IIP.B. Mallikarjuna I Dept. of ECE, II Dept. of CSE I, II Visvesvaraya Technological University, Belgaum, India, International Journal of Advanced Research in Computer Science & Technology (IJARCST 2014). 2(3).
- [9] Shen, Rui Min, Yiyang Tang, and Tong Zhen Zhang. 2001. The intelligent assessment system in Web-based distance learning education. Frontiers in Education Conference, 2001. 31st Annual. Vol. 1. IEEE.
- [10] Thakar Pooja. 2015. Performance analysis and prediction in educational data mining: A research travelogue. arXiv preprint arXiv:1509.05176.
- [11] Leng, Gang, Thomas Martin Mc Ginnity and Girijesh Prasad. 2006. Design for self-organizing fuzzy neural networks based on genetic algorithms. IEEE Transactions on Fuzzy Systems. 14.6: 755-766.
- [12] Shuman, Larry J., *et al.* 2004. Can our students recognize and resolve ethical dilemmas. CD) Proceedings, 2004 American Society for Engineering Education Conference.
- [13] Hameed Ibrahim A. and Claus G. Sorensen. 2010. Fuzzy systems in education: a more reliable system for student evaluation. Fuzzy Systems. In Tech.
- [14] A. Lghaswyneh, Odai Falah Mohammad. 2014. The Impact of Implementing Customer Relationship Management in Umniah Mobile Company Jordan. European Journal of Business and Social Sciences. 2.10: 169-181.
- [15] Kabakchieva Dorina. 2012. Student performance prediction by using data mining classification algorithms. International Journal of Computer Science and Management Research. 1.4: 686-690.
- [16] Sumitha R. and E. S. Vinothkumar. 2016. Prediction of Students Outcome Using Data Mining Techniques. International Journal of Scientific Engineering and Applied Science (IJSEAS) 2.6: 8.
- [17] P. Asha, Dr. T. Jebarajan. 2017. SOTARM: Size of Transaction based Association Rule Mining Algorithm. Turkish Journal of Electrical Engineering and Computer Sciences. 25(1): 278-291.
- [18] P. Asha, Dr. S. Srinivasan. 2016. Analyzing the Associations between Infected Genes using Data Mining Techniques. International Journal of Data Mining and Bioinformatics, Inderscience Publishers. 15(3): 250-271.
- [19] P. Asha, Dr. T. Jebarajan. 2013. A survey on efficient incremental algorithm for mining high utility item sets in distributed and dynamic database. in the International Journal of Emerging Technology and Advanced Engineering, ISSN 2250-2459.
- [20] P. Asha, Dr. T. Jebarajan. 2014. Improved parallel pattern growth data mining algorithm. in the International Review on Computers and Software, ISSN 1828-6003, 9(1): 80-87.