Visual Analytics for Faculty Teaching Practice and Performance: A Case Study of Educational Data Mining

Atikah Khalid¹, Mohd Hilmi Hasan², and Azlinda Abdul Malik³
Centre for Research in Data Science, Computer and Information Sciences Department, Universiti Teknologi PETRONAS, 32610 Seri Iskandar, Perak, Malaysia

Abstract

Faculty teaching practice and performance have become one of the utmost importance factors in developing students’ quality in academics. The performance of the faculty plays an important role in academic institutions. Evaluating the faculty members' performance helps to gather critical information and discover new ways of improving them. In this paper, the proposed system can be used as a comprehensive system for evaluating, reporting and analyzing data with a promising audience by utilizing the visual analytics platform in using the educational mining techniques. Based on different parameters, the faculty teaching practice and performance are evaluated and projected by building models. The sample data is collected, preprocessed, and model learning is done using Decision Tree, Support Vector Machine (SVM) and Artificial Neural Network (ANN) in this evaluation. Besides, an analysis of the variable importance for each classifier model is done to see which questions appear in determining the success of faculty members' performance. The idea of this paper is to indicate the effectiveness of Visual Analytics for Faculty Teaching Practice and Performance using Educational Mining Techniques on Student's Self-Reflection Tool (SSRT) survey.

Keywords: Educational data mining, Visual analytics, teaching practice, Teaching performance.
1. Introduction

Assessing faculty teaching performance is essential for assurance of student learning and helping the administrative to make informed decision (Lyde et. al., 2016). Similarly, to increase the efficiency of delivering quality education, it is important to assess the performance of two major stakeholders, namely students and faculty (Saxena and Bhatnagar, 2018). The term faculty teaching practice and performance is based on the learning experience, student overall performance, and actual subject performance.

The most used method for evaluating the quality of faculty teaching practice and performance is through the evaluation form filled by the students from the course and lecturer questionnaire. The ongoing debate on the reliability and validity of these students’ evaluation concerns primarily on the students not having enough experience and maturity to evaluate the course and the lecturer, students' evaluations being affected by the popularity of the course and/or lecturer, grades given by the lecturer and course being compulsory or elective. However, students are the main source of learning environment information, and their evaluations are highly needed to rate the quality, effectiveness, and satisfaction of course content, instruction method, textbook, and assignment. Student evaluations are primarily used to improve the quality of the course and teaching, and in some institutions, they become part of the evaluation process for the staff appraisal system. These can then be used to more effectively allocate resources and staff, make better decisions about educational activities to improve the success of students, increase the learning outcome of students, increase the retention rate of students, decrease the dropout rate of students and reduce the cost of system processes.

In recent years, the emerging data analytics techniques have been the most useful technology that gives beneficial purposes for the university to analyze and evaluate faculty teaching practice and performance by data collection through the course and evaluation forms filled by the students. According to the study in Kokina et al. (Kokina et. al, 2017), data analytics are important components for effective performance assessment and growth management. The data collected can also be utilized for assessment and classification by building data models using data mining techniques. Data mining is giving valuable benefits to higher education, and this is known as educational data mining (Deepak et al., 2016). This educational mining technique aims to extract the data from the dataset and turns it for a recognizable structure for future use.

Measuring the standard of teaching effectiveness and transparency in higher education has a long and well-researched tradition. Nonetheless, the questions of what "effective" means and how it is measured continue to challenge faculty and administrators in university. The main problem faced by Higher Educational Institutions (HEIs) is the challenge of improving academic quality (Saxena and Bhatnagar, 2018). Higher education has an important role both for the student and the university itself. However, most of the instructive mining research focus on modelling and forecasting student performance, and very few research models on faculty performance are available (Saxena and Bhatnagar, 2018). Moreover, according to the findings found by Lyde et al., evaluation rating forms filled by students are the most commonly used indicator of success in faculty members' teaching performance yet many parties did not make the best use of these datasets (Lyde et. al., 2016). Hence, by evaluating
and analyzing the faculty teaching practice and performance, it will help to represent an aid for the growth and maintain the quality of the university and development of the students.

This paper aims to propose a solution for the university to evaluate the faculty teaching practice and performance by using the visual analytics platform and the classification techniques. The solution must be able to provide analysis of data through data visualization on the dashboard. In this paper, the propose visual analytics dashboard for faculty teaching practice and performance is presented. Furthermore, this paper also includes comparative study of three classification models namely Decision Tree, Support Vector Machines (SVM) and Artificial Neural Network (ANN) for enhancing the capability of the visual analytics dashboard to predict future faculty performance.

This paper is organized as the following. The next section discusses related works, then section 3 describes the methodology used in this research. The subsequent section, i.e. section 4 presents the results and analysis. Then section 5 concludes the paper and outlines the future directions.

2. Related Works

Based on past research, a few studies have been conducted related to building models using classification techniques for faculty teaching practice and performance evaluation. According to Saxena and Bhatnagar, data mining techniques in higher education are increasingly being applied to provide insights into the academic and resolve organizational issues which imply the increase of the effectiveness of the management (Saxena and Bhatnagar, 2018). This will help to make use of the evaluation rating forms filled by students to indicate the success in faculty members' teaching performance.

Nonetheless, most works on education mining focus mainly on modelling the performance of students and a very small number of study model on the performance of faculty. Accordingly, higher education has become particularly important in making successful careers for graduate students and evaluating faculty teaching practice and performance evaluation is vital to maintain academic quality in developing students’ careers (Deepak et al., 2016). The quality of the faculty plays a crucial role in academic institutions. Besides that, maintaining a good quality of faculty teaching practice and performance will help to attract more quality students which in return improves student admission and enrolment.

Numerous studies have attempted to explain performance evaluation in Higher Education Institutions. Agaoglu found that education mining research focuses on modelling the performance of students rather than the performance of instructors in faculty (Agaoglu, 2016). Moreover, faculty teaching practice and performance are given the top priority for maintaining good quality in institutions to build the careers of the students and in modelling the performance (Deepak et al., 2016). Later, Saxena and Bhatnagar, extended their analysis and revealed that it is important to evaluate the performance of two major stakeholders, namely students and faculty, to increase the effectiveness of delivering quality education (Saxena and Bhatnagar, 2018). However, studies on visual analytics on faculty teaching and performance were scarce. Hence, this study aims to investigate the benefits of the outcome in
evaluating the teaching practice and performance to solve educational and administrative problems in higher education.

The application of data mining in the domain of education has been growing. The work presented in Deepak et al. focuses on faculty performance by using a SVM algorithm for classification of data to classify the performance of the faculty (Deepak et al., 2016). This can be supported by Agaoglu - findings that indicate the effectiveness of using data mining techniques by four different classification techniques, namely decision tree algorithms, SVM, ANN and discriminant analysis in course evaluation and higher education mining to classify the instructor performance (Agaoglu, 2016). In placing more emphasis, Saxena and Bhatnagar states that it is observed that the Decision Tree is the best technique amongst five applied techniques, Logical Regression, Decision Tree, SVM, ANN and Naive Bayes for analysis of faculty performance evaluation (Saxena and Bhatnagar, 2018). Hence, in this study, Decision Tree, SVM and ANN are selected for the classification models as these were the top three classification algorithms often used in previous studies.

3. Materials and Methods

3.1 Dataset

The dataset for this research is taken from Computer and Information Sciences Department (CISD) in Universiti Teknologi PETRONAS (UTP). It is the results from the system known as Student’s Self-Reflection Tool (SSRT), which is used to evaluate students’ satisfaction and learning outcomes. The evaluation by students is done according to course’s lecturer. SSRT evaluation is done based on the eight questions shown in Table 1.

Table 1: Evaluation questions and details

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Possible Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question1</td>
<td>I am able to apply the knowledge gained from this course.</td>
<td>{1, 2, 3, 4, 5}</td>
</tr>
<tr>
<td>Question2</td>
<td>I had a clear understanding of the aims and goals of this course.</td>
<td>{1, 2, 3, 4, 5}</td>
</tr>
<tr>
<td>Question3</td>
<td>The requirements of the assessments assigned were made clear to me.</td>
<td>{1, 2, 3, 4, 5}</td>
</tr>
<tr>
<td>Question4</td>
<td>The instructor/lecturer was able to facilitate the learning process.</td>
<td>{1, 2, 3, 4, 5}</td>
</tr>
<tr>
<td>Question5</td>
<td>I am able to make connection between this course and other courses that I have learnt.</td>
<td>{1, 2, 3, 4, 5}</td>
</tr>
<tr>
<td>Question6</td>
<td>I was provided with various resources to help me learn.</td>
<td>{1, 2, 3, 4, 5}</td>
</tr>
<tr>
<td>Question7</td>
<td>I always received adequate and timely feedback from the instructor/lecturer.</td>
<td>{1, 2, 3, 4, 5}</td>
</tr>
<tr>
<td>Question8</td>
<td>The way the course was conducted created conducive environment for student-instructor/lecturer interactions.</td>
<td>{1, 2, 3, 4, 5}</td>
</tr>
</tbody>
</table>

The possible values in Table 1 represent the following evaluation ratings: 1 = Strongly Disagree, 2 = Disagree, 3 = Neural, 4 = Agree, and 5 = Strongly Agree
3.2 Visualization Method

The visualization is carried out by using Microsoft Power BI tools. Various kinds of visualizations are used namely pie chart, table, stack chart, line graphs and many others.

3.3 Classification Methods

As mentioned earlier, this research compares three classification models namely Decision Tree, Support Vector Machines (SVM) and Artificial Neural Network (ANN) for enhancing the capability of the visual analytics dashboard to predict future faculty performance. These three algorithms are described as the following.

3.3.1 Decision Tree

A decision tree is a method for decision support and its possible consequences, including outcomes of chance events, resource costs, and usefulness. Tree-based learning algorithms are mostly based on supervised methods of learning with high accuracy, stability, and interpretative ease and are used for both linear and non-linear relationship modelling. The advantage of the decision tree is known that the decision tree algorithms are reasonably fast and accurate and consume less time (Frederick, 2016). Non-technical professionals can easily understand the rules generated by this algorithm. However, the disadvantage of the decision tree is they are unstable, which means that a small change in the data can result in a major change in the optimal decision tree structure.

A decision tree can be utilized to model regression and classification problems. Its mechanism is based on stratifying or segmenting the predictor space into several regions. The set of rules used in splitting the predictor space are organized in a tree-shape structure, as shown in Figure 1(a).

3.3.2 Support Vector Machine

In 1995, Vladimir N. Vapnik founded Support Vector Machines (SVM). This has become one of the most famous supervised classification and regression learning algorithm. By constructing an N-dimensional Hyper plane, the classification performed by SVM is perfected. The four basic concepts of SVM are hyper planes, maximum margin hyper plane, maximum margin and support vectors. Moreover, SVM is a group of supervised methods of learning used for classification, regression and detect outliers. SVM works in high-dimensional spaces as well as in cases where the number of dimensions is greater than the number of samples. The advantage of SVM is that it works relatively well when there is a clear margin of separation between classes while the disadvantage is the SVM algorithm is not suitable for large data sets. Figure 1(b) shows the diagram of SVM and its basic components.
3.3.3. Artificial Neural Networks

Artificial neural networks (ANNs) are computational systems inspired by biological neural networks that make up human brains which slowly improve their performance through learning. ANN is a weighted set of input/output connected units with layers between the input and output. The layers can be single or multiple layers. The advantage of ANN algorithms is highly accurate in data recognition and very high accuracy in classification (Frederick, 2016). However, the disadvantage of the ANN is that to converge, it needs a large amount of training data and it may take longer to train than other models. ANN may also suffer from overfitting, which occurs when training is too long, and the network begins to consider noise as part of the pattern. Figure 1(c) shows the diagram of ANN.

4. Results and Discussion

The proposed system, known as Visual Analytics Dashboard for the Faculty Teaching Practice and Performance using Educational Mining Techniques contains the dashboard for the following:

- Main page (Figure 2)
- The first page is the Performance by Faculty Members (Figure 3)
- The second page is the Performance by Questions (Figure 4)
- The third page is the Overall Overview (Figure 5)
- The last page is the Performance Prediction (Figure 6)
Figure 1: (a) Decision tree diagram for Titanic dataset (Polhemus, 2018; Kumar, 2020) (b) SVM and its basic components (Javatpoint, 2020) (c) Artificial Neural Network’s diagram (Facundo, 2017)
Figure 3 shows the Performance by Faculty Members i.e. viewing the overall ranking of the faculty members. The ranking is calculated based on their total average rating of all the questions rated by the students. The features include; (1) Filtering by Semester to see which lecturer that perform well in that particular semester, (2) Filtering by Year to see which lecturer that perform well in that particular year, (3) Filtering by Cycle to see which lecturer
that perform well in that particular cycle, (4) Filtering by Lecturer to see the lecturer’s average rating, ranking and particular semester, year and cycle the lecturer is teaching.

Figure 4 shows the Performance by Questions page i.e. viewing the overall faculty teaching practice and performance based on questions. The features include; (1) Filtering the Lecturer to see the course that the lecturer is teaching, the highest to the lowest questions rank rated by the students based on all the courses and the average total rank throughout the cycle, semester and year, (2) Filtering the Course to see which lecturers are teaching the course, the highest to the lowest questions rank rated by the students based on the course, (3) The bar graph to see the highest to the lowest questions rank based on a particular lecturer and all courses or specific course, and (4) Line chart to see how the lecturer performs throughout the cycle, semester and year based on the average total rating.

Figure 5 shows the Overall Overview page to see the overall performance of the faculty members. The features include; (1) Filtering by Semester to see how many lecturers obtain High, Moderate or Low rating and highest to the lowest questions rank in that particular semester, and (2) Filtering by Year to see how many lecturers obtain High, Moderate or Low rating and highest to the lowest questions rank in that particular year. Figure 6 shows the Performance Prediction page where the questions are ranked based on the ratings by students. These predictions show which questions are the highest important of all questions and having the greatest relevance with the faculty performance.

Figure 7 shows the questions being ranked by importance according to the three machine learning classification algorithms namely SWM, decision tree and ANN. For SVM, the top three main important variables are Question 8 - “The way the course was conducted created conducive environment for student-instructor/lecturer interactions”, Question 4 - “The instructor/lecturer was able to facilitate the learning process” and Question 3 - “The requirements of the assessments assigned were made clear to me”. For decision tree, the top three main important variables are Question 7 - “I always received adequate and timely feedback from the instructor/lecturer”, Question 8 - “The way the course was conducted created conducive environment for student-instructor/lecturer interactions” and Question 5 - “I am able to make connection between this course and other courses that I have learnt”. For ANN, the top three main important variables are Question 5 - “I am able to make connection between this course and other courses that I have learnt”, Question 6 - “I was provided with various resources to help me learn” and Question 8 - “The way the course was conducted created conducive environment for student-instructor/lecturer interactions”.
Figure 4: Performance by Questions Page

Figure 5: Overall Overview Page
Figure 6: Performance Prediction Page
5. Conclusion

The increasing use of various types of data analytics to inform strategic management in analytical reasoning for decision making is growing in various fields nowadays. By using the analytics tools for performance evaluation, the current evaluation system can be extended to a simpler, easier and more precise performance evaluation system. This application can be used for data storage, retrieval, visualization and analysis, as well as for precise reporting and performance evaluation. Furthermore, the use of the analytics platform to carry out the evaluation process can exponentially reduce the cost of the necessary resources. The use of various data visualization techniques will facilitate the authorities to evaluate the performance of a single faculty, a group of lecturers and an entire department.

This paper proposes a visual analytics and data mining techniques in higher education to provide information on the faculty teaching practice and performance. The data obtained from the Student's Self-Reflection Tool (SSRT) survey is utilized for visual analytics in evaluating faculty teaching practice and performance using the educational mining techniques. It then
provides a context for discussing more sophisticated statistical analysis techniques such as Decision Tree, Support Vector Machine (SVM) and Artificial Neural Network (ANN).

Acknowledgment

The research is part of the continuous work under Yayasan Universiti Teknologi PETRONAS grant No. 015LC0-029.

References


