Improving Teaching and Learning Experience in Engineering Education using Sentiment Analysis Techniques.

Wandeep Kaur¹, Vimala Balakrishnan¹, Baljit Singh²

¹Faculty of Computer Science & Information Technology, Universiti Malaya, Malaysia.
²Faculty of Mechanical Engineering, Universiti Teknologi Mara (UiTM), Shah Alam, Selangor, Malaysia.

Abstract. Students are the golden commodity when it comes to ringing in revenue for academic institution. Therefore, it is vital to ensure the opinions and feedback of students is taken seriously to ensure continuous improvement in the teaching and learning experience. In an era of digital information and rich opinionated text being easily available, it is crucial to look into different forms of data analysis from which vital information can be extracted. Sentiment and emotion analysis are one such area of research that looks to extract implicit information from written text and analyse data that would be able to provide a deeper insight compared to conventional measures. This paper is an extension of a previously conducted study of analysing emotion as well as sentiment of students’ feedback taking Thermal Engineering (MEC551) from Universiti Teknologi Mara (UiTM). Supervised learning technique was adopted and data analysed revealed students were biased towards assignment and quizzes as these would help improve their carry forwards for the subject and the preference of chapters to the exam was more for conduction and convection compared to others which had more mathematical related calculation.

1. Introduction
Evaluating teaching and learning process at any education institution is not only crucial but also vital in determining the continuous growth of said institute. This is because such feedback from students help institutes to gain an insight as to how students are perceiving the subject and helps teachers to better improve the teaching process for future reference. In an age where opinions and feedbacks are deemed important, gathering information via a Likert scale would only cause important information to be lost in translation (Joseph & Varghese, 2019). Therefore, a new approach needs to be adopted when it comes to analysing student feedback and this can be done using sentiment and emotion analysis techniques.

Majority of higher learning education adopt the student response system (SRS) at the end of each semester in order to acquire student feedback. Online learning platforms such as Coursera (www.coursera.org) and Udemy (www.udemy.com) rely on such feedback as a measuring tool in order to improve their content. Students are constantly providing feedback on the free courses provided by sharing their opinion with millions of potential students who may be looking to sign up. Such is the era of rich text that even Malaysian students are referring to online forums such as EasyUni (www.easyuni.my) to comment on their educational experiences at local colleges and universities. Such feedback can directly impact the number of future student intakes who are potential customers for such higher learning institutes.
Although sentiment analysis plays an essential role when it comes to effectively assessing student feedback (Esparza et al., 2017; Rani & Kumar, 2017), the analysis only classifies each feedback as either being positive, negative or neutral. With the availability of vast amounts of self-expressive and opinionated data, it is imperative to move forward from only classifying positive and negative sentiments to distinguishing discrete emotion that is layered within positive and negative text. This is because every positive or negative text cannot be treated as an equal (Sailunaz et al., 2018). For example, both anger and fear can be classified as negative sentiment; however, literature has shown anger is an emotion that provokes a positive outcome as people are motivated to make a change but fear on the other hand instils pessimistic responses (Sailunaz et al., 2018).

This paper is an extension of a preliminary study on student feedback (Kaur et al., 2017) that is looking to not only classify feedback extracted from students taking Thermal Engineering (MEC 551) course from the Faculty of Mechanical Engineering of University Teknologi Mara (UiTM) according to sentiment but also emotion with respect to Plutchik (2003) wheel of emotion. The extracted posts were subjugated to supervised machine learning algorithms such as Naïve Bayes (NB) and Support Vector Machine (SVM), implemented using Weka. The algorithms are compared for evaluation purpose using accuracy, precision and recall measures. The remainder of this paper is organized as follow; the next section discusses the literature followed by methodology, results, discussion and finally conclusion.

2. Literature Review

This section looks to provide a brief discussion on sentiment and emotion analysis followed by a discussion on various related work within the domain of student feedback analysis.

2.1. Sentiment Analysis

According to Rani et al. (2017), teaching and learning outcomes can be evaluated directly (assessment of students’ work such as assignments, quizzes, exams etc.) and indirectly (students’ learning experience and teaching quality). Sentiment and emotion analysis are a form of indirect assessment looking to gather information regarding student learning experience in order to improve teaching quality from the feedback provided.

A text can be classified as positive with respect to the number of positive related words extracted from it and negative similarly (Nithya & Maheswari, 2014). A neutral sentiment is assigned to a text when the algorithm is unable to detect positive or negative words from it. For example:

“The quiz on Chapter 3 had too many confusing questions”

“The open book exam today gave me hope to pass!”

As it can be observed from above, the first sentence clearly shows a negative sentiment towards the questions asked in the quiz (implied from the word confusing) while the second sentence shows a positive sentiment towards the given exam (implied from the word pass).

There are three approaches to conducting a sentiment analysis study on text; machine learning approach, lexicon-based approach and hybrid (Figure 1). The machine learning approach adopts machine learning algorithms in order to classify text while lexicon-based approach is more laborious with the creation of manual lexicon dictionaries. The hybrid approach is a combination of both machine learning and lexicon-based approaches. This paper looks into adopting the supervised learning
2.2. Emotion Analysis

Emotion analysis is an extension of sentiment analysis and defined as a linguistic process of identifying emotions expressed in written text (Sailunaz et al., 2018). Yadollahi et al. (2017) found a strong correlation between sentiment and emotion where an emotion may lead an individual to build an opinion on something and similarly an opinion may invoke an emotion in others. Furthermore, an extracted text may hold contradicting opinions and emotion as example:

“My supervisor thinks it’s a good decision for me to take on the scholarship overseas although he is sad to see me go”

The example above shows a positive sentiment towards the possibility of pursuing studies overseas, however the emotion related to it shows sadness. Additionally, a text may have a sentiment attached to it yet not have an emotion. For example:

“I have not studied for the mid-terms”
“The results of the experimentation are disappointing”

As it can be observed, the first sentence carries a negative sentiment but there is no emotion attached to it. Similarly, the second sentence shows the same negative sentiment but the emotion detected is disgust as perceived by the usage of the word disappointing.

Literature has identified four main approaches to emotion detection (Figure 2). The keyword-based approach looks into the grammatical aspect of a text where Noun, Verb, Adjective and Adverb (words that have been identified as words that carry emotions) are matched to emotion dictionaries (Yadollahi et al., 2017). The machine learning, lexicon-based and hybrid approaches as similar to sentiment analysis approaches explained in section 2.1. This paper will look into adopting the lexicon-based approach adopting the NRC Emotion Lexicon for the purpose of emotion detection (Mohammad & Turney, 2013).
3. Methodology
This section will discuss the methodology adopted (Figure 3). Extracted feedback is classified according to sentiment polarity (positive, negative and neutral) and emotion detection with respect to Plutchik (2003) wheel of emotion (anger, anticipation, disgust, fear, joy, sadness, surprise and trust) using the NRC Emotion Lexicon (Mohammad et al., 2013). Each phase of the methodology will be discussed in detail in the following sub-sections.

3.1. Data Collection
A Google form was created with the intention of gathering feedback from students who registered for Thermal Engineering (MEC 551) course from the Faculty of Mechanical Engineering of University Teknologi Mara (UiTM). These students were a mixed bunch of Year 2 to Year 5 students who were attempting this subject between the first time to the third time. The questions on the feedback form were designed after a brainstorming session between the lecturers who were teaching said subject, verbal conversations with students attempting said subject as well as researchers working on sentiment analysis within the education domain. The google form link was distributed to the students during the last week of the lecture. After duration of 6 weeks, 121 students feedback forms were received. The cut-off point of the collection was decided upon as the number of feedbacks received started to dwindle and the number remained stagnant for a week. All feedbacks were treated with utmost confidentiality as no collection on names, student matric number or emails were collected during this phase in order to protect students’ anonymity.
3.2. Data Preprocessing
The standard data pre-processing steps were adopted in this paper which include tokenization, stemming, normalization and changing all sentences to lower case. Apart from the aforementioned, the dataset also had to be cleaned from possible human typo errors, elimination of answers that were too short (less than 3 words in length) and removal of content that was irrelevant to analysing sentiment and emotions (punctuations, stop words such as the, a, an etc.). Although, this dataset came clean from errors, some of the feedbacks had to be removed as they were too short to perform analysis over. This brought the final count of the dataset to 103.

3.3. Sentiment & Emotion Identification
Two of the most widely used sentiment classification algorithm were used for this paper; Naïve Bayes (NB) and Support Vector Machine (SVM) (Sun et al., 2017). The algorithm was built using Weka 3.7. The cross-validation operator was set at 10-fold where the algorithm then proceeds to automatically split the fed dataset between testing and training. The corpus uploaded to the algorithm was analysed according to three major aspects namely feedback on the teaching and learning experience (includes classroom ambience, availability of materials online and feedback on lecturer), syllabus of the subject (breakdown on the chapters of the subject) and feedback on other student assessment measures (quizzes, mid-terms, exams and assignments). The training and testing phase of the algorithm was set manually as 80% for training and 20% for testing purposes.

For emotion identification on the other hand, Mohammad et al. (2013) built the NRC Emotion Lexicon by assigning each word in the lexicon dictionary to an emotion vector. Therefore, when the algorithm identifies a word in the student feedback, the corresponding emotion in the lexicon list is returned. For example;

Sample sentence 1: “The quiz was so easy!”

Would return the following emotion vectors:

| Table 1. Emotion Vector for Sample Sentence 1 |
|---------------------------------------------|
|     | Anger | Anticipation | Disgust | Fear | Joy | Sadness | Surprise | Trust |
|-----|-------|--------------|---------|------|-----|---------|----------|-------|
|     | 0     | 0            | 0       | 0    | 1   | 0       | 0        | 0     |

The use of this lexicon was founded under the assumption that a text would be able to have both negative and positive emotion as well as showcase more than one emotion at a time. In case of such occurrence, the algorithm would then work to find the aspect upon which the emotion is attached to and assign the final emotion accordingly. For example:

Sample sentence 2:
“I was so scared going to the exam and I really thought I was going to fail the subject but thank God I managed to scrape through and pass”

The above sample would return the results as shown in Table 2 and based on the count, and the algorithm would identify the emotion fear (related to the keyword “scared” and “fail”) is related to the aspect of exam.
Table 2. Emotion Vector for Sample Sentence 2

| Emotion | Value |
|---------|-------|
| Anger   | 0     |
| Anticipation | 0   |
| Disgust | 0     |
| Fear    | 2     |
| Joy     | 1     |
| Sadness | 0     |
| Surprise| 0     |
| Trust   | 0     |

4. Results and Discussion

This section will look into the results obtained during the phase of this study. Results will be displayed and discussed according to three features: chapter modules (breakdown on each chapter), feedback on other student assessment (exams, assignment, quizzes etc) and teaching evaluation (feedback on lecturer as well as classroom ambience). Results will be displayed for both sentiment and emotion respectively. The evaluation metrics used for this paper was pre-built within Weka itself, hence a direct comparison of f-measure scores of both Naïve Bayes (NB) and Support Vector Machine (SVM) are shown in Table 3. All results are displayed in percentage form.

Table 3. Result comparison NB and SVM.

| Feature                  | NB   | SVM  |
|--------------------------|------|------|
| Chapter Modules          | 94.5 | 92   |
| Assessment Measures      | 87.1 | 91   |
| Teaching Evaluation      | 75   | 70   |

From the table, it can be noted that SVM generally produced the best results for all features except teaching evaluation (70%). This could be possibly because the teaching evaluation feedback was wordier compared to the other features and NB has shown to perform better when classifying longer text compared to SVM (Sun et al., 2017).

4.1. Chapters Modules Results

Figure 4 shows sentiment analysis for each chapter thought in Thermal Engineering (MEC 551). As the chapters progressed within the subject, the number of negative feedbacks compared to positive feedbacks started to increase. Figure 5 shows the breakdown of emotion attached to said subjects.

Figure 4. Sentiment classification based on chapter modules

Only negative emotions (anger, disgust, fear and sadness) results are displayed in Figure 5 as these emotions were more apparent when it came to classification process of this feature. Students were mostly
fearful of the last few chapters as they claim it involved more calculation hence following the subject was difficult. Additionally, data also revealed students were more hesitant when it came to the last chapter because it involves chemistry knowledge which they do not have much exposure to.

![Figure 5](image-url)

**Figure 5.** Emotion detection based on chapter modules

### 4.2. Assessment Measures

Figure 6 shows the emotion classification for assessment. Students favoured more assignments compared to tests and quizzes yet from the feedback they claimed the number of tests should be increased as this would help them obtain better carry forwards as they headed to finals.

![Figure 6](image-url)

**Figure 6.** Emotion detection for assessment measures

### 4.3. Teaching Evaluation

When it came to evaluating lecturers, students generally had a very positive feedback. Although they have mentioned that they would prefer more time spent during lectures and tutorials so that they would be able to clear doubts that they would have within the class. Additionally, students did find the learning environment not to be conducive as the classrooms and lecture halls all were facing air conditioning issues. Figure 7 shows the sentiment classification of this feature.
5. Conclusion
This study was an extension from a previously conducted study. More aspects from the feedback were extracted for analysis purpose and the format of collecting feedback was also changed to accommodate the objective of this study. Machine learning algorithms (Naïve Bayes and Support Vector Machine) were used for sentiment classification purpose and NRC Emotion lexicon was used for emotion analysis built using Weka 3.7 platform.

The results found in this phase of study met the objective of mining emotions as well as sentiment. This can be seen from the results discussed that students were more expressive when it came to providing feedback. Hence the results obtained can be used to provide an insight on how to better improve the teaching and learning experience in order to improve overall student quality.

For future research, this study would like to conduct a content analysis study to analyse the form of feedback that came through from the students. There are also intentions to look into how this study can be expanded to other subjects thought within the faculty.

References
[1] Esparza, G. G., de-Luna, A., Zezzatti, A. O., Hernandez, A., Ponce, J., Álvarez, M., . . . de Jesus Nava, J. (2017). A sentiment analysis model to analyze students reviews of teacher performance using support vector machines. Paper presented at the International Symposium on Distributed Computing and Artificial Intelligence.
[2] Joseph, G., & Varghese, V. (2019). Analyzing Airbnb Customer Experience Feedback Using Text Mining Big Data and Innovation in Tourism, Travel, and Hospitality (pp. 147-162): Springer.
[3] Kaur, W., Balakrishnan, V., & Singh, B. (2017). Social Media Sentiment Analysis of Thermal Engineering Students for Continuous Quality Improvement in Engineering Education. Journal of Mechanical Engineering, Special Edition (4(1)), 263-272.
[4] Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word–emotion association lexicon. Computational Intelligence, 29(3), 436-465.
[5] Nithya, R., & Maheswari, D. (2014). Sentiment analysis on unstructured review. Paper presented at the Proceedings - 2014 International Conference on Intelligent Computing Applications, ICICA 2014.
[6] Plutchik, R. (2003). Emotions and life: Perspectives from psychology, biology, and evolution: American Psychological Association.
[7] Rani, S., & Kumar, P. (2017). A sentiment analysis system to improve teaching and learning. Computer, 50(5), 36-43.
[8] Sailunaz, K., Dhaliwal, M., Rokne, J., & Alhajj, R. (2018). Emotion detection from text and speech: a survey. Social Network Analysis and Mining, 8(1), 28.
[9] Sun, S., Luo, C., & Chen, J. (2017). A review of natural language processing techniques for opinion mining systems. Information Fusion, 36, 10-25.
[10] Yadollahi, A., Shahraki, A. G., & Zaiane, O. R. (2017). Current state of text sentiment analysis from opinion to emotion mining. ACM Computing Surveys (CSUR), 50(2), 25.