International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/CSEIT22847

Student Feedback Analysis with Recommendations

Chethan G. S, Harshitha H S, Meghana Bekal, Nithya V Shet, Shama G
Department of ISE, J N N College of Engineering, Shivamogga, Karnataka, India

Article Info
Publication Issue :
Volume 8, Issue 4
July-August-2022

Page Number : 56-62

Article History
Accepted: 02 July 2022
Published: 11 July 2022

ABSTRACT

When modern institutions use their student data, they can better understand their students' educational experiences. Teachers are better able to educate their kids as a result of this. Big Data is also being used to transform the educational system in order to provide a well-rounded education to pupils. Analyzing how well the teaching has been effective for the students is an important criterion in teaching. Students' feedback is an important component that should be encouraged in order to improve the learning experience. Each faculty member is rated on a scale of 1 to 5, with 5 being the highest, based on student feedback. To capture and process emotions from feedback, Naive Bayes classifiers and simple text mining algorithms were applied. Rating scores are clustered using the K-means clustering technique.We use sentiment analysis techniques to analyse student feedback in this work.

Keywords: Student Feedback, Systematic Literature Review, Sentiment Analysis.

I. INTRODUCTION

The term "big data" refers to a dataset that continues to increase to the point where it is challenging to handle using traditional database administration concepts and techniques. It's possible that the issue has something to do with data collection, depository, sharing, search, analytics, and visualisation, among other things. The three dimensions of big data are volume, velocity, and variety.

i. Volume: The amount of data in terabytes and petabytes is enormous.
ii. Velocity : To optimise its value to the business, it should be employed while streaming into the enterprise.
iii. Variety: It encompasses more than just structured data, including unstructured data of all varieties: text,audio,video,posts,log files etc.

A Big Data platform should give a solution which is designed specifically with the needs of the enterprise in the mind[1].

Huge amounts of educational data are generated, and many educational institutions are having trouble figuring out how to effectively harness and analyse this data in order to provide continuously improving education. Nowadays, educational data is growing in quantity, diversity, and speed.

In order to extract value and make well-informed judgments, this created data must be managed effectively.[2].

Once the student submit feedback by login to the portal,the feedback is analyzed through the analyzer.It is classified into positive,negative and neutral polarity.This polarity helps to calculate whether the performance is good or bad or to calculate whether the performance is good or bad or moderate.The system's objective is to assist teachers in improving their teaching methods based on feedback received[3].

II. LITERATURE SURVEY

For sentiment analysis of student input, the researchers utilised a combination of machine learning and lexicon-based approaches. Textual feedback, it is often collected at the conclusion of a semester, gives important insights regarding overall teaching quality and practical methods for improving teaching methodology.The article also compared the results of other sentiment analysis APIs to the hybrid approach proposed.

Dataset Description:

All comments were gathered from our institute's educational portal for this study. Sentiment polarity labels were manually added to the dataset.*{ positive, negative,neutral}*.

Preprocessing :

Punctuation, Tokenization, Case
Conversion ,Stop words are removed.

Data Partition :

At random, divide into train and test sets.The dataset was utilised for training 70% of the time, with the rest being used for testing.
Split into train and test sets at random. 70% of the dataset was used for training, with the remainder used for evaluation.

Feature Extraction :

Following data separation, both the training and testing datasets were used to extract features.Unigrams, bigrams, TF-IDF, and lexicon-based features were used to train the hybrid model for sentiment analysis.[4].

Naive Bayes' Classifier:

The Naive Bayesian network classification algorithm is a well-known Bayesian network classification algorithm.A probabilistic classifier is Bayes' classifier. It uses the Naive (Strong) independence assumption and is based on Bayes' theorem.The core premise of the Naive Bayes technique is to find the probabilities of categories given a text document by using the joint probabilities of words and categories.

Algorithm 1:

NaiveBayes' classifier

Input: a document d
A fixed set of classes C = {c1,c2,……cj}
Output: a predicted class c

1. Preprocessing:

ii. Feedback is being gathered from the standard benchmark databank.
iii. Positive and negative responses were separated into two folders, positive.txt and negative.txt.
iv. Two blank lists were made:The first is for good feedback, and the second is for negative feedback.
v. Positive and negative feedback sentences were divided up, with the terms 'pos' and 'neg' applied to each and saved in two separate lists.
vi. The dictionary retained 3/4 of these sentences for training purposes, whereas the remaining 1/4 were retained for testing purposes.

2.Using the chi squared test, they determined the score of each of the remaining terms, and instead of using all of them, only the best of them were used. The classifier was honed with the help of the newly constructed databank.

3.In reference sets, labelled sentences were correctly kept, while the predicatively labelled variation was retained in test sets.

4. As needed, metrics were calculated.[5].

The Naive Bayes classifier can be used to identify feedback accurately based on the degree of influence of various factors on feedback. Furthermore, a Nave Bayes classifier covers more ground faster than discriminative models like logistic regression since it requires less training data[6].

Introduce a recommendation system based on the analysis.The methods used in this paper are Sentimental analysis,Latent Direct Allocation (LDA), Opinion Mining using Lexicon based approach.

The sentiment Analysis process is as below,

1. Preprocessing.
2. A lexicon of expressions.
3. Tagging for polarity.

The number of times a word is used.

4. .The number of times a word is used.
5. Attitude of the words.
6. Overall attitude.
7. Word cloud visualization.
8. Sentiment score [7].

Handling qualitative student opinions while generating automatic reports is a difficult task. Based on the findings of a two-layered LSTM model, We present a method for supervised aspect-based opinion mining. The first layer anticipates and then determines the direction of the feedback-described features (positive, negative, or neutral)[8]. Classification is an important aspect of data mining that can be done using a variety of inductive learning methods. Because of its benefits, decision trees are one of the most widely used strategies. In ensemble classifiers, they are commonly employed as a fundamental model. However, learning from data streams with these traditional inductive learning models or algorithms remains a challenge [9].

Algorithm 2: Recommendations based on the analysis of Student feedback system.

Documents D1, D2,...Dd are the sources of information.

Output: Indication of a change in required aspects.

Step 1: Select documents D1, D2…,Dm.

Step 2 : To the document that was created, apply LDA

Zn~Multinomial(θ)

Step 3: Determine polarity of aspects [-3<P<3]

$$P = \sum_{i=1}^{n} PTsi / N$$

Step 4: If (P0) is true, then

Suggestion: 'Change the teaching method.'

Otherwise, suggest 'Continue the existing technique'[10] if (P>0).

SMS text processing, reading, and categorization is part of data preparation.The most basic model, the "corrected" model (which corrects spelling errors), and the "sentiment" model (which extends the "corrected" model by doing sentiment mining) were built. [11].

The student comments obtained via Google form is input data that serves as training data for the system. When given test samples, the trained system uses machine learning techniques to categorize the sentences as negative, neutral, or positive. This outcome is depicted graphically[12].

Instructors might be rated as 'Bad,' 'Average,' or 'Good.' Each one weighs 0.1, 0.2, or 0.3. We add the overall weight of the student's response once we have generated values for all five weights. This weight, as previously stated, is based on a 1.5 total value.This weight value, together with the student's assessment of that specific faculty member, is added into the database.Following the mapping of the weights into the database, to calculate the overall weight for that specific star rating, we add the total counts under

each column of the rating for each faculty.. The entire weight is then multiplied[13].

The EPDM + ML (Data mining and machine learning are all part of the educational process) model,based on descriptive decision theory, which studies the rationality of learners' judgments through textual data quantification and statistical analysis, the model was built and implemented using a combination of text mining and machine learning technologies. The implementation is done in 2 phases: First,the main functions of EPDM model is applied to analyse textual data(Feedback).It also includes cleaning and filtering of databank.The emotional valence and sentiment scores of the comments are extracted and compared to the data in the SET instrument. In the second phase, based on the gender, sentiment, and emotional valence scores of the students, an ML classification model is constructed that predicts what a student's suggestion for teachers is.[14].

The students rate the faculty with the features such as,

1. Overall learning with experience
2. Delivery of high-quality instruction

Teaching resources of high quality

3. Teaching resources of high quality

Exchange of information with lecturers

4. Exchange of information with lecturers
5. College classmates' communication [15].

III. METHODOLOGY

The methodology of this research is shown in Fig. 1, here the first four cells will reserve the feedback gathered from the student which are interpreted and stored in the list. Instead of collecting it from the students, feedbacks are taken from the dataset. The Splitter class will be used to split and tokenize each sentence in the feedback. Text mining is a sentiment analysis technique. Text mining converts unstructured data such as words, phrases, and sentences into structured data. To make preparation easier, the entire document is transformed to lower case. The text is then tokenized before the matrix forms are generated. Let us consider the example for a student's evaluation of the teacher in the classroom, “The concepts explained by him were not clear, needs improvement”. As seen, these two sentences are tokenized and saved in a list of lists.[[‘The’,‘concepts’,‘explained’,‘by’,‘him’, ‘were’, ‘not’, ‘clear’], [‘needs’,‘improvement’]]. Bag vectors serve as key words in this document, which is separated into words. Stop words are the most frequently used and thus removed.Lemmatization is the process of stripping a word of its inflectional endings and restoring it to its base or dictionary formText mining, the Naive Bayes algorithm, and a multipoint rating scale ranging from one to five stars are utilised to grade a faculty. After that, the Naive Bayes classifier will develop an output class for the faculty member whose feedback it has received. The Bayes theorem states that conditional probability can be used to classify an input. The feedback received from the students is very important in determining the characteristics of a good teacher.Some of these include instructor demeanour, subject expertise, and overall teaching efficacy.

The classifier algorithm's mathematical formula is shown in Eq1,

$$P(c|x)=P(x|c)P(c)/P(x) \quad (1)$$

$$P(c|X)=P(x1|c) * P(x2|c) * ... * P(xn|c) * P(c)$$

The numbers x1, x2,..., xn in the mathematical formula above represent the total number of matches identified between the input and reference data sets for the same characteristic under a given level of danger. P is found by multiplying the total number of rows in a common data set by the total number of rows in our entire standard data set (c). A number is assigned to each risk category. The risk class that provides the highest result from the preceding

calculation is assigned to a given input. Standard data sets have been created and saved in the correct database settings.The Naive Bayes classifier is given a list that contains the final emotion for each attribute from the previous phase, and the algorithm is run. The input values are compared to those in our reference data set.It raises the value of the attribute 'x' by one when it finds a match. Thetest dataset served as the basis for evaluating taught models. The following evaluation metrics were used:

All paragraphs must be indented. All paragraphs must be justified, i.e. both left justified and right-justified.

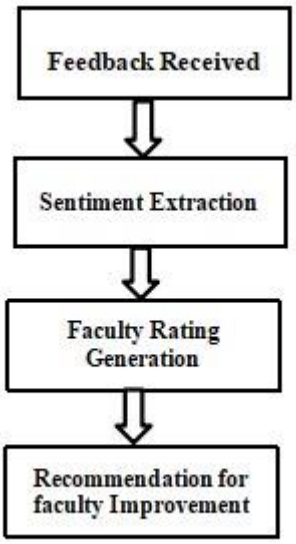

FIGURE 1:PROPOSED METHODOLOGY

Accuracy : As indicated in Eq 2, accuracy is defined as the ratio of the number of true predictions made by the model to the number of rows in the dataset.

Accuracy=#correctpredictions/#datapoints (2)

F-Measure : In the multi-class classification problem, F-Measure is another often used measure. The geometric mean of precision and recall is what it's called. The ratio between the right predictions and the total predictions made by the system is known as precision. The ratio of the model's correct predictions to the total number of true emotion labels is known as recall.F-measure is a useful indicator for assessing a model's performance when the data is extremely skewed, as indicated in the equation below.

F-measure=(2*Precision*Recall)/(Precision+Recall)

The suggested system will assist us in establishing a recommendation system that has the following features.: The suggestion would be made in cases where public opinion turns out to be negative. Tutors would have a clear picture of the areas that needed to be addressed as a result of this. It is also vital to be able to provide recommendations when modifications are needed in order to boost grades. After the student feedback opinion was extraction is completed,an unambiguous picture of the student's perspective on the course is acquired, indicating whether it was adequate or required adjustments any one of components. It's not the case sufficient to collect and analyse student input to decide if it is positive or bad.

The following stages can be used to design a recommendation system.

i. Choosing the data: The information set is aa token collection used to extract polarity and point of view
ii. Analysis of public opinion on a chosen dataset: Student feedback is analysed to determine if it is favourable or negative. The technique is lexicon-based, with polarity ratings ranging from -3 to +3. After that, opinion mining is performed to evaluate whether a factor is somewhat positive, faintly positive, very Positive, slightly negative, moderately negative, and strongly negative are all words that can be used to describe a situation.Python programes can provide tutors with recommendations for areas of student feedback that have been deemed unsatisfactory by opinion mining.

.

IV. CONCLUSION

A data mining methodology is proposed in this system depending on specific characteristics of the institution, the best method to learn about a professor is through student feedback. By evaluating the pupils' contributions, we were able to successfully categorize the faculties into their relevant groupings. This system assists the faculty in improving overall in order to improve as a professor. As a result, the institute's authorities are able to maintain its quality.

V. REFERENCES

[1]. Sachchidanand Singh, Nirmala Singh, "Design Big Data Analytics," Inproceedings of International Conference On Communication, Information & Computing Technology (ICCICT),2012

[2]. AMR A. Munshi and Ahmad Alhindi, " Big Data Platform for Educational Analytics", IEEE,vol.9,pp.52883-52890 doi: 10.1109/ACCESS.2021.3070737

[3]. Dr.Vijaykumar S.Bidve, Pratiksha Patawari, Rucha Raut, Arya Kadam,

[4]. Zarmeen Nasim, Quratulain Rajput† and Sajjad Haider, "Sentiment Analysis of Student Feedback Using Machine Learning and Lexicon Based Approaches" Conference Paper July 2017 DOI: 10.1109/ICRIIS.2017.8002475

[5]. Lopamudra Dey, Sanjay Chakraborty, Anuraag Biswas, Beepa Bose and Sweta Tiwar,"Sentiment Analysis of Review Datasets Using Naïve Bayes' and K-NN Classifier",In journal "International Journal of Information Engineering.and Electronic Business", Vol. 4, pp. 54-62, 2016,doi: 10.5815/ijieeb.2016.04.07.

[6]. Sandhya Maitraa*, Sushila Madanb "Mining authentic student feedbackforfaculty using Naïve Bayes classifier" scientific committee of the International Conference on Computational Intelligence and Data Science (ICCIDS2018). Doi: 10.1016/j.procs.2018.05.032

[7]. Quratulain Rajput, Sajjad Haider, and Sayeed Ghani, "Lexicon-Based Sentiment Analysis of Teachers' Evaluation",Vol. 2016,doi: 10.1155/2016/2385429.

[8]. Irum Sindhu1, Sher Muhammad Daudtota2 and Kamal Badar3, "Aspect Based Opinion Mining On Student's Feedback For Faculty Teaching Performance Evaluation" Citation information:DOI10.1109/ACCESS.2019.2928872, IEEE Access.

[9]. "Improved Sentiment Analysis for Teaching Evaluation Using Feature Selection and Voting Ensemble Learning Integration",In "2016 2nd IEEE International Conference on Computer and Communications(ICCC)", pp.1222-1225.doi: 10.1109/CompComm.2016.7924899.

[10]. Karunya K, Aarthy S, Karthika R and Jegatha Deborah L, "Analysis of Student Feedback and Recommendation to Tutors." 2020.

[11]. Chee Kian Leong, Yew Haur Lee and Wai Keong Mak, "Expert Systems with Applications." 2012.

[12]. Daneena Deeksha Dsouza, Deepika, Divya P Nayak, "Sentimental Analysis of Student Feedback using Machine Learning Techniques" International Journal of Recent Technology and Engineering (IJRTE)ISSN: 2277-3878, Volume-8, Issue-1S4, June 2019.

[13]. Krishnaveni and Rohith R pai, "Faculty rating system based on student Feedback using Sentimental Analysis" ,2017, IEEE.

[14]. Analysis of Student Feedback and Recommendation to Tutors, "Towards teaching analytics: a contextual model for analysis of students" evaluation of teaching through text mining and machine learning classification" 2021.

[15]. Sian Lun Lau, TzeYingSim "Feedback of University Students Online Delivery Learning During the COVID-19 Pandemic Period". 2020

IEEE Conference on e-Learning, e-Management and e-Services (IC3e).

Cite this article as :

Chethan G. S, Harshitha H S, Meghana Bekal, Nithya V Shet, Shama G, "Student Feedback Analysis with Recommendations", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 8 Issue 4, pp. 56-62, July-August 2022.
Available at doi : https://doi.org/10.32628/CSEIT22847
Journal URL : https://ijsrcseit.com/CSEIT22847