A Data Mining Framework for Analyzing Students' Feedback of Assessment

Zainab Mutlaq Ibrahim, Mohamed Bader-El-Den, Mihaela Cocea
(zainab.mutlaq-ibrahim, mohamed.bader, mihaela.cocea)@port.ac.uk

University of Portsmouth, Lion Terrace, PO1 3HE, UK

Abstract. Assessment constitutes a fundamental part of an academic learning process due to its importance in testing students gaining knowledge and finalizing their grades. This study aims to develop a data mining based framework for analyzing students' assessment feedback that will be obtained from social media sites and/or text feedback. The study consists of three stages: The first stage is to build a model that automatically detect the polarity of student feedback using sentiment analysis methods. The second stage is to build a model that automatically classify issues of assessment. And finally, test the correlation between issue(s) and students' performance. The research uses different popular algorithms for text classification to analyze students' feedback of assessment to enhance learning process.

Keywords: Assessment· Decision Tree· Machine learning algorithms· Naive Bays· Random Forest· Sentiment analysis· Support Vector Machines.

1   Introduction

Analyzing students' feedback of assessment can point out issues they may have in accomplishing its components. Many students squander time in completing assessment trying to figure out what the thought processes of the marker might be, what s/he wants to hear or read instead of developing their own skills and understanding in evaluating what the assessment component asked for [1], that should remind the assessment author to clarify what exactly the assessment asked for. More examples of assessment issues are: length, format, validity, reliability, late rectification, no remedy, issues with teaching and curriculum, and disruptive.

Massive online open courses(MOOCs) bring new challenges to the learning process in general and assessment in particular as MOOC unit includes a huge number of candidates from different cultures and backgrounds who use different languages and may be different accents and jargon of the same languages. This will need extra effort and different techniques to design and simplify the assessment to test the right knowledge and skills.

To understand more about assessment the following section is to highlight its main aspects:

1.1   Assessment

Assessment is defined as all procedures that evaluate student's knowledge, understanding, abilities, or skills according to the quality assurance agency for higher education [2].

Students receive feedback from their tutors during the semester and at the end of it. The latter is called suumative assessment that sums up all student achievements and leads towards awarding the final grade, while the first, which is given to student during the semester with the main aim to recognize students strengths and weaknesses to guide them accordingly [3].

Assessment types include written and oral exams, essays, reports, portfolios, presentation, projects, posters, theses, and many more. All form of assessment have general advantages and disadvantages[1]. However, advantages to student A can be disadvantage to student B. For example essays shows depth of learning which suits a good writer who has a great writing skills, but not to other students who don't have good writing skills. Struyven and Black [4,5] revealed that students' perceptions of assessment remarkably affected their approaches to learning and studying, this means that different design and style of assessment would guide them in choosing the right method of studying to in achieve better results.

Black and William[5] added that assessment influences learning in three ways: provides motivation to students; highlights the important part of the curriculum; and helps students to evaluate and judge the effectiveness of their learning.

Despite the importance of Students' feedback on assessment as a constructive act of learners in higher education, there are relatively limited studies and reviews that considering students' perspectives [4].

Institutions seek students' feedback using questions-based surveys in which choices are provided to choose from.This is good to evaluate the impact of issues that are defined by the survey but not issues that are defined by students themselves. Also this may make any study biased to the terms of the survey questions and choices, finite to the number of its questions, and not specific to a particular component of learning such as assessment.

This research aims to study students' feedback of assessment in higher education. The proposed approach of this research has three stages: first, to take students' feedback on assessment in a written text format and assign a sentiment to each entry as(positive, negative). The sentiment will be assigned to all instances that include instances talking about all advantages and disadvantages of assessment; second, to detect issues of assessment from students' negative instances; finally, to integrate the result from the second stage with actual students' marks and attendance.

First stage aims to see the extent of students' satisfaction of assessment. To do that we intend to use sentiment analysis methods. Second stage aims to identify and develop a set of labels to classify issues accordingly, in this stage we use classification methods. Final stage aims to test the correlation among the following variables: Mark, attendance, and issues that student suffers from, In this stage classification and statistical methods will be used.

Thus, The following section is addressing the research questions:

1.2   Research Importance and Questions

Analyzing students' feedback of assessment can lead to identifying issues that students struggle with and allow decision-makers to propose a suitable solution to tackle them to enhance the learning process
The research is important in educational data mining field, as the literature did not reveal any study that applied data mining models on assessment practices data sets.
This research is also important as it will produce optimal data mining models ready to apply on big data sets, such as online, and massive online open courses (MOOCs) feedback.
In this research, we aim to answer the following questions:

- How to automatically detect the polarity of students' feedback of assessment?
- How to automatically detect issues of assessment?
- What is the best method to visualize the correlation between performance and detected issues?

1.3   Research Contribution

This research will contribute the following elements to the field of knowledge:

- Proposes effective data mining framework that detect the polarity of the students' feedback.
- Proposes effective data mining framework that detect issues of students' feedback of assessment.
- Proposes effective data mining framework that test the correlation between detected issues and performance.
- As it is difficult to find assessment related data set, this research generates one from student general feedback.

The rest of this paper is structured as follow: section two is literature review, section three is the methodology, and section four is the Early results and Future Work.

2   Literature Review

Researchers studied students' experiences that affected their performance[6,7]using social media data. Their intentions were to identify engineering students' issues regarding learning in general and were not concentrating on a specific subject such as assessment. They [6,7] used data mining and natural language processing methods. Studies[4,8,9,10] analyzed assessment, Struyven, Janssens and Dochy [4] reviewed articles and documents(between 1980-2002) that related to assessment and evaluation from students' point of view .They found that students' perceptions of assessment and their approaches for learning are strongly related. Students perceived the multiple choice format as more favourable than the essay

format with exception of female students and students who have strong learning skills[4].

Cruickshank studied the use of exams at postgraduate level, the language and cultural issues faced by international students, and the impact of international-isation on the United Kingdom higher education sector [8]. His outcome recommendations were: Institutions should increase exam time by 15 minutes to allow reading, reduce exam weighting marks, and assessment should test the students' knowledge and not their language skills[8].

Trotter studied students' perception of continuous summative assessment and its impact on their motivation,approaches to learning, and changes to their learning environment[9]. He concluded that although the process is time consuming and hard work, it significantly enhanced the learning environment [9].

Alqarran analysed a two universities' methods of assessing students, his study recommended that institutions should encourage and support different methods of assessments[10].

Our research approach is similar to [6,7,11] in using text data from surveys and social media sites but it will go further to test the correlation among the findings from our research and actual output(mark and attendance) from students' files. we are going to identify issue(s) of assessment from the students' feedback unlike [4,8,9,10] who defined specific issues of assessment and studied them.

3   Research Methodology

As mentioned above in the introduction section and shown in figure 1 the research approach divided into three stages: Sentiment analysis; Issue detection; and performance-Issues correlation.
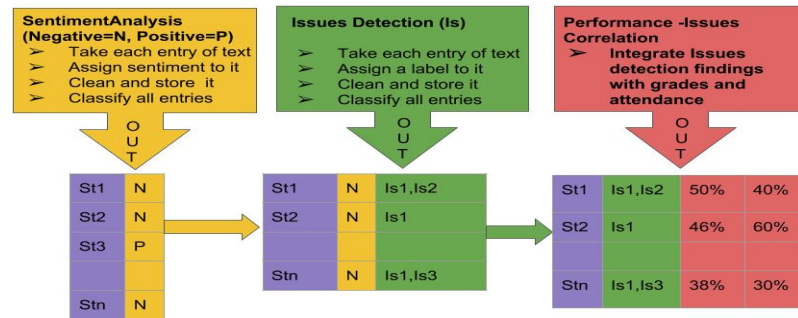

Fig. 1. Analysing process of students' feedback, St= Student Number, Is= Issue

3.1   Data Collections

The first two phases will be based on text data feedback that will be obtained from the students in text format. For the third phases, numerical data (Grades and attendance will be taken from actual students' records) will be combined with the result from sentiment analysis and issues detection phases.The feedback will be collected shortly after the students complete their assessment to ensure the integrity of the collected feedback. Also, we will investigate the use of previous feedback which is normally collected by universities at the end of each teaching block.

3.2   Sentiment Analysis and Issue Detection

In this section, text classification techniques and methods are used, the general framework is presented in figure 2
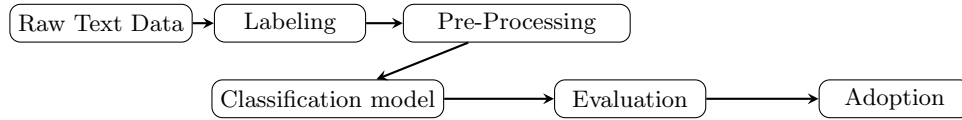

Fig. 2. General text classification framework

Labeling Text feedback in general needs an in depth analyzing as it is normally contain large amount of informal words, jargon, abbreviations local slang words, mis-spelling words, and sarcasm which make the meaning extraction process difficult. Chen [6] tried a popular topic modelling algorithm called Latent Dirichlet Allocation(LDA), it produced senseless word groups with a lot of overlapping words across different topics. They [6] decided to function an in depth analyzing process to have a qualitative look at the data to recognize the quality and minimize the margin of error as they categories these entries .

Pre-processing Cleaning data enhances the output accuracy and minimize data dimension. It depends on the source of data which can be a hand written text or social media and blogs sites as they use special characters.Researches [12,13,14,15,16] used one or more technique(s) of the following:tokenization, remove stop words, needless punctuation, exclamation, question marks,any additional unnecessary symbols,and special marks. And modify words contain uppercase letters or special marks.

Feature Selection Feature selection is also called variable selection or attribute selection, it is used to improve classification effectiveness and computational performance[17], the most popular used features are N-Gram, and Part of Speech (PoS) features[18]

Classification model In this stage, we will investigate a wide range of state-of-the-art classification algorithms [19] such as but not limited to: Naive Bays (NB), Support Vector Machine (SVM), Desioson Tree (DT), and Random Forest(RF).

Evaluation and Adoption Evaluation is the process of using specific metrics to assess how good is the developed classification model. The most popular metrics are: accuracy, precision, recall, and the F-measure.

Early Results Early results showed that about 25% of the first data set commented on the assessment procedures in particular in spit of the fact that the data set was a general feedback .
Also regarding the first stage which is detecting the polarity of the assessment feedback showed significant performance of Support Vector machine models.

3.3 Integrating Students' Grades and Attendance with detected issues

This section represents the final stage of the research, it is to project the students' grades, attendance with detected issue(s) that student found it/them as an obstacle(s) to achieve better. In this stage we aim to test the correlation among detected issues, grade, and attendance, to achieve that, data mining algorithms and statistical methods will be used.

4   Future Work

In particular, this research is to propose an effective data mining framework to study students' feedback of assessment, to detect issues of its procedures, inform the decision-makers of these issues to update and modify accordingly. The main aim is to enhance the learning process.
The research is still in early stage, The next step is to complete collecting data and use it to explore issues of assessment. Then integrate the output of issue detection stage with students' performance data.

References

1. Murphy F. Module Design and Enhancement, Assessment Types. 2009. Accessed: 2017-09-07.
2. Understanding assessment: its role in safeguarding academic standards and quality in higher education. 2012.
3. Strakova Zuzana. Promoting learning through assessment. 2016.
4. Janssens S Struyven K, Dochy F. Students' perceptions about evaluation and assessment in higher education: a review. 2005.
5. Wiliam D. Black P. Assesment and classroom learning. *Assessment in education:priciples, policy and practice*, 5(1):68–75, 1998.

6. Madhavan K Chen X, Vorvoreanu M. mining social media data for understanding students'learning experiences. 2015.
7. Pagare P. Recognizing Students' Problem Using Social Media Data. *International Journal of Computer Science and Mobile Computing*, 4(6):440–446, 2015.
8. Cruickshank P. The Perceptions of Postgraduate International Students of Examinations. *Journal of Perspectives in Applied Academic Practice*, 4(3), 2016.
9. E Trotter. Student perception of continous summative assessment. 2006.
10. Alquraan Mahmoud R. A cross-cultural study of students' perceptions of assessment practices in higher education. *Education, Business and Society: Contemporary Middle Eastern Issues*, 7(4):293–315, August 2014.
11. Nabeela Altrabsheh, Mihaela Cocea, and Sanaz Fallahkhair. Sentiment analysis: towards a tool for analysing real-time students feedback. In *Tools with Artificial Intelligence (ICTAI), 2014 IEEE 26th International Conference on*, pages 419–423. IEEE, 2014.
12. Tippakorn Rungkasiri Wilas Chamlertwat, Pattarasinee Bhattarakosol and Choochart Haruechaiyasak. Discovering consumer insight from twitter via sentiment analysis. *Journal of Universal Computer Science*, 18(8):973–992, August 2012.
13. M. Cooper W.B. Claster and P. Sallis. Thailand – tourism and conflict: Modeling sentiment from twitter tweets using naive bayes and unsupervised artificial neural nets. pages 89–94, 2010.
14. K. Mouthami, K. N. Devi, and V. M. Bhaskaran. Sentiment analysis and classification based on textual reviews. In *2013 International Conference on Information Communication and Embedded Systems (ICICES)*, pages 271–276, Feb 2013.
15. Alexander Pak and Patrick Paroubek. Twitter based system: Using twitter for disambiguating sentiment ambiguous adjectives. 2010.
16. Alexander Pak and Patrick Paroubek. Twitter as a corpus for sentiment analysis and opinion mining. 2010.
17. Yang Y. Rogati m. High-Performing Feature Selection for Text Classification. 2002.
18. Fallahkhair S Altrabsheh N, Cocea M. Predicting learning-related emotions from students' textual classroom feedback via Twitter. 2015.
19. Pollyanna Gonçalves Marcos André Gonçalves Fabrício Benevenuto Filipe N Ribeiro, Matheus Araújo. entiBench - a benchmark comparison of state-of-the-practice sentiment analysis methods. 2016.