Using Sentiment Analysis to Explore Student Feedback: A Lexical Approach

https://doi.org/10.3991/ijet.v18i09.38101

Rdouan Faizi(✉)
ENSIAS, Mohammed V University in Rabat, Rabat, Morocco
rdouan.faizi@um5.ac.ma

Abstract—Given the increasing abundance of online courses over the last couple of years, new forms of student feedback, which are less frequently used by teachers, have been generated in massive amounts. Nonetheless, extracting and processing this student generated content manually is costly and time consuming. In this respect, our objective in this paper is to propose a lexical-based approach that can predict the underlying sentiments of each student review, thus, enabling teachers to assess to what extent students are satisfied with the online learning resources and teaching practices. To enhance the performance of the proposed approach, a new education sentiment lexicon was built and incorporated into the model. After its implementation on a dataset that was extracted from the Web, this sentiment analysis lexical approach has proven to correctly predict the sentiment polarities of the great majority (i.e. 86.45%) of student feedback.

Keywords—education, student feedback, sentiment analysis, lexicon

1 Introduction

Collecting information from students about their experience as learners has always been part and parcel of any educational system. Though there are various methods to assess students' learning and teaching strategies, student feedback is considered as one of the most effective and most reliable methods that are used to assess the effectiveness and quality of both learning and teaching [1, 2]. By retrieving and analyzing this feedback, teachers can get valuable insights into how students are learning as well as how they are engaged with course materials.

Feedback is traditionally collected by administering anonymous surveys to students using open-ended questions, or a mixture of open-ended and closed-ended questions. In addition to using questionnaires, feedback can also be gathered through class discussion, especially in cases in which the class is relatively small. In recent years, different forms of student feedback, which are less frequently used by educators and academic institutions, have emerged. In fact, due the growing proliferation of learning resources on the Internet over the last decade, the online platforms hosting these educational materials are not only an environment where learners can access courses, but also an open space where they can post comments as well as questions or answers to various issues or learning problems that need clarification irrespective of time and space constraints [3, 4]. These new forms of feedback, thus, provide students with a 24/7 opportunity to make their voices heard and reveal information which may not have been available in standard surveys [5, 6].

However, given the voluminous quantities of student feedback that are available online, extracting and processing this data manually is costly and time consuming. Thus, given the fact that student feedback can arouse positive emotions or feelings such as satisfaction and happiness, or negative ones such as sadness or anger [7], our objective in the present article is to propose a sentiment analysis approach that enables teachers and educators to automatically identify the perceptions and attitudes of students towards learning materials and teaching strategies.

Sentiment analysis refers to the application of Natural Language Processing and text analysis techniques so as to extract subjective information from textual data and to determine whether the sentiments expressed in the text are positive, negative or neutral [8]. In recent years, given the massive amounts of data that have been generated online, there has been a steady increase in interest from businesses and researchers in sentiment analysis and its application in informing decision making. Sentiment analysis applications have already been implemented in a variety of sectors. Nevertheless, one of the domains in which these systems has recently been gaining ground is education [9, 10, 11, 12, 13].

The rest of this article is organized as follows. Section Two reviews some of the recent research works that have explored student feedback using sentiment analysis. Section Three presents our lexicon-based approach for analyzing the sentiments expressed in student feedback. Section Four presents and discusses the results related to the implementation of our approach. Finally, Section Four summarizes the findings of this paper.

2 Related work

Given the growing amounts of student feedback generated on social media and on e-learning platforms, numerous research studies have recently been carried to extract value from this data using different sentiment analysis techniques. In this respect, in an attempt to assess teacher performance before, during and after the COVID pandemic, Jiménez et al. [14] submitted questionnaires to students from a Brazilian university. Based on this student feedback, the authors used the Bidirectional Encoder Representations from Transformers (BERT), which is a pre-trained language model [15], to identify the students' sentiments expressed in the comments. The results of their research study revealed that the perceptions of students towards the performance of teachers online (during the pandemic) were positive, which clearly proved that the performance of faculty in distance education was better than in in-person courses.

In their research work Jimmy & Prasetyo [16] applied sentiment analysis to classify undergraduate student feedback on teaching conduct. For this purpose, they collected data from Indonesian higher education students and empirically evaluated the performance of three sentiment analysis classifiers, namely Naive Bayes (NB), Support Vector Machine (SVM), and Decision Tree. Amongst these models, SVM was found to perform the best, in terms of accuracy, in classifying the sentiments of student feedback.

For their parts, Almosawi & Mahmood [17] built a lexical-based sentiment analysis approach to determine the polarity of students' feedback. In this vein, they collected data by using an open-ended questionnaire and created a dictionary of opinions in the field of higher education. After the implementation of their lexicon-based approach to identify the polarity of student feedback, the findings have revealed that it scored (82%) for the positive class and 40% for the negative.

In their research study, Umair et al. [18] undertook a comparative analysis of student feedback before and during the COVID pandemic. Student feedback was collected by means of WhatsApp and Google forms and then fed into two supervised machine learning algorithms, namely NB and SVM for classification. A comparison of the performance of both models on the dataset showed that SVM works best for text polarity classification. It was also found out that the online learning adopted during the pandemic is associated with more negative reviews as compared to the blended teaching mode.

To assess the effectiveness of teaching and learning, Mabunda et al. [19] developed a sentiment model to analyze the feedback provided by students. After training machine learning models such as SVM, Multinomial NB, Random Forest, K-Nearest Neighbors and Neural Networks on the student dataset they obtained from Kaggle, it was noted that the latter model was more effective in predicting the sentiments of students towards teaching practices.

To gain insights into the opinions of Malaysian university students regarding the quality of e-learning systems, Baragash et al. [20] analyzed the data they collected from Twitter. After data preparation, the authors used a sentiment analysis-based machine learning model (i.e. SVM) to classify students' opinions and RapidMiner to determine the sentiment of tweets and the accuracy of the algorithm. The findings of their study demonstrated that most students have a positive opinion about e-learning systems in Malaysian universities as 65% of reviews were classified as positive.

3 Proposed approach

Sentiment Analysis is carried out by making use of various techniques. Yet, these can be roughly divided up into two major categories, namely machine learning and lexicon based approaches. In the first category, a machine learning algorithm is trained to classify sentiments based on both the words and their order in the sentence. However, the success of this set of models greatly depends on the quality of the training dataset as well as on the algorithm used.

On the other hand, the lexicon-based approach is based on extracting and computing the polarities of sentiment words by using a sentiment lexicon. Each word in this sentiment lexicon has a score that represents its polarity, namely positive, negative or neutral and each document is scored by aggregating the sentiment scores of all the terms in that document.

Since the annotation of datasets on which machine learning models should be trained and tested requires a lot of efforts, we opted for VADER (*Valence Aware Dictionary and sEntiment Reasoner for Sentiment Analysis*) for the sentiment classification of student feedback [21]. VADER is a lexicon and rule-based sentiment analysis tool that is specifically fined-tuned to recognize the sentiments expressed in web-based media [22]. It makes use of a dictionary that maps lexical features or words to emotion intensities or sentiment scores. A word can, thus, be annotated as either positive or negative depending on its semantic orientation. In addition to calculating the polarity (positive/negative) of an input text, this rule based sentiment classifier can also quantify the intensity or strength of the emotion the text has on the basis of the word order and sensitive relationships between the terms. The underlying sentiment of a given review or document, is then computed by the compound score. This metric sums the total of lexicon ratings which have been normalized to be between −1 (most extreme negative) and +1 (most extreme positive). The architecture of the proposed model is illustrated below.

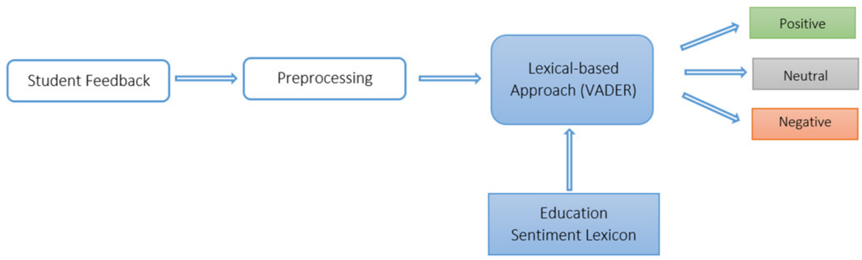

Fig. 1. Architecture of the proposed lexical model

However, though over 7500 tokens are listed in VADER dictionary, there are many entries that are not available in this lexicon. In fact, words such as '*basic*', '*enroll*', '*unenroll*', '*mediocre*', '*elementary*' and '*quit*', which are likely to be used in a comment posted by a student and have the potential to affect the sentiment score of the feedback, are absent as entries. To make up for this limitation, we built an education sentiment lexicon (cf. Figure 1 above) that is mainly composed of words that are likely to appear in student feedback together with their sentiment scores. For instance, words such as '*enroll*' and '*unenroll*' are assigned +1 (i.e. positive) and −1 (i.e. negative) sentiment scores, respectively. If such tokens are part of the student feedback but not added in the lexicon, they would be considered as neutral and so is the feedback.

4 Experiments and results

To test the performance and validity of the proposed model, it was implemented on student feedback that was extracted from Coursera, which is an e-learning platform that hosts hundreds of courses in a variety of disciplines. This dataset is mainly composed of reviews that leaners posted about eight information technology courses or tutorials that they have been enrolled in together with the titles of these courses. However, the titles of these online courses have been replaced with new IDs in the present study.

Before using this dataset, it should be pre-processed. In actual fact, since raw data is usually noisy given that it often contains useless information, which is likely to affect the performance and prediction of the model, the retrieved student feedback had to undergo preprocessing. For this purpose, reviews that are not written in English were deleted. Unnecessary characters such as numbers, hyperlinks, tags were also discarded. However, preprocessing tasks such lowercasing, punctuation and stopwords removal have been ignored in this particular study. This is simply because VADER tends to take capitalization such as "GOOD" vs. "good" and punctuation marks as "!" into consideration when assigning sentiments. Furthermore, as stopwords such as negation particles like "*not*", and degree modifiers (e.g. "*very*") have been found out to be quite helpful when it comes to identifying negative emotions, such items have not been subject to pre-processing.

After this slight pre-processing, the total dataset composed of a total of 3000 reviews were fed into the proposed lexical model to see how performant the latter is in predicting the underlying sentiment that each review exhibits. The implementation of the model on the dataset yielded the following results.

Table 1. A sample of reviews and their assigned sentiments

	Reviews	Courses	Clean_Reviews	Score	Compound	Sentiment
0	This course gave me a great experience. I am k...	C-1	This course gave great experience. I keen lear...	{'neg': 0.0, 'neu': 0.625, 'pos': 0.375, 'comp...	0.7650	Positive
1	The subject is very interesting but the teachi...	C-1	The subject interesting teaching not good. The...	{'neg': 0.165, 'neu': 0.743, 'pos': 0.092, 'co...	-0.7183	Negative
2	Compactly and concisely organized to show the ...	C-1	Compactly concisely organized show application...	{'neg': 0.0, 'neu': 1.0, 'pos': 0.0, 'compound...	0.0000	Neutral
3	A repetition of other courses in the track, go...	C-1	A repetition course track, good reinforcement ...	{'neg': 0.168, 'neu': 0.611, 'pos': 0.221, 'co...	0.1779	Positive
4	Everything is very good, I enjoy this course a...	C-1	Everything good, I enjoy course I learned lot....	{'neg': 0.0, 'neu': 0.336, 'pos': 0.664, 'comp...	0.8713	Positive

As Table 1 clearly shows, each student review is associated with four values, namely *neg, neu, pos*, which provide a proportion of the text that counts as negative, neutral, and positive, respectively. The fourth value (i.e. compound) is the normalized sum of the first three values. The first review, for instance, has no negative information (neg = 0), but it has some positive and neutral tones (pos = 0.375 and neu = 0.625). Thus, since the compound score of the review is ≥ 0.05 (i.e. 0.765), the overall sentiment of this comment is positive.

A thorough sentiment classification of the dataset reveals that a large portion of the reviews posted by learners about courses are positive. This is illustrated by Figure 2:

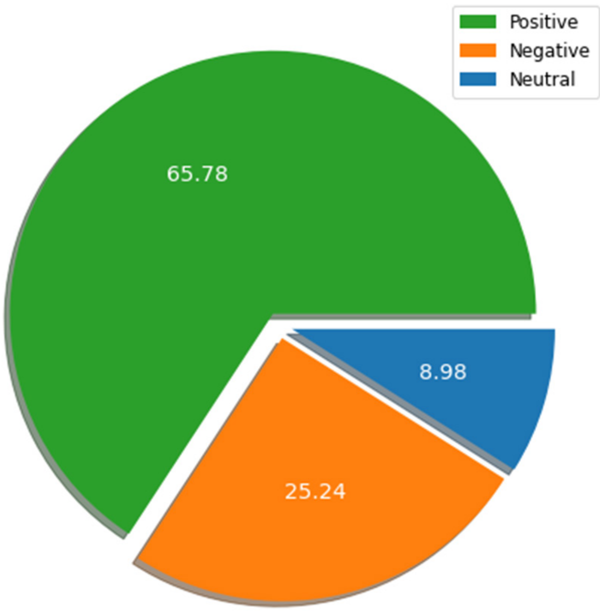

Fig. 2. Overall sentiment classification of student feedback

As Figure 2 shows, 65.78% of the whole feedback is positive, which explicitly demonstrates that positive words outweigh negative terms in these instances. However, while about a quarter of these reviews correlates with negative sentiments given the omnipresence of negative words, the minority (i.e. 8.98%) of the comments is characterized by the absence of sentiment words or include approximately the same ratio of positive and negative items. Accordingly, they are classified as neutral.

The proposed sentiment classifier does not only provide a global overview of how positive or negative student feedback generally is, but it is also capable of identifying the ratio of positive, negative and neutral sentiments that are associated with specific courses. Consider Figure 3.

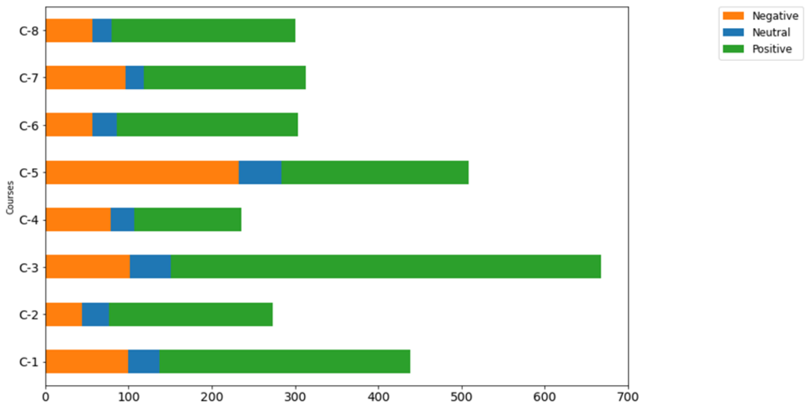

Fig. 3. Ratio of sentiment polarity per course

A close look at Figure 3 shows that apart from Course C-5 in which the majority of student feedback is negative, all other courses are characterized, with varying degrees, by the predominance of positive reviews. Through this feedback, faculty can spot what parts of a given course their students are good at and in what particular areas or issues the students are having learning problems.

To see to what extent the proposed lexical approach can predict the underlying sentiments of new student feedback, we extracted an additional dataset (i.e.1500 reviews) from Coursera and manually labelled each review as positive, negative or neutral. After going through the same pre-processing phases outlined above, the resulting reviews were fed into the proposed lexical approach. To evaluate the performance of the model, we used accuracy as an evaluation metric. This is illustrated in Table 2 given below.

Table 2. Performance of the proposed approach

Lexical Approach	Accuracy Score
Without Education Sentiment Lexicon	77.65%
With Education Sentiment Lexicon	86.45%

As the table shows, the lexical approach achieved an accuracy score of 77.65% in predicting the correct sentiment for the overall student feedback. However, since the number of lexical words on which VADER is based is very limited, we opted for the construction of an education sentiment lexicon and for its incorporation in the system. The integration of this dedicated lexicon has improved the accuracy of the approach to 86.45%. Such good results clearly demonstrate the outstanding performance of the proposed approach over the baseline model.

Taking into consideration the findings above, it is quite clear the proposed approach can help instructors get the perceptions and opinions of students towards the learning resources, teaching strategies as well as towards the teacher's performance [23]. In fact, based on the sentiment classification of feedback performed by the proposed model, teachers can monitor the learning experience of their students and to discover and address their concerns. This can include the learning difficulties they have encountered, the things to change, add or keep in the course together with the strengths and weaknesses that students perceive in teaching practices along with potential areas for improvement.

Moreover, as opposed to traditional approaches which collect and analyze student feedback at the end-of-course evaluations or at the end of the semester or year, that is, at a time in which the chances of enhancing the course materials and learning have passed, this sentiment analysis classifier can collect and analyze student feedback in real-time or near-real-time and can enable faculty to act on the feedback in a timely manner. Accordingly, it does not only allow instructors to adjust resources and teaching methods while the course is still being taught, but it also enables current learners to have answers to questions that they have asked and to reap the benefits of the suggestions they have made. However, this does in no way means that instructors should make all of the suggested changes, but they can acknowledge or assure the students that they have heard their voice.

5 Conclusion

Given the growing amounts of student generated content on the Web, the objective of this paper was to propose a sentiment analysis lexical-based approach that can classify this student feedback as being positive, negative or neutral. To improve the performance of the proposed approach, an education sentiment lexicon was constructed and integrated into the model. After the implementation of our approach on a dataset that was extracted from Coursera, it was revealed that the sentiments of the great majority of students' reviews were correctly predicted. This lexical approach can, thus, enable educators to assess the degree of satisfaction or dissatisfaction that students feel towards learning resources and teaching practices.

6 References

[1] Rollett, W., Bijlsma, H., & Röhl, S. (2021). Student feedback on teaching in schools: Using student perceptions for the development of teaching and teachers. https://doi.org/10.1007/978-3-030-75150-0

[2] Nair, C. S., & Mertova, P. (Eds.). (2011). Student feedback: The cornerstone to an effective quality assurance system in higher education. Elsevier. https://doi.org/10.1533/9781780632582

[3] Faizi, R. (2018). Moroccan higher education students' and teachers' perceptions towards using Web 2.0 Technologies in language learning and teaching. *Knowledge Management & E-Learning, 10*(1), 86–96. https://doi.org/10.34105/j.kmel.2018.10.005
[4] Faizi, R. (2018). Teachers' perceptions towards using Web 2.0 in language learning and teaching. *Education and Information Technologies*, *23*(3), 1219–1230. https://doi.org/10.1007/s10639-017-9661-7
[5] Shah, M., Pabel, A., & Martin-Sardesai, A. (2019). Assessing Google reviews to monitor student experience. *International Journal of Education Management, 34*(3), 610–625. https://doi.org/10.1108/IJEM-06-2019-0200
[6] Raman, R., & Nedungadi, P. (2020). Adoption of Web-Enabled Student Evaluation of Teaching (WESET). *International Journal of Emerging Technologies in Learning (iJET), 15*(24), pp. 191–207. https://doi.org/10.3991/ijet.v15i24.17159
[7] Röhl, S., Bijlsma, H., & Rollett, W. (2021). The Process Model of Student Feedback on Teaching (SFT): A Theoretical Framework and Introductory Remarks. In *Student Feedback on Teaching in Schools* (pp. 1–11). Springer, Cham. https://doi.org/10.1007/978-3-030-75150-0_1
[8] Liu, B. (2020). Sentiment analysis: Mining opinions, sentiments, and emotions. Cambridge university press. https://doi.org/10.1017/9781108639286
[9] El Fkihi, S., Ezzahid, S. S., Faizi, R. & Chiheb, R. (2018). "Formative Assessment in the Era of Big Data Analytics," Proceedings of the 32nd International Business Information Management Association (IBIMA), ISBN: 978-0-9998551-1-9, 15-16.
[10] Faizi, R., & El Fkihi, S. (2022). A Sentiment Analysis Based Approach for Exploring Student Feedback. In *International Conference on Innovative Technologies and Learning* (pp. 52–59). Springer, Cham. https://doi.org/10.1007/978-3-031-15273-3_6
[11] Faizi, R. (2021). A Sentiment-Based Approach to Predict Learners' Perceptions Towards YouTube Educational Videos. In *International Conference on Innovations in Bio-Inspired Computing and Applications* (pp. 549–556). Springer, Cham. https://doi.org/10.1007/978-3-030-96299-9_52
[12] Lampropoulos, G., & Keramopoulos, E. (2022). Virtual reality in education: A comparative social media data and sentiment analysis study. *International Journal of Recent Contributions from Engineering, Science & IT (iJES)*, *10*(03), pp. 19–32. https://doi.org/10.3991/ijes.v10i03.34057
[13] Niu, M. (2022). Classification of learning sentiments of college students based on topic discussion texts of online learning platforms. *International Journal of Emerging Technologies in Learning (iJET)*, *17*(24), pp. 42–56. https://doi.org/10.3991/ijet.v17i24.35951
[14] Jiménez, H. G., Casanova, M. A., Finamore, A. C., & Simões, G. (2021). Sentiment analysis of student surveys—A case study on assessing the impact of the COVID-19 pandemic on higher education teaching. *International Educational Data Mining Society*.
[15] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
[16] Jimmy, & Prasetyo, V. R. (2022). Sentiment analysis on feedback of higher education teaching conduct: An empirical evaluation of methods. In *AIP Conference Proceedings* (Vol. 2470, No. 1, p. 050006). AIP Publishing LLC. https://doi.org/10.1063/5.0080182
[17] Almosawi, M. M., & Mahmood, S. A. (2022). Lexicon-based approach for sentiment analysis to student feedback. Webology (ISSN: 1735-188X), 19(1).
[18] Umair, M., Hakim, A., Hussain, A., & Naseem, S. (2021). Sentiment analysis of students' feedback before and after COVID-19 pandemic. *International Journal on Emerging Technologies*, *12*, 177–182.

[19] Mabunda, J. G. K., Jadhav, A., & Ajoodha, R. (2021). Sentiment Analysis of Student Textual Feedback to Improve Teaching. In Interdisciplinary Research in Technology and Management (pp. 643–651). CRC Press. https://doi.org/10.1201/9781003202240-100
[20] Baragash, R., Aldowah, H., & Umar, I. N. (2022). Students' perceptions of E-learning in Malaysian Universities: Sentiment analysis based machine learning approach. *Journal of Information Technology Education: Research*, *21*, 439–463. https://doi.org/10.28945/5024
[21] Hutto, C., & Gilbert, E. (2014, May). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the international AAAI conference on web and social media* (Vol. 8, No. 1, pp. 216–225). https://doi.org/10.1609/icwsm.v8i1.14550
[22] Hossain, M. S., Rahman, M. F., & Uddin, M. K. (2022). Analyzing and Predicting Learner Sentiment Toward Specialty Schools Using Machine Learning Techniques. In *Applying Data Science and Learning Analytics Throughout a Learner's Lifespan*. IGI Global. https://doi.org/10.4018/978-1-7998-9644-9.ch007
[23] Esparza, G. G., de-Luna, A., Zezzatti, A. O., Hernandez, A., Ponce, J., Álvarez, M., … & Jesus Nava, J. D. (2017, June). A sentiment analysis model to analyze students reviews of teacher performance using support vector machines. In International Symposium on Distributed Computing and Artificial Intelligence (pp. 157–164). Springer, Cham. https://doi.org/10.1007/978-3-319-62410-5_19

7 Author

Rdouan Faizi received his Ph.D. from Mohammed V University in Rabat. He is currently a Full Professor at the Higher National School of Computer Science and Systems Analysis (ENSIAS), Mohammed V University, Rabat, Morocco. His research interests focus on Natural Language Processing and E-learning.

Article submitted 2023-01-15. Resubmitted 2023-03-02. Final acceptance 2023-03-03. Final version published as submitted by the authors.