Career Prediction based on Academic Performance and Skills using Machine Learning (CPS)

Anujit Deb
ID: UG02-49-18-012
Sadia Anha Tonna
ID: UG02-49-18-004
Niaz Ahmed Naem
ID: UG02-49-18-003
Department of Computer Science and Engineering

Advisor
Dr. D Stalin Alex
Professor
Department of Computer Science and Engineering
State University of Bangladesh

Abstract

In the current scenario, the students need to identify their area of interest in an academic field to opt for the right career courses they are interested in and capable of going through. The students have to go through many options to draw a career path. This paper predicts the career an engineering student can select after graduation using machine learning classification techniques based on academic performance and skills. We will describe the machine learning classification techniques to help students support their decision-making. The machine learning algorithms are presented here; we will compare and analyse the classifier's results developed by this algorithm. We will be discussing our classification in machine learning algorithms to predict the career options for engineering students.The research objective is to find the factors that can affect students' decision to choose the right career path using machine learning techniques.

1 Keywords

Random Forest, Machine Learning, Decision Tree, Classifier, Accuracy, Career Prediction

2 Introduction

Engineering is one of the best career streams apart from medicine, which most students are opting for, some due to interest while some due to parental pressure, as it is the most defined career option in the world. Many engineering students come out of college every year. Many students choose their stream after their graduation. Opting for the right carrier has become a complex science nowadays as there are multiple career options and job competitions in the market. Researchers have even suggested machine learning classification technology to explore the right career option. Without proper guidance from professional services, students face problems choosing the right career path. They often mismatch their career path regarding their personality skills and interest. Students are even forced to opt for a career stream in engineering as pressure from family and the greed for high pay. The students in the past who have passed engineering and started working for MNC but still lack interest and skills make them unhappy. As a result, the upcoming generation has now started to opt for the streams that interest them. Machine learning technique for career guidance has been developed for engineering graduates who have completed their graduation or in the last semester and are still confused about which part of the field they should opt for. It's a big challenge for those students to make the correct decision regarding the career they opt for, as their complete

future depends upon this. Thus, we have considered other aspects that will help us choose the right career path based not just on the academic score but also on their personality, which is important for making their decision.

3 Literature Review

In this section, we have reviewed some papers in the related area. Iqbal et al. have discussed various machine learning techniques to predict grades after students in various courses. Models such as matrix factorisation, classification, and regression are used to analyse the collected data from ITU, Pakistan. They have evaluated performance by using machine learning techniques, and it has been found that RPM is the best among various machine learning techniques. (Iqbal,2017)

Vaidu et al. has implemented machine learning techniques based on student performance to predict their employability skills. They have used KNN and Naïve Bayes models to classify the students into numerous groups. The prediction of the students' employability from the KNN algorithm is 95.33Bayes is 67.67 accuracy (vaidu, 2017)

To predict our future performance after students, ByungHak et al. have used a GritNet algorithm based on deep learning. As per the logistic regression, GritNet gives more accurate results according to this research paper. They have taken data from the Udacity Nanodegree Programs. (B.H,2018)

Jie et al. also proposed a machine learning approach to predict student performance in degree programs. in this investigation, the past as well as present performance of the students is evaluated. It uses a bi-layered structure that compromises multiple phase predictors and a data-driven approach based on the efficient factors to base prediction. This research paper has shown that the proposed method gives a more accurate result than the benchmark approaches. (Jie,2017)

The machine learning algorithms examined by Pojon Murat et al. are used to predict student performance. Pojon Murat et al. has used three different algorithms, Linear regression, Naïve Bayes classification, and decision tree, on two separate data sets, Roberson and another one is featured engineering version. As per the result, Naïve Bayes is the best technique used for the first data set as it gives an accuracy of 98for the second database as it gives an accuracy of 78Singh, M. et al. have used some machine learning techniques to predict the academic performances of the students' subjects wise in their engineering field. To analyse the subject's scores based on the previous semester, they predict the success scores of the students in the ongoing courses. For this purpose, decision tree classifier and Naive Bayesian techniques have been used, and it has been shown that the decision tree gives a more accurate result as compared to Naïve Bayes. (Singh, M., 2013)

Using machine learning techniques like Support Vector Machine, Random Forest, Gradient Boosting, and Naïve Bayes, Bendangnuksung et al. predict the student's performance, whether they will fail or get a pass in the previous semester. As per the prediction, the accuracy rate of Random Forest is higher than other algorithms,

Pushpa et al. have used the DNN model, Deep Neural Network, to predict student performances. The research paper by Pushpa et al. compares the DNN with machine learning algorithms like Naïve Bayes, ANN, and Decision Tree. According to this, DNN achieves 84.3which is better than Machine Learning Techniques. (Pushpa,2017)

4 Proposed Work

Intense websites and web applications help students know their suitable career paths. Still, the drawback of this system is that they only use personality traits to predict the career, which might not give a consistent result. Similarly, numerous websites suggest students opt for a career as per their interests. These systems cannot understand whether the student can survive in that particular field or not. Beth Dietz-Uhler Janet E. Hurn's paper suggests a need for learning analytics to predict and improve student performance to enlighten the importance of students' interests, trends, abilities etc. (UD Beth, 2013). Lokesh Katore, Jayant Umale, Bhakti Ratnaparkhi's paper predicted the different classifiers have different accuracy in predicting a student's career (KS Lokesh, 2015).10

5 Conclusion

In this project we have to create a classification model to predict career options for an undergraduate student. Various input features such as student's mark percentage in 10th, 12th, B.Tech/Diploma, skills in communication etc., are taken into consideration, and output variable was career options a student can choose which were classified as Government Job, M.Tech/ME/MS, MBA, Others, Private Job. We have proposed the six most popular machine learning classification algorithms, i.e., k - nearest neighbour, support vector machine, stochastic gradient descent, decision tree, logistic regression and neural network. The accuracy of each algorithm is evaluated, and the performance sequence of each algorithm is as follows: K - Nearest Neighbor ¿ Gaussian Naive Bayes ¿ Stochastic Gradient Descent ¿ Support Vector Machine ¿ Decision Tree ¿ Logistic Regression ¿ Neural Network ¿ Random Forest. Execution time is also calculated for each algorithm, and the sequence of the algorithm is as follows: Decision Tree ¡ Gaussian Naive Bayes ¡ K - Nearest Neighbor ¡ Stochastic Gradient Descent ¡ Random Forest ¡ Logistic Regression ¡ Support Vector Machine ¡ Neural Network. we calculated the confusion matrix to check the performance of each algorithm and, finally, the ROC curve, which represented the performance of each algorithm on each class of our problem.

References

Iqbal, Z., Qadir, J., Mian, A. N., Kamiran, F., Machine LearningBased Student Grade Prediction: A Case Study, (2017). Arxiv Preprint Arxiv:1708.08744.

Vadivu, G., K. Sornalakshmi,. Applying Machine Learning Algorithms for Student Employability Prediction Using R International Journal of Pharmaceutical Sciences Review and Research (ISN 0976-044X) March-April, (2017) 38-41.

Kim, B. H., Vizitei, E., Ganapathi, V., Gritnet: Student Performance Prediction with Deep Learning. Arxiv Preprint Arxiv:1804.07405, (2018).

Xu, J., Moon, K. H., Van Der Schaar, M., A Machine Learning Approach for Tracking and Predicting Student Performance in Degree Programs. IEEE Journal of Selected Topics in Signal Processing, 11(5) (2017) 742-753.

Pojon, M., Using Machine Learning to Predict Student Performance (Master's Thesis), (2017).

Singh, M., Singh, J., Machine Learning Techniques for Prediction

of Subject Scores: A Comparative Study, (2013). [7] Bendangnuksung, P. P., Students' Performance Prediction Using A Deep Neural Network. International Journal of Applied Engineering Research, 13(2) (2018) 1171-1176.

Pushpa, S. K., Manjunath, T. N., Mrunal, T. V., Singh, A., Suhas,C., Class Result Prediction Using Machine Learning. in 2017 International Conference on Smart Technologies for Smart Nation (Smarttechcon), (2017) 1208-1212). IEEE.

[9] Gerritsen, L., Predicting Student Performance with Neural Network. Tilburg University, Netherlands, (2017).

[10] Solis, M., Moreira, T., Gonzalez, R., Fernandez, T., Hernandez,M., Perspectives to Predict Dropout in University Students with Machine Learning. IEEE International Work-Conference on Bioinspired intelligence (IWOBI), (2018) 1-6. IEEE.