

Extracting Patterns from Students' Feedback using Mixed Method

Th. Shanta Kumar

Department of Computer Science and Engineering
Girijananda Chowdhury Institute of Management and Technology, Assam, India
thshantakumar@gimt-guwahati.ac.in

Available online at: www.ijcseonline.org

Abstract—It is common practice to collect feedbacks from the students in order to evaluate the performance of their teachers. Questionnaire is one of the best ways of collecting this information. Analysis of this questionnaire is crucial and not easy especially if it contains both quantitative and qualitative measures. This paper provides a comprehensive analysis of students' evaluation of faculties' teaching Computer Science at GIMT. The 5-points scale being used in quantitative analysis is being converted to 2-points scale and is supplemented by result from the qualitative analysis.

Keywords—mixed research method; students' feedback; Likert scale

I. INTRODUCTION

Most of the colleges and universities have started taking Students Evaluation of Faculties (SEF) for accessing their performance. This is a means of maintaining the standard of teaching-learning process. There are debates as to why should the students be given the right to judge their teachers, yet majority of the people take it positively and as a chance to rectify oneself, in case it comes out to be negative. Normally, these feedbacks are taken using surveys to be answered by the students either in the middle of the semester or at the end of the semester. The survey is designed so as to extract their feelings regarding the course, clarity of the course objectives, faculty's standard of delivery, punctuality and sincerity in the class, encouragement of the students for queries and discussion and also the overall view of the student towards the course. The objective of SEF is to mine feedback from the students so as to motivate their faculties to teach and help them learn better. Faculties accept SEF results to improve their class-room performance; at the same time it is considered as a measure to judge the performance by the authorities for the employees' promotions and career enhancement programs. In some parts, it has been seen that the SEF results are given to third party for educational purposes like recruitment of new faculties and selecting graduate students for teaching awards [9]. It is also used to engage the faculties with positive feedbacks to design and review curriculums.

As the result is being used in different applications, higher authorities in the managerial level prefer to convert the result into a single grand average of the scores. This can be sometimes misleading given the nature of interaction between the faculty and the students which is quite complex in nature. Reference [3] for instance, showed that little relationship exists between learning and the evaluations which are situational and not applicable to all the individuals. It is concluded that as we measure

learning more objectively it would result to less likelihood of its relation to the evaluations.

One of the most popular scales used in questionnaire is the Likert scale developed by Renis Likert, the detail of which can be seen in [7]. This scale is used as they are relatively easier to develop and analyse. Here it may be noted that there is a separation between Likert-type and Likert-scale. A clear boundary between these two can be seen in [2]. They suggested that for Likert-type, only parametric techniques should be used while for Likert-scale parametric techniques should be used.

This study is based on the exercises done by collecting feedbacks from the students regarding their respective teachers in the Department of Computer Science and Engineering, Girijananda Institute of Management and Technology, Assam. The questionnaire used is targeted to check the instructor in terms of his/her punctuality, knowledge about the subject, sincerity, motivation in terms of discussion and queries and updating herself with the global happenings. The feedback of 20 faculties was collected from 103 students. Section II and III present the quantitative and qualitative analysis of the problem respectively. Section IV gives the experimental results while interpretation, limitation and future enhancement is presented in Section V.

II. QUANTITATIVE METHOD FOR THE PRESENT STUDY

A. Preliminaries

Mixed research method is one that mixes or combines quantitative and qualitative research techniques in the same framework [6]. This combination of the techniques is useful in situations where one technique can supplement each other in the same context. Reference [4] had specified the three phases for this kind of method:

- Research Conceptualisation*: Determining the mixed goal of the study.
- Research Planning*: Selecting the mixed sampling and research design.

c) *Research Implementation*: Collecting, analysing and interpreting.

SEF is a perfect platform where the problem can be formulated into the above three phases and the required objective can be obtained. Along with the questions a column is being added to write the remarks regarding the concern faculty. A glance into some of the remarks can prove anyone the inner feeling of the students regarding the faculty. SEF consist of 7 items, 6 of which are closed-ended (5=Strongly Agree, 4=Agree, 3=Neutral, 2=Disagree, 1=Strongly Disagree) and 1 open-ended (remark about the faculty). Table I displays part of the questionnaire.

B. Data Description

SEF is conducted towards the end of the semester. According to the college policy only students having attendance percentage of 75 and above are selected for the purpose. The data set consist of the questionnaire collected in the fall semester of 2016 by the Department of Computer Science and Engineering, GIMT.

Students are to select any of the options {Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree} as answers to six questions anonymously, followed by a remark about the faculty. The 5-point Likert scale is used for the purpose. The questions of the evaluation form are given in Table 1.

TABLE I. QUESTIONS OF THE EVALUATION FORM

SI	Questions
1	Faculty is well prepared for the class
2	Responding to students query is satisfactory
3	Faculty is punctual and regular to the class
4	Communication skill is good
5	Has knowledge of emerging technologies related to the subject
6	Cares for learning of students

C. Delusion by Representing with a Grandmean

As already mentioned, feedbacks from the students are taken in order to understand the faculties' attitude towards the students and the students' view about their teacher so that it helps the faculties to teach better. This exercise is quite complex in nature and leads to a never ending debate on whether feedback from the undergraduates are worth for a major decision making. In one hand, they are immature and demand direction; on the other hand certain deeper inside can be seen at times. Thus, the importance of such an exercise cannot be ruled out.

The common method of taking grand average is inappropriate as the variables under considerations are non-numeric in nature. Equation (1) represents the formula to calculate the grand mean:

$$grandmean(x) = \frac{1}{nm} \sum_{j=1}^m \sum_{i=1}^n x_{ij} = mean(mean_{j=1:m}(s(x_j))) \quad (1)$$

where $s(x_j) = \{x_{ij} : i=1..n\}$. Many colleges and universities use averages and standard deviations. A final grand average is calculated and awarded as an overall rating of the faculty, which is misleading. There are simple examples that prove this fact [1] and [5]. For each faculty,

the average of each question is calculated, after which a grand average of these averages is found out. Such a single valued grading is misleading. Likert type scoring is another use of ordinal, thus taking averages is incorrect. For example, let $x_1=4$ (Agree) and $x_2=2$ (Disagree). When average of the two is calculated we get $(x_1+x_2)/2=(4+2)/2=3$ (Neutral). This means that if one student agrees to a query while the other disagrees, then the result is a single observation with value "Neutral" which is not the case. Reference [5] suggests that even the use of mode would be better than the mean in situations where the remarks need to be a single value. Assuming that there is a requirement of reducing this 5-point scale to either 2 or 3 in order to make managerial decision then a better representation can be put in two different ways:

a) 3-point scale:

Agree=sum of frequencies of Strongly Agree and Agree

Neutral= frequency of Neutral

Disagree=sum of frequencies of Strongly Disagree and Disagree

b) 2-point scale:

Non-Disagree=sum of frequencies of Strongly Agree, Agree and Neutral

Disagree=sum of frequencies of Strongly Disagree and Disagree

This paper concentrates on the 5-point scale converted to 2-point scale for grading the faculties whose feedback is under consideration. This is done so by converting the actual value to percentage. *Non-Disagree* being a positive indicator of the faculty is represented as *Strength* while *Disagree* as *Weakness*.

III. QUANTIFYING QUALITATIVE INFORMATION

A. Sentiment Analysis from Remarks

The process of computationally identifying and analyzing people's opinions expressed about a product, services events etc is called sentiment analysis or opinion mining. This term came up first in [10].

Words that are used to express positive or negative sentiments are called sentiment words [8]. Examples of positive sentiment words are nice, awesome, best, amazing etc. and poor, horrible, hopeless etc. are negative sentiment words. In most cases, sentiment words are formed using adjectives and adverbs while some nouns and verbs are also used for the same. Examples of such nouns are crap, junk while verbs such as love are also used. Apart from these words sentiment phrases and idioms are also used to express sentiments. Extractions of such sentiment words are the most crucial part of this analysis.

Remarks entered in the feedback are unstructured data. Pre-processing of the data is required to check the format and character sets before Term Based Method is applied.

B. Data Structures Maintained

Two lists *pos_sentiments* and *neg_sentiments* are maintained to keep tract of the positive and negative remarks. The part content of the former is shown in Table II.

TABLE II. PART VALUES OF POSITIVE SENTIMENTS

SI	words	SI	words
1	friend	11	establish
2	know	12	like
3	satisfy	13	kind
4	perfect	14	happy
5	want	15	awesome
6	care	16	enjoy
7	help	17	favourite
8	love	18	excellent
9	know	19	cool
10	thank	20	intelligent

IV. EXPERIMENTS

The values of *strength* (Strongly Agree, Agree, Neutral) and *weakness* (Disagree, Strongly Disagree) from quantitative analysis can be plotted along with *positive remarks* and *negative remarks* from the qualitative analysis. Also, taking averages of the two components

$$\left. \begin{aligned} \text{mean_strength} &= (\text{strength} + \text{pos_remark})/2 \\ \text{mean_weakness} &= (\text{weakness} + \text{neg_remark})/2 \end{aligned} \right\} \quad (2)$$

can help in situation where single values are in demand. The calculated values of (2) are shown in Table III.

TABLE III. MEAN OF CORRESPONDING COMPONENTS

SI	mean_strength	mean_weakness
1	97.81	2.19
2	96.76	0.85
3	80.91	18.68
4	84.70	19.94
5	95.91	3.90
6	88.39	9.54
7	96.77	1.61
8	65.69	30.70
9	82.31	11.02
10	85.89	15.83
11	82.20	19.27
12	87.03	11.30
13	84.40	13.94
14	80.00	26.03
15	92.26	8.76
16	75.08	16.27
17	90.06	7.11
18	83.54	9.72
19	85.42	12.04
20	96.57	0.98

The relations between the faculties and their scores (in percentage) from the feedback, showing the similarities between the attributes of the two components is shown in Fig. 1.

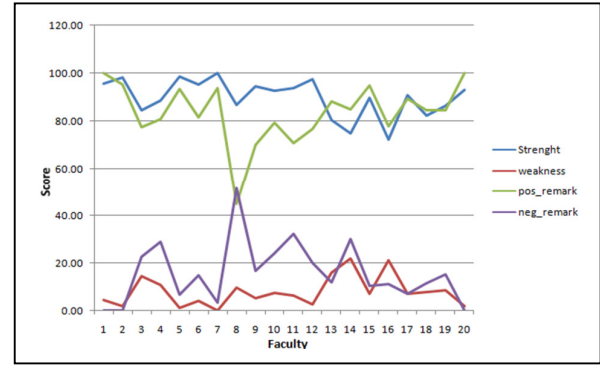


Figure 1: Comparison of Strength, Weakness, Positive Remark, Negative Remark

V. CONCLUSION AND DISCUSSION

The students' evaluation of faculties is popular among educational Institutes. Such an exercise of finding the relationship between the students and teachers in teaching and learning process has been debated for years. This study reveals some of the internal feelings of the students towards their teachers which otherwise are suppressed within themselves. It has been seen that the *strength* (Strongly Agree, Agree, Neutral) from quantitative analysis maps with positive remarks of the qualitative analysis; also the *weakness* (Disagree, Strongly Disagree) from quantitative analysis matches with the negative remarks of the qualitative analysis, which can be considered as a verification for the evaluation process.

The result thus obtained can be used to motivate the faculties to perform better where required and grant incentives to the ones who deserves for their hard work. Another implication of the feedback collection process is that the students are indirectly motivated to attend classes regularly so that they become part of the exercise in giving feedback to their teachers, as good attendance is the criteria for SEF.

One of the limitations of this study is the inability of handling complex and sarcastic sentences. The remarks from the students are assumed to be in simple sentences, which is true in many cases but the presence of complex sentences can't be ruled out. It is also true that there are sentences without sentiment words while they imply opinions. This study can be enhanced by applying text mining algorithms as a part of the qualitative analysis on the remarks submitted by the students.

REFERENCES

- [1] Adams, E. W., R. F. Fagot, and R. E. Robinson (1965, June). A theory of appropriate statistics. *Psychometrika* 30(2), 99–127.
- [2] Boone, H. and D. A. Boone (2012). Analyzing Likert Data. *Journal of Extension* 50(2).
- [3] Clayson, D.E. (2004). A test of reciprocity effect in the student evaluation of instructors in marketing classes. *Marketing Education Review*, 14(2), 11–21.
- [4] Collins, K. M. T., Onwuegbuzie, A. J., & Sutton, I. L. (2006). A model incorporating the rationale and purpose for conducting mixed methods research in special education and beyond. *Learning Disabilities: A Contemporary Journal*, 4, 67–100.

- [5] Fokoue, E. and G., Necla (2013), "Data Mining and Machine Learning Techniques for Extracting Patterns in Students' Evaluations of Instructors". Accessed from <http://scholarworks.rit.edu/article/1746>
- [6] Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14-26. doi:10.3102/0013189X033007014, <http://dx.doi.org/10.3102/0013189X033007014>
- [7] Likert, R. (1932, June). A Technique for the Measurement of Attitudes. *Archives of Psychology* 22(140), 5-55.
- [8] Liu, B. (2012). *Sentiment Analysis and Opinion Mining*, Morgan & Claypool Publishers.
- [9] Kulik, J. A. (2001). Student ratings: Validity, utility, and controversy. In M. Theall, P. C. Abrami, & L. A. Mets (Eds.), *The student ratings debate: Are they valid? How can we best use them?* *New Directions for Institutional Research*, No. 109 (pp. 9-25). San Francisco: Jossey-Bass.
- [10] Nasukawa, Tetsuya and Jeonghee Yi. (2003), Sentiment analysis: Capturing favorability using natural language processing. in *Proceedings of the KCAP-03, 2nd Intl. Conf. on Knowledge Capture*.
- [11] Turney, Peter D.(2002). Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. In *Proceedings of Annual Meeting of the Association for Computational Linguistics (ACL-2002)*.

Author Profile

Thokchom Shanta Kumar, is currently working as Associate Professor in Department of Computer Science and Engineering, Girijanjan Chowdhury Institute of Management and Technology, Assam, India. He received his Master in Computer Applications from Jorhat Engineering College, Assam, India and PhD from Himachal Pradesh University, Shimla, India. His area of research is Clustering Techniques, Data Warehousing and Visualisation. He has published book chapters, journals and conference paper in Data Mining and its applications in Education, Business and Disaster Management.

