

# Faculty Rating System Based on Student Feeding using Word Analysis

Rohit G, Mohammed Shahwaan Aliku, Pooja Shankar Shinde, Shivashankamma,

Mrs. Ashwini

*M2 Student, ISE / National Institute of Engineering, Mysuru, Karnataka, India*

*Associate Professor, ISE / National Institute of Engineering, Mysuru, Karnataka, India*

---

**Abstract:** Educational Data Mining is a prominent area of knowledge assessment in the field of education using data mining algorithms. In this paper we have used a number of learning algorithms to successfully measure academic faculty on the basis of the feedback sent by students. Our proposed model uses visual analysis and machine learning algorithms to capture emotions from the students' response system. This model provides an accurate and efficient way to measure the intelligence of a particular educational institution. Through this proposed model the intelligence will be tested and measured with specific parameters that will help us improve the quality of learning and education.

Index Terms - Naive Bayes Group Partner, Data Mining, Text mining.

## 1. INTRODUCTION

In a large country like India where millions of students are constantly engaging with the education sector through various educational institutions, tracking each internship that is relevant to a particular educational institution is important. It helps to keep track of the quality of education in educational institutions. Our application aims to help achieve the same. Based on the feedback submitted by students, each remedy is rated on a scale from 1 star to 5 stars. Data mining algorithms such as the Naive Bayes classifier and basic text mining methods have been used to detect the feedback from the response and process it. Other instruments were also assigned to the food items included by each student according to their academic performance, the authenticity he or she displayed by submitting feedback and other factors. Algorithms are executed using the python programming language. Through the successful processing of response forms and optimization algorithms, output values are generated and updated in the database and are finally displayed. App endpoints are built using HTML, CSS and JavaScript. Server side communication is done using PHP.

### A. Existing Solutions

There are many ways to solve the problem here. Most of the existing analysis algorithms were designed for binary classification, which means they assign ideas or updates to respiratory categories such as Positive or Negative. The concept of product review using text mining has become very popular recently.

For the authors of they focus on how we can use the basic analysis of reviews brought by consumers for various products. It describes the various levels at which a Sentiment analysis can be performed. Traditional ways of dividing something into good or bad. This paper introduces the concept called multi-point scales and is known as the infence rating we have incorporated into our project. The model discussed in this paper takes a set of reviews as input, and then executes them using three key steps, Data Correction, Review Analysis and Sentiment Classification. The results produced by such a model are the classification of reviews, test sentences, or ideas identified in reviews.

In analysis is defined as using the information of digital news articles about its positive or negative feeling about a particular politician. This paper suggests a simple model for analyzing the unity of a digital newspaper using the Naive

Bayes classifier method. The model uses a first set of data that will be updated when new information appears. The system discussed in comprises product reviews from various web sources and uses a machine learning algorithm to predict users' reactions to the product. The only requirements are that the user should know the specific name of the product they wish to purchase and only the existing English reviews can be used.

This paper successfully illustrates how product reviews can be differentiated based on the semantics they carry. This helps us to reduce crossing traffic through various sites. This program currently focuses on text segmentation only. In the future, the system can be trained to process icons in updates. Different languages can also be added to the system. The efficiency of the algorithm can be increased by using clustering methods

## **B. Our Approach**

Mining is another widely used method of analyzing emotions. It is a process of obtaining high quality information in a text. We use text mining followed by the use of the Naive Bayes algorithm to measure intelligence. We use a multi-point scale for a range of 1 to 5 stars to measure intelligence. First of all we get feedback from students about a faculty based on four qualities that we have found to be important for a good teacher. It is a professional ethics, knowledge of subject matter, communication skills of the subject and a thorough teaching of the subject. Once the response has been received, we are conducting a Naive Bayes production phase for the production phase of the study. The main benefits of our program are:

1. Instead of classifying intelligence in bipolar classes as good or bad in a typical situation, we use a multiple-point scale to measure teacher.
2. The student can provide his / her technical description in a practical rather than objective way which is the most commonly used method.
3. The basic text mining technique is used to enable the reader to present the answer in a supportive way.
4. Each answer submitted by the student is given a particular weight depending on their performance and the commitment they have filled out in the answer form. This helps in increasing the accuracy of the fullimeter calculation by reducing errors in the standard response system.

## **2 DESAILED APPROACH**

The user connects to the front end of the application. The leadership team response is collected from the student using the feedback form. We store the attributes in the teacher's Department database, Student Name for Feedback, Name of student for which the answer is sent, Student ID and Technical Team, Student CGPA, Student gender and four major criteria for assessing the aforementioned intelligence, to store student data. All information entered by the student is then looked up to specific attributes in the database. Required attribute values are processed to include specific emotions from the response using text mining, which is later assigned to the Naive Bayes classifier as input. The normal flow of the system is shown in figure 1.

## **A. Journal of Mining**

For text industries we use the NLTK module (Natural Language Toolkit) which contains the Splitter section. First, all four cells that store the response received from the reader are read and stored in the python list. Each sentence was then separated and typed using the Splitter section. Each punctuation sentence was stored in a python list. After typing all user-entered sentences in relation to a specific attribute of the teacher, the entire alphabet is stored in another list ie; in the list view

For example, let's take a look at this answer given to the student about the teacher's behavior in the classroom, "he had a hard time at first! But, now you are friendly". These two sentences are typed and stored in a list as shown below,

```
[[ 'In the beginning', 'he', 'was', 'very', 'strongly', '!', 'But', 'now', 'he', 'is', 'friendly' ]]
```

When a sentence is taken the next step is to look at the feelings described in the input. In the above case, a list is generated to store only the emotions as shown below,

```
[ 'Very strict', 'friendly' ]
```

Now, once this is done we need to find that one word that best describes the performance of the teacher by looking at all the feelings we received in the input. To this end we created two tables of sentiment used by many users on a regular basis with their respective values. The first table is a list of possible emotions and their values corresponding to the behavior of the skills team and the second table preserves the possible emotions and their values corresponding to the knowledge, communication skills and overall teaching effectiveness. Given below are the tables, Table 1 and Table 2.

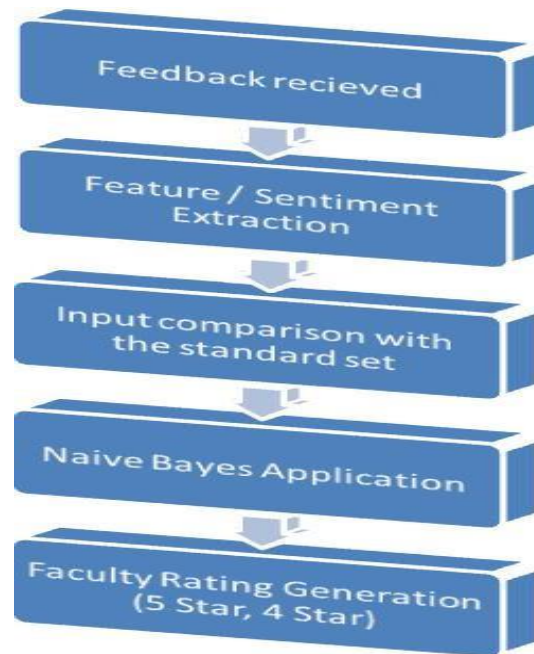


Fig. 1: Overall System Flow

So, in our example above with the list storing the words

**Table 1.** The most commonly used words to describe the behavior attribute and the values assigned to them.

Word	Value
very rude	1
very poor	2
very bad	3
very strict	4
rude	5
poor	6
bad	7
strict	8
arrogant	9
not nice	10
short tempered	11
unpleasant	12
not good	13
not polite	14
doesn't understand	15
not bad	16
not strict	17
ok	18
indifferent	19
nice	20
polite	21
understands	22
good	23
very polite	24
very nice	25
calm	26
cool	27
very good	28

['Very moderate', 'friendly'] output name could not be 'strictly'. We find this by adding the values of both words in the list above and finding the middle value. After that we find the final value we collect in the integer. Finally we write the corresponding name and the last value we calculated, in the list using the Naive Bayes classifier. We do the same with the other 3 main attributes.

## B. Ideas

The first thought is that, since most of the most frequently used words in the faculty answer are mentioned as shown in table 1 and table 2, we assume that the student makes the most of those feelings while submitting the answer. In the rare case of not finding a match of a student's position and name on our table, the value is listed in the checklist where Naive Bayes using algorithm. The second assumption is, that the reader submitting a response completes that form and comments on at least one aspect of the specific ability for successful verification of the response. If a student tries to submit an empty response form, the application rejects it. The third thought is that the words most often used for faculty response are identified in the

appropriate relationship with the teacher and the student. Any type of junk input will not be processed.

**Table 2.** The most commonly used words to describe the knowledge, communication skills and overall teaching effectiveness attributes and the values assigned to them.

Word	Value
very poor	1
very bad	2
poor	3
bad	4
doesn't know anything	5
not good	6
not bad	7
must improve	8
doesn't know much	9
ok	10
not interesting	11
average	12
not great	13
not perfect	14
can improve	15
good	16
interesting	17
has improved	18
very good	19
great	20
perfect	21
superb	22
fabulous	23
awesome	24
brilliant	25
excellent	26
best	27
outstanding	28
pleasant	29
friendly	30
superb	31
fabulous	32
awesome	33
sincere	34
dedicated	35

### C. Use of Naive Bayes Classifier

Naive Bayes is an algorithm used for the input of a particular output class also known as the risk class. It is an exemplary approach to supervised learning. It is based on a Bayes theorem that uses the concept of probability. Given below is the

mathematical formula for the classifier algorithm,

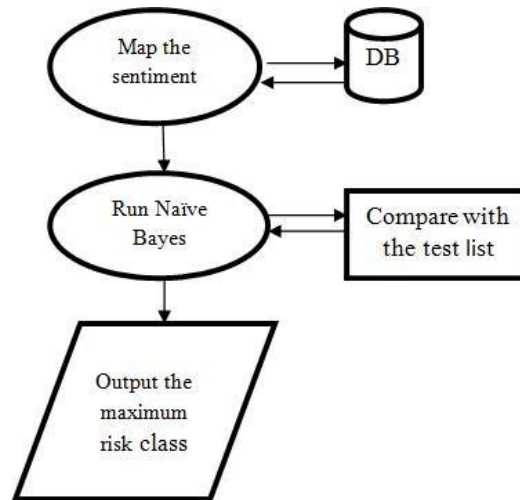


Fig. 2: Application of the Naive Bayes Classifier.

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$

$$P(c|X) = P(x_1|c) * P(x_2|c) * ... * P(x_n|c) * P(c)$$

In the above mathematical formula,  $x_1, x_2, \dots, x_n$  represents the amount of match between the input and the standard data set for that attribute under a specific risk class. 'C' is the total number of rows in the standard data set for a particular risk category and  $P(c)$  is the total number of rows for the standard data under a specific risk category for the total number of rows in all standard data sets. We calculate this value for all risk classes. The input provided belongs to the risk category that produces the maximum value from the above formula. In our case there are five harmful categories that vary from 1-5 stars. Standard data sets are designed and stored in specific attributes in the database.

From our previous step, the list with the last encouragement / name of each attribute is fed into the Naive Bayes classifier as the input and algorithm are created. Comparing input values with those in our standard data set. In each game you get an 'x' value for that attribute to be increased. The number of rows for each risk category and the total number of rows in the standard data set are also calculated. An accident class that generates a maximum amount of time in which the algorithm is generated is the class of that subject. Therefore, afterwards the results are released including all information such as the name of the student submitting the answer, the name of *Student weight restoration*

In our normal cases we calculate the sum of the factors such as marks, earnings, losses etc. by adding the total value of the whole and then dividing the total by the total number or the number of instances. The same can be used here to find the average skill level. But there have been a few problems because of how the feedback sent by the reader can affect professional ratings. One such example is a student acting on behalf of a teacher and intentionally giving a negative response or another student sending a response form on their behalf for a certain amount of garbage in a few seconds. Our program here focuses on how to solve this problem. Although this problem is not completely overcome, it can certainly be reduced to some extent. So, assign specific instruments and feedback sent to students.

Calculating the average skill level, instead of dividing the total number of competencies with the number of students submitting a response, will produce a more accurate result when we calculate the student's weighting method we are talking about here. The education team should provide a brief response of the student before the student writes a response about them. Typically we take into account the 5 key elements by which we give weight to a particular student's response. They are student CGPA, attendance, class involvement, integrity and the time spent by the student to convey the idea of a particular skill.

The process of calculating the weight is shown in figure 3.

In our case the student response can range from 0.5 to 1.5. This is because no matter what the particular student is, his or her answer will not be completely blocked and not given as high a value as 2. If given a value such as 2, it's the same as saying that a student's answer is as important as the answer of two students, which is wrong. Each element is assigned an equal weight of 0.3. In case of the total weight of the response sent by the student below the 0,5 rate, it gives us a value of 0,5.

First, the CGPA of the student submitting the answer form is mapped to the details and other information contained therein. The CGPA value (10 point scale) is divided into 10 followed by 3.33 to produce a value between range 0 and 0.3. Similarly, the time spent by the student (in seconds) is captured and mapped onto a digital map. Now assign a specific weight based on the accuracy shown by the student in completing the answer form. Assign instruments between 0 to 0,3 range according to the time spent by the student to complete the answer form. Table 3 shows the amount of weight assigned to the amount of time spent-the faculty for which the response was submitted, the Similarly, student attendance (in percentages) submitted by student ID, faculty ID and so on in the data and the average the faculty is included in the database. Assign instruments rating for the particular faculty. The use of the Naive Bayes between ranges of 0 to 0,3 depending on the learner. Table 4 Classifier is shown in Figure 2.shows the amount of weight assigned to the attendees.

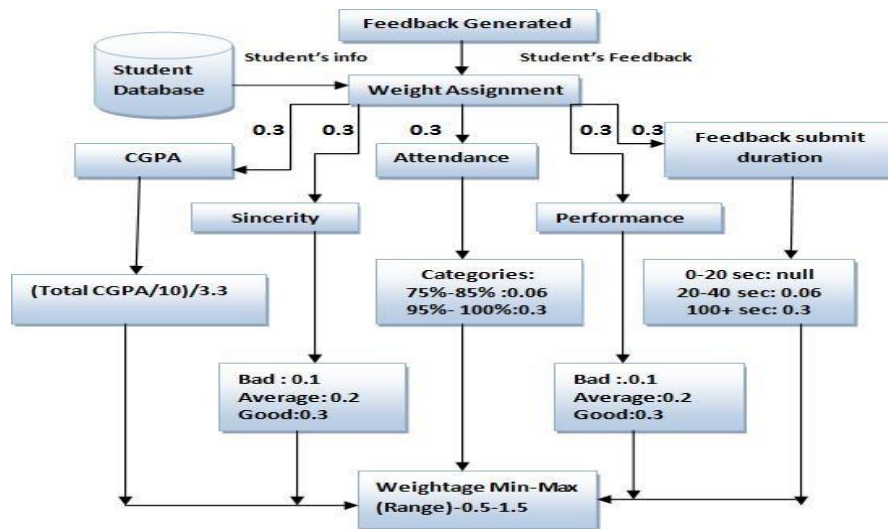


Fig. 3: Assigning Weights to the student feedbacks.

Table 3. The time spent by the student on a feedback form and the weight assigned to that respective feedback.

Time (in seconds)	Weight
0-30	0.0
30-60	0.06
60-90	0.12
90-120	0.18
120-150	0.24
More than 150	0.30

Table 4. Weight assigned according to the attendance of a student.

Attendance (in percentage)	Weight
< 75	0.0
75-80	0.06
80-85	0.12
85-90	0.18
90-95	0.24
95-100	0.30

Creativity measures student performance in the classroom and determines 'Bad', 'Moderate' and 'Good'. Each has a weight value of 0,1, 0,2 and 0,3 respectively. As long as we generate the sum of all the five instruments, we put them together to determine the weight of the response the student gave. This weight as mentioned earlier exceeds the total value of 1.5. Ultimately this amount of weight is added to the data, the ratio the student has provided to that particular organization, for calculating the average technical rating. When weights are already included in the information menu, we take the total calculation under each measurement column for each head, to find the total weight of that particular star rating. Next, we multiply the weight value of each measurement by the appropriate measurement value. Eg: If it is less than a 2-star rating column the total weight is 24 we multiply by 24 by 2 which is 48, to produce the sum of the average of the equations under that section. We also do the same in other columns. Now, we take the total calculation of the average value under each column as mentioned above and put them all together to produce the average value of a particular technology and the total weight by adding all the metals under each category. Finally we divide the average value of the average by the total weight to produce an average rating for a particular skill. The same procedure is done for all intelligence.

### III. CONCLUSION

In our program we have proposed a data mining method to measure the capacity of an educational institution based on specific factors. Student feedback is the best way to find out about faculty. By analyzing the food items sent by students using the basic text mining method and the Naive Bayes classifier we were able to classify the skills into appropriate categories (1 star - 5 star). We also introduce concepts of the weight of student feedback and the relevant student who helps overcome the obstacles encountered in traditional response programs. When a loggity logs into his portal they are able to know what their strengths and areas need to improve. This helps them develop more fully in order to become better. This helps the management of the center to maintain its quality. Our current program measures those skills in the teaching profession. As a future project we may extend this rating system to measure other non-teaching staffs in an institution that has a unique set. And the method of text mining can be further developed to reduce the amount of comparisons.

### ACKNOWLEDGMENT

We thank Amma for the inspiration to our work. We thank Dr. Sajeev G P for helpful Suggestions.

### REFERENCES

- [1] Cane W.K. Leung & Stephen C.F. Chan. Sentimental Analysis of Product Reviews. The Hong Kong Polytechnic University, Hong Kong SAR, 2008.
- [2] Yustinus Eko Soelistio and Martinus Raditia Sigit Surendra. Simple Text Mining For Sentiment Analysis of Political Figure Using Naive Bayes Classier Method. CoRR, 2015.
- [3] Jyoti Jain, Asst. Prof. Archana Shinde & Prachi Panchal, Sentiment Anal- ysis using Machine Learning. Department of Information Technology,



Sinhgad Academy of Engineering, Maharashtra, India, 2016

- [4] E.J. Fortuny, T.D. Smedt, D. Martens & W. Daelemans , Media Coverage In Times of Political Crisis: A Text Mining Approach, Expert Systems with Applications, Sciendirect, 2012.
- [5] G.F. Luger. Artificial Intelligence Structures And Strategies For Complex Problem Solving, Pearson Education Inc., 182-185., 2009.
- [6] Yamanishi, K., & Li, H. Mining open answers in questionnaire data. IEEE Intelligent Systems, 17(5),pp. 58-63, 2002.
- [7] Jiawei Han, Micheline Kamber & Jian Pei. Data Mining Concepts and Techniques. Waltham: Morgan Kaufmann Publishers, 2012.
- [8] Liu, B., Hsu, W., & Ma, Y. Integrating classification and association rule mining. Proceedings of Knowledge Discovery and Data Mining, pp. 80- 86, 1998.
- [9] C. P. Prathibhamol, Ashok, A. Solving multi label problems with clustering and nearest neighbor by consideration of labels, Advances in Intelligent Systems and Computing, vol. 425, pp. 511-520, 2016.