International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2017 IJSRCSEIT | Volume 2 | Issue 6 | ISSN : 2456-3307

# Discovering Enhanced Pedagogical Practice Using Data Mining Techniques for Evaluation of Programming Language Teaching

**M. Ramla*[1], M. Ramesh[2]**
*[1]Department of Computer Applications, SRM Institute of Science and Technology, Chennai, Tamil Nadu, India
[2]Department of Computer Science, SRM Institute of Science and Technology, Chennai, Tamil Nadu, India

## ABSTRACT

Evaluating the performance of teachers is a crucial aspect in the effective running of an institution. Teachers are knowledge creators and should be knowledgeable too. Although there are several analytical tools that help in this assessment, this paper especially focuses on the teaching evaluation for programming courses for the computer science subjects. The primary objective is to apply classification techniques to the prediction of performance of teachers. A survey with 150 students of BCA degree of SRM Institute of Science and Technology was made. The various fitting attributes contributing to the performance are identified. A predictive classification model has been built using different classifiers in python and results are tabulated.

**Keywords :** Teaching Evaluation, Predictive model, classification, Educational Data Mining

## I. INTRODUCTION

Teachers and textbooks play an important role in providing students with generative knowledge.[7] Teachers must provide an effective learning environment and pedagogical considerations that guides the students. As the standard of the education system is declining, effective teaching evaluation is of paramount concern. The outcome of these evaluations includes effective college management, appraisal for teachers, improving quality and productivity of the institution, identify the areas of concentration of the teachers, increasing the enrolment of students, decreasing dropout rates and increasing the success rate of the students. There exists a dichotomy in all universities. On the one hand, we have young people with no appetite for research and hungry for programming skills.[11] On the other hand, large majority of professors who have no time or desire to impart those sorts of skills to the students. Throughout a course, students are asked to write programs of relatively limited scope.[6] But they must work as a team constructing complex programs. Learning to program is hard with lot of difficulties and deficits. There are high dropout rates because of programming courses. This can be rectified by correct organization of knowledge given by teacher.

## II. RELATED WORKS

Fateh Ahmadi et. al has analysed and predicted acceptance of a teacher for continuing the teaching in faculty using WEKA tool.2 S. Mardikyan and B. Badur has identified the relevant and irrelevant factors to explain the teaching performance of teachers.3In 2013, Ajay Kumar Pal and Saurabh Pal used Naïve Bayes, ID3, CART and LAD Tree for the evaluation of teacher's performance.5 In short,Educational Data Mining has given rise to hypothesis concerned with the scientific study of human sciences.[4]

## III. METHODOLOGY

Educational Data mining is an evolving field exploring pedagogical data and has gained its momentum in the recent times.1 The teachers are evaluated based on feedback from students or through the performance appraisal of the teachers which results in massive data. This research work integrates the background information of the teachers with the data collected from the students to perform classification.

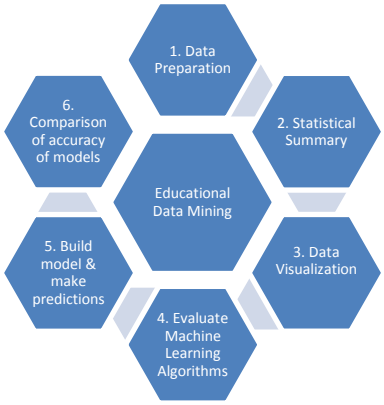


**Figure 1.1** Illustration of methodology

This work considers the close ended questions collected as feedback from the students. It also integrates the test assessment of the individual staffs and introduces a new predictive classification model to assess the teachers with special attention to programming courses. In the elaborate review of the literature, the plethoras of factors that contribute to the teaching evaluation based on the student's expectation are furnished.

| S.No | Attributes | Type |
|---|---|---|
| 1 | Teaching Strategy | Categorical |
| 2 | Attitude | 0-Bad 1-Good |
| 3 | Level of knowledge | Ordinal |
| 4 | Experience in teaching the course | Numerical |
| 5 | Professional Contacts | Categorical Y/N |
| 6 | Lesson Planning | Categorical Y/N |
| 7 | Monitor student progression | Categorical Y/N |
| 8 | Content Delivery | Ordinal |
| 9 | Classroom Management | 0-Bad 1-Good |
| 10 | Voice Modulation | Categorical Y/N |
| 11 | Speed of delivery | 0-Bad 1-Good |
| 12 | Presentation | Ordinal |
| 13 | DoubtClearing | Categorical Y/N |
| 14 | Update with tech trends | Categorical Y/N |
| 15 | Research Activities | Categorical Y/N |

As programming languages is the crux of the computer courses, special concern is given to these courses because it requires different skills that stand apart. The various identified attributes that contributes to effective teaching of programming course are listed below with their description. The various predictors are measured using 6-point Likert Scale that reads

1. Extremely satisfied
2. Very satisfied
3. Somewhat satisfied
4. Somewhat dissatisfied
5. Very dissatisfied
6. Extremely dissatisfied

| Variables | Description | Possible Values |
|---|---|---|
| LT | Logical Thinking | 1 to 6 |
| PS | Problem Solving | 1 to 6 |
| CS | Coding Skills | 1 to 6 |
| SWU | Software Usage | 1 to 6 |
| DS | Debugging Skills | 1 to 6 |
| HO | Hands on Training | 1 to 6 |
| CA | Content Arrangement | 1 to 6 |
| AO | Application Orientation | 1 to 6 |
| AC | Any Certification | Yes/No |
| Class | Classification | 1/0 |

## IV. EXPERIMENTAL SETUP

A data set of 150 instances has been prepared in a csv file format. No missing values were reported in the data. The prediction processes are built using python. The distributions of variables are depicted in the histogram.

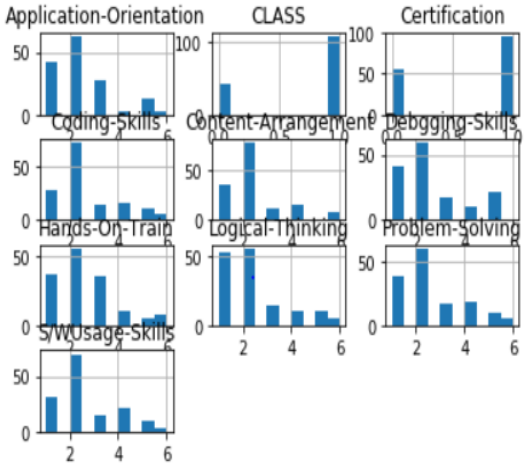


**Figure 1.2** Histogram Plots

The data set was split into two, 80% was used for training our models and 20% was used for validation. also 10-fold cross validation was used to estimate the accuracy of the model.

We evaluated six different classifiers such as Linear Regression, Linear Discriminant Analysis, CART, K-Nearest Neighbour, Naïve Bayes and Support Vector Machine. This is a good mixture of linear and non-linear algorithms.

## V. RESULTS AND DISCUSSION

The various classification algorithms used to build the predictive model is listed below with their accuracy.

```
LR: 0.925000 (0.078617)
LDA: 0.958333 (0.041667)
KNN: 0.950000 (0.040825)
CART: 0.966667 (0.040825)
NB: 0.941667 (0.065085)
SVM: 0.983333 (0.033333)
```

We find that the Support Vector Machine (SVM) gives the high accuracy score. A comparison of mean accuracy of each model is box plotted.

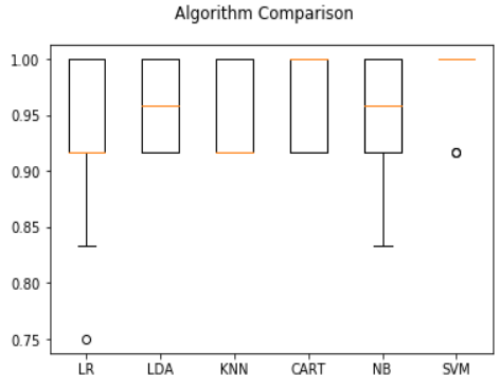


**Figure 1.3** Algorithm Comparisons

The SVM is the most accurate model that we tested and now the same is applied on the validation test to find the accuracy of the model.

```
0.933333333333
[[10  1]
 [ 1 18]]
             precision    recall  f1-score   support

          0       0.91      0.91      0.91        11
          1       0.95      0.95      0.95        19

avg / total       0.93      0.93      0.93        30
```

**Figure 1.4** Classification Report

We see that the accuracy percentage is 0.93. The precision, recall, accuracy and F-measure of each class shows excellent results

## VI.CONCLUSION

Educational data mining explores the pedagogical data by applying various machine learning techniques. But most of them uses the close-end questions with simple answers as multiple choice and easily process it using statistical evaluation. Open ended questions can also be gathered and can be used to recommend effective teaching process. The work can be extended by expanding the data set size and more distinct attributes can be included to improve the overall accuracy of the prediction process.

## VII. REFERENCES


[1]. Asanbe M.O, Osofisan A.O, William W.F, "Teachers" Performance Evaluation in Higher Educational Institution using Data Mining Technique", International Journal of Applied Information Systems, Vol10-No 7, 2015

[2]. Fateh Ahmadi, Shiri Ahmed Abadi, "Data Mining in Teacher Evaluation System using WEKA ", International Journal of Computer Applications, Vol63-No10,2013

[3]. Sona MARD˙IKYAN, Bertan BADUR "Analyzing Teaching Performance of Instructors Using Data Mining Techniques", Informatics in Education, 2011, Vol.10 No.2

[4]. C.Anuradha and T.Velmurugan, "A Comparative Analysis on the Evaluation of Classification Algorithms in the Prediction of Students Performance", Indian Journal of Science and Technology, Vol 8(15), 2015

[5]. Ajay Kumar Pal, Saurabh Pal "Evaluation of Teacher"s Performance: A Data Mining Approach", International Journal of Computer Science and Mobile Computing

[6]. Anthony Robins, Janet Rountree, Nathan Rountree, "Learning and Teaching Programming: A review and Discussion", Computer Science Education, Vol 13 No 2, 2003

[7]. Judith Gal-Ezer, "Teaching Software Designing Skills"

[8]. Chakrit Pong-inwong, Wararat, " Teaching evaluation using Data Mining on Moodle LMS forum", Information Science and Service Science and Data Mining, 6th International Conference on New trends, 2012

[9]. Fabio de Paul, "Speak Well or complain about your teacher:A contribution to educational data mining in the evaluation of teaching practices", Computer in Education, 2016

[10]. E.Deepak, G.Sai Pooja, "SVM Kernel based predictive analytics on faculty performance evaluation", Inventive Computation Technologies. 2015.

[11]. AvinashKak,"Teaching Programming", 2016