Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 102 (2016) 137 – 142

Procedia Computer Science

12th International Conference on Application of Fuzzy Systems and Soft Computing, ICAFS 2016, 29-30 August 2016, Vienna, Austria

Using data mining to predict instructor performance

Ahmed Mohamed Ahmed[a,b*], Ahmet Rizaner[c], Ali Hakan Ulusoy[c]

[a,*] *College of Computer Science, University of Bahri, Sudan*
[b] *Department of Mathematics, Eastern Mediterranean University, North Cyprus, Turkey*
[c] *Department of Information Technology, Eastern Mediterranean University, North Cyprus, Turkey*

Abstract

During these decades, data mining has become one of the effective tools for data analysis and knowledge management system, so that there are many areas which adapted data mining approach to solve their problems. Using data mining in education to enhance the education system is still relatively new. This paper focuses on predicting the instructor performance and investigates the factors that affect students' achievements to improve the education system quality. Turkey Student Evaluation records dataset is considered and run on different data classifier such as J48 Decision Tree, Multilayer Perception, Naïve Bayes, and Sequential Minimal Optimization. Comparison of all the four classifiers is conducted to predict the accuracy and to find the best performing classification algorithm among all. The conclusions of this study are very promising and provide another point of view to evaluate student performance. It also highlights the importance of employing data mining tools in the field of education. The results show that using the attribute evaluation method on the dataset increases the prediction performance accuracy.

Peer-review under responsibility of the Organizing Committee of ICAFS 2016

Keywords: Data mining; decision tree; multilayer perception; naïve bayes; sequential minimal optimization

1. Introduction

Nowadays Data Mining (DM) has attracted a lot attention in data analysis area, and it became recognizable new tool for data analysis that can be used to extract valuable and meaningful knowledge from data. DM offers promising ways to uncover hidden patterns within large amounts of data. These hidden patterns can potentially be used to predict future behavior[1]. Accordingly, DM has been adopted by many researchers to solve real-world problems in various domains such as marketing, stock market, telecommunication, industrials, health care, medical and customer relationship. Recently a reasonable number of researches have been conducted to apply DM techniques in the education area in ordered to classify and predict student performance in numerous education institutes. Employing DM techniques in education is promising because of the tremendous opportunities in this area.

Peer-review under responsibility of the Organizing Committee of ICAFS 2016
doi:10.1016/j.procs.2016.09.380

Moreover, education systems claim new approaches which improve quality, efficiency, and achievement[2]. Mostly DM is utilized in education to investigate the impact of pedagogical strategies on students, and how students understand the course[3].

The academic performance of students based on several factors. The most important factors are the attributes such as the previous academic records, economic status, family background, and demographic data, and the prediction methods. Thus most of the research in this area relayed on the attributes specified student data. This paper attempts to investigate the data associated with the student evaluation for the instructors to improve the quality of education and indicate the factors that affect the student performance. The prediction of student performance is mainly related to the quality of teaching process[4]. In this paper, some data classification algorithms are applied to Turkey Student Evaluation dataset to predict student achievement, investigate instructor's performance, and find the best classification algorithm in accordance with high accuracy.

2. Background

In the past years, several studies have addressed DM for educational purposes. Minaei-Bidgolim was one of the first authors who used genetic algorithms classified students' performance in order to predict their final grades[5]. DM in education from 1995 to 2005 was reviewed in[3] which became a significant research paper in this field. A student's academic success (classified into low, medium, and high-risk classes) using different DM methods such as Decision Tree (DT), and Neural Network (NN) was predicted in[6]. The research in[7] attempted to investigate the reason of failure in the two core classes (Mathematics and Portuguese) of two secondary school students from the Alentejo Region of Portugal. The result showed that both DT and NN algorithms had the predictive accuracy as 72% for a four-class dataset.

The most efficient machine learning technique in predicting the final grade of Ionian University Informatics postgraduate students was investigated in[8]. It was found that Naïve Base (NB) and K-Nearest Neighbors (K-NN) algorithm accurately predicted the students' final performances exactly when it included a small number of instances. The research in[9] attempted to apply DM techniques by using Microsoft and Weka on small student datasets. The results proved that the prediction was significantly successful by both technologies. In[10], a new approach was introduced to predict student performance by using comment DM that employed Latent Dirichlet Allocation (LDA) and Support Vector Machine (SVM). The results were very promising.

3. Dataset Description

Most of the research in the past frequently used Cumulative Grade Point Average (CGPA) and internal assessment attributes to predict student performance. Another common attribute used frequently was the students' demographic and external assessments. Several researches have been conducted to enhance educational systems using above mentioned data. This paper attempts to improve the quality of education system by utilizing the data that associated with student evaluation for their instructors. The dataset is collected from University of California-Irvine (UCI) Machine Learning Repository and contains a total of 5,820 evaluation scores provided by students from Gazi University in Ankara, Turkey. There is a total of 28 course specific questions and additional 5 attributes. Q1-Q28 are all Likert-type in which responses are scored as {1, 2, 3, 4, 5}. The class label attribute is suggested to be result and takes values also as {1, 2, 3, 4, 5} whereas result values that are greater than 3 are considered to be very good , result values that are equal to 3 are considered as good , and result values that are less than 3 are considered to be bad. Moreover, the level of attendance values are taken as {0, 1, 2, 3, 4} whereas values less 2 are considered to be weak, values equal to 2 are considered to be medium, and values greater than 2 are considered to be good. Furthermore, level of difficulty of the course values are taken as {1, 2, 3, 4, 5}, whereas values less than 3 are considered to be low, values equal to 3 are medium, and values greater than 3 noted as high. The attributes information given below are divided into two sections: First one contains general information and the second holds evaluation questions:

instr: Instructor's identifier; values taken from {1,2,3}
class: Course code (descriptor); values taken from {1-13}
repeat: Number of times the student is taking this course; values taken from {0, 1,2,3,...}
attendance: Code of the level of attendance; values from {0, 1, 2, 3, 4}
difficulty: Level of difficulty of the course as perceived by the student; values taken from {1, 2, 3,4,5}

Q1-Q28 are all Likert-type, meaning that the values are taken as {1, 2, 3, 4, 5}

Q1: The semester course content, teaching method and evaluation system were provided at the start.
Q2: The course aims and objectives were clearly stated at the beginning of the period.
Q3: The course was worth the amount of credit assigned to it.
Q4: The course was taught according to the syllabus announced on the first day of class.
Q5: The class discussions, homework assignments, applications and studies were satisfactory.
Q6: The textbook and other courses resources were sufficient and up to date.
Q7: The course allowed field work, applications, laboratory, discussion and other studies.
Q8: The quizzes, assignments, projects and exams contributed to helping the learning.
Q9: I greatly enjoyed the class and was eager to actively participate during the lectures.
Q10: My initial expectations about the course were met at the end of the period or year.
Q11: The course was relevant and beneficial to my professional development.
Q12: The course helped me look at life and the world with a new perspective.
Q13: The instructor's knowledge was relevant and up to date.
Q14: The instructor came prepared for classes.
Q15: The instructor taught in accordance with the announced lesson plan.
Q16: The instructor was committed to the course and was understandable.
Q17: The instructor arrived on time for classes.
Q18: The instructor had a smooth and easy to follow delivery/speech.
Q19: The instructor made effective use of class hours.
Q20: The instructor explained the course and was eager to be helpful to students.
Q21: The instructor demonstrated a positive approach to students.
Q22: The instructor was open and respectful of the views of students about the course.
Q23: The instructor encouraged participation in the course.
Q24: The instructor gave relevant homework assignments/projects, and helped/guided students.
Q25: The instructor responded to questions about the course inside and outside of the course.
Q26: The instructor's evaluation system (midterm and final questions, projects, assignments, etc.) effectively measured the course objectives.
Q27: The instructor provided solutions to exams and discussed them with students.
Q28: The instructor treated all students in a right and objective manner.

4. Methodology and Data Mining Methods

4.1. Methodology:

DM techniques can be used as tools to develop and improve the quality of education system and also to enhance school resource management. For instance, there are several interesting questions for this domain that could be answered using DM:

- Is it possible to improve education quality based on student evaluation for courses?
- What is the quality of courses can be offered to attract more students?
- What are the main reasons that affect instructor’s performance?
- Is possible to predict instructor’s performance?
- What are the factors that affect student’s achievement?

Here the research focuses on the last two questions. Modelling instructor's performance is a useful method for both educators and students since it can help a better understanding of education achievement.

4.2. Data Mining Methods:

DM techniques are used to extract valuable information from a huge amount of data. Different DM techniques have been introduced during the past decades. Selecting the most suitable technique to mine the data is the step that leads to the correct road map. The main goal is to build a DM classification model that enables us to clarify the factors that affect student performance. In this paper, four well-known DM techniques that are J48 DT, MLP, NB, and SMO are used.

5. Experiment Design

For the implementation of the first four consecutive classification tasks, we have used Weka workbench. Experiments are conducted in four consecutive steps. In step one, attribute evaluation is performed using the OneR algorithm to clarify which attribute has the greatest potential impact on every class in the dataset[11]. Weka ranker method is applied to justify ranks of attributes using 5 fold cross validation. The results show that the attributes Q23, Q27, Q26, Q21, Q22, Q28, Q25, and Q24 have the greatest impact on the dataset, whereas others like attendance, instructor, class, and Q1 are considered redundant because they obtained the lowest impact on the dataset.

In step two, the attributes that have the highest impacts (according to the results of the previous step the best 24 attributes considered) are selected and the four suggested DM techniques are conducted after removing the last ten attributes with lower impacts on the dataset which are (attendance, instructor, class, Q1, difficulty, number of repeats, Q2, Q3, Q13, Q5). Table 1 shows the accuracy of the prediction when mentioned algorithms applied on the dataset after attribute evaluation is done.

In step three, all dataset of Turkey Student Evaluation is tested and analyzed with the mentioned four classification algorithms. The dataset is divided into two sets where 66% is used for training, and 34% is used for testing. The model is built using the training set and tested by the test set. A comparison of the accuracy of all classifiers is presented in Table 1. The results show that J48 DT algorithm achieves the best performance compared to the other algorithms with an accuracy of 84.8%.

Table 1: Prediction accuracy results after attribute evaluation process and when algorithms run on all dataset

Algorithm	Performance accuracy after attribute evaluation process for attributes with highest impacts	Performance accuracy when algorithms run on all data set for all attributes
J48 DT	85.1%	84.8%
NB	84.3%	83.3%
SMO	85.8%	84.5%
MLP	84.6%	82.5%

In step four, some experiments are conducted in order to investigate the performance of instructors. The aim of this analysis is to determine the performance of each instructor individually and to investigate the factors that affect their achievements. The four suggested algorithms are run on the dataset that is organized as explained below. The evaluation records of the courses that are taught by each instructor are combined together in one dataset file. Since we have three instructors, the data set is grouped into three distinct files. The results of the experiments are summarized in Table 2 and 3.

In step five, the four suggested algorithms are applied on the dataset of each instructor as mentioned above in step four after removing the worst ranked attributes that have the lowest impact on the dataset such as in step 2.

Those are attendance, instructor, class, Q1, difficulty, the number of repeats, Q2, Q3, Q13, Q5, and considered the best 24 attributes. The results are explained as in Table 4.

Table 2: Instructors, courses, and numbers of students evaluated for each instructor

Instructor	Course Code	Total Number of Students
1	2, 7, 10	776
2	1, 6, 11, 13	1,444
3	3, 4, 5, 8, 9, 12,13	3,601

Table 3: Performance accuracy of each instructor individually

Algorithm	Performance accuracy for instructor 1	Performance accuracy for instructor 2	Performance accuracy for instructor 3
J48 DT	85.4%	85.7%	82.8%
NB	85.5%	86.8%	82.0%
MLP	86.2%	87.4%	82.8%
SMO	87.0%	85.4%	83.0%

Table 4: Performance accuracy of instructors for attributes that have the highest impact on dataset

Algorithm	Performance accuracy for instructor 1	Performance Accuracy for instructor 2	Performance accuracy for instructor 3
J48 DT	85.6%	86.4%	83.0%
NB	85.9%	87.3%	82.8%
MLP	85.6%	87.8%	83.5%
SMO	85.2%	86.4%	83.8%

6. Results and Discussion

Different methods and techniques of DM are utilized in this paper. The data set is tested and analyzed using four data classifiers which are J48 DT, MLP, NB, and SMO. Comparison of accuracy of all algorithms is performed during the prediction process. It is found that using the attribute evaluation method on the dataset is helpful in order to predict the instructor performance. The most important attributes in the dataset are selected then the mentioned above algorithms are run on the dataset. Table 1 shows those attributes that have a strong effect on student's achievement. This means that these attributes are more significant in predicting the instructors' performances and accurately describe their experiences. On the other hand Table 1 shows that SMO performs better than other algorithms with an accuracy level of 85.8%. Furthermore, from Table 1 it is also observed that J48 DT algorithm outperforms other algorithms when applied on all the dataset with an accuracy level of 84.8%.

Another interesting issue is observed from the results which show that the performance of an instructor is mainly affected by the number of courses that is taught. Table 4 shows that all classification algorithms obtained lower prediction accuracy when running on instructor 3 dataset file, compared to the prediction accuracy obtained by those algorithms when run on instructor 1 and 2 dataset file. By comparing of all classifiers, SMO and MLP algorithms performed best among all classifiers with accuracies as 87.0%, 86.2% respectively for instructor 1 dataset as shown in Table 3. While the accuracy of SMO degraded, MLP continued to give the best performance with an accuracy of 87.2% for instructor 2 dataset as shown in Table 3. On another hand the results show that the performance accuracy increased when the worst ranked attributes are removed compared to the case when the algorithms run on the dataset with all attributes. Table 4 indicates the performance accuracy of instructors for attributes that have the highest impact on dataset after removing the worst attributes. It can be noticed from the results that the performance accuracies of all algorithms in Table 4 are better than the accuracies obtained by these algorithms with all attributes that are presented in Table 3 except for MLP and SMO that performs well on instructor 1 dataset in Table 3 whereas their performance degrades when run on instructor 1 dataset after removing worst attributes as show in Table 4.

7. Conclusion

We concluded that using data of student evaluation for courses is useful to predict the factors that affect their achievement and also to predict instructors' performance. Moreover, it is another point of view to improve educational quality which is vital to attract students while most of the researchers used CGPA and internal assessment attributes to predict students' performance to enhance educational system. Furthermore removing the worst ranked attributes that have a lower impact on dataset increased the algorithms performance accuracies.

References

1 Mendes RR, de Voznika FB, Freitas AA, Novella JC. Discovering fuzzy classification rules with genetic programming and co-evolution. *Principles of Data Mining and Knowledge Discovery*, Springer Berlin Heidelberg, 2001, p. 314-325.

2 Peña-Ayala A. Educational data mining: A survey and a data mining-based analysis of recent works. *Expert systems with applications*, 2014, vol. 41, no. 4. p. 432-462.

3 Andonie R. Extreme data mining: Inference from small dataset. *International Journal of Computers Communications and Control*, 2010, vol. **5**, no. **3**. p. 280-291.

4 Romero C, Ventura S. Educational data mining: A survey from 1995 to 2005. *Expert systems with applications*, 2007, vol. 33, no. 1, p. 135-146.

5 Minaei-Bidgoli B, Punch WF. Using genetic algorithms for data mining optimization in an educational web based system. *Genetic and Evolutionary Computation*, Springer Berlin Heidelberg, 2003, p. 2252-2263.

6 Superby JF, Vandamme JP, Meskens N. Determination of factors influencing the achievement of the first-year university students using data mining method. *Workshop on Educational Data Mining*, 2006, p. 37-44.

7 Cortez P, Silva A. Using data mining to predict secondary school student performance. *EUROSIS, A. Brito and J. Teixeira* (Eds.) 2008, pp. 5-12.

8 Koutina M, Kermanidis K L. Predicting postgraduate students' performance using machine learning technique. *Artificial Intelligence Applications and Innovations*, Springer Berlin Heidelberg 2011, p. 159-168.

9 Natek S, Zwilling M. Student data mining solution–knowledge management system related to higher education institution. *Expert systems with applications*, 2014, vol. 41, no.14, p. 6400-6407.

10 Sorour SE, Mine T, Goda K, Hirokawa S. Estimation of student performance by considering consecutive lesson. *4th International Congress on Advanced Applied Informatics*, 2015, p.121-126.

11 Forman G, Cohen I. Learning from little: Comparison of classifiers given little training. *Knowledge Discovery in Databases: PKDD*, Springer Berlin Heidelberg, 2004, p. 161-172.