SSRG International Journal of Computer Science and Engineering
ISSN: 2348 – 8387 / https://doi.org/10.14445/23488387/IJCSE-V9I3P102

Volume 9 Issue 3, 5-20, March 2022
© 2022 Seventh Sense Research Group®

Original Article

Career Prediction Classifiers based on Academic Performance and Skills using Machine Learning

Akanksha Pandey[1], L S Maurya[2]

[1,2] Computer Science & Engineering, Shri Ram Murti Smarak College of Engineering and Technology, 13 km, Bareilly-Nainital Road, Ram Murti Puram, Bareilly, U.P, India.

Received: 09 February 2022 Revised: 21 March 2022 Accepted: 27 March 2022 Published: 31 March 2022

***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 discuss our classification in machine learning algorithms to predict the career options for engineering students. The different criteria used to scrutinise the results achieved by these classifiers are accuracy score, confusion matrix, heatmap, percentage accuracy score, and classification report. The research objective is to find the factors that can affect students' decision to choose the right career path using machine learning techniques.*

***Keywords -** Career, Machine Learning, Prediction, Python, Skills, Supervised Learning.*

1. Introduction

Engineering is one of the best career streams apart from medicine, which most students are opting for, some due to interest and 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.

Students face problems choosing the right career path without proper guidance from professional services. 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 choose, as their complete future depends upon this. Thus, we have considered other aspects that will help us choose the right career path based on the academic score and personality, which is important for making their decision.

This research aims to develop a classification model for predicting career options after engineering undergraduate students. This study aims to govern whether the student's academic performance is determined by aptitude or personality and develop a model to analyse students' performance. Machine learning is the best technique to automatically identify and analyse the data and use them to perform the predictions automatically. Processes for many data are analysed to discover patterns and rules after data. A computer can easily process and define these generated rules and patterns to characterise the new data. It is an automatic process that helps to improve the updated data. Therefore, as a result, it helps the students improve their learning activities and helps them analyse their career path for the future.

Classification techniques like KNN, SVM, SGD, Logistics Regression, Decision Tree, and Neural Network are applied to calculate educational performance at the global level after students. The prediction model in terms of student-related variables is assessed.

In this paper, the future career options for students are predicted based on their skills, interest, hobbies, links, etc. the rules learned are portrayed in the context of the decision tree. In this work, the SVM algorithm is preferred to improve prediction accuracy. The application of Machine Learning algorithms involves data analysis, visualisation of data, performance prediction, providing feedback and recommendations, and grouping of students. For this, student performance data is pre-processed to extract features and select them. It also involves data cleaning, tokenisation, sentimental analysis, and removal of words. At the end of the preceding processes, the student's final performance is obtained, which will help them analyse the best career options they should opt for after graduation in engineering. The rest of the paper is summarised as follows:-

Section 2 addresses the literature review. Section 3 represents the proposed work. Section 4 discusses the research design and methodology. Section 5 describes the results and discussions of our research work. Section 6 shows the conclusions of the entire research. Finally, Section 7 lists the references.

2. 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 using machine learning techniques, and it has been found that RPM is the best among various machine learning techniques. (Iqbal,2017)

Vaidu et al. have 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 95.33% accurate, which is all for the Naïve Bayes is 67.67% accurate. (vadiu, 2017)

To predict our future performance after students, Byung-Hak 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 Program. (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 the performance of the students is evaluated. It uses a bi-layered structure that compromises multiple phase predictors and a data-driven approach based on 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. have used three different algorithms, Linear regression, Naïve Bayes classification, and decision tree, on two separate data sets, Roberson and another one featuring an engineering version. As per the result, Naïve Bayes is the best technique used for the first data set as it gives an accuracy of 98%, while the Decision Tree is the best technique for the second database as it gives an accuracy of 78%.

Singh, 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 than 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, that is, 89.06% (Bendangnuksung, 2018).

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 Trees. According to this, DNN achieves 84.3% accuracy, which is better than Machine Learning Techniques. (Pushpa,2017)

Gerritsen L. et al. used data from Learning Management System about educational data using Neural Networks to predict student performance. For this paper, a Moodle Log data set is considered that has a file that contains 4601 students' information. In this paper, the performance of Neural Networks is compared with six classifiers named K-Nearest Neighbor, Naive Bayes, Decision Tree, Support Vector Machine, Logistic Regression, and Random Forest. According to this paper, Neural Network is more accurate than the other six classifiers (Gerritsen, L., 2017).

To predict the University's dropout student list, Hernandez et al. have used four techniques of Machine Learning to analyse the performance. These four techniques include Random Forest, Support Vector Machines, Logistics Regression, and Neural Networks. For this research paper, the dataset of institute Technological de Costa Rica (ITCR) students is used who enrolled between 2011 and 2016. Among these four algorithms, the Random Forest algorithm is preferred to be best to predict University dropouts. (Hernandez,2018)

García-Peñalvo et al. have used the new artificial Neural Networks to predict the students' career paths based on the dataset. He proposed a data-driven system to collect the data to predict all the future career paths available. (García-Peñalvo, 2018)

The machine learning algorithms such as the SVM decision tree and X.G. boost are used by K.Sripath Roy et al. to create a model of student career predictions. Among these algorithms, the Support Vector Machine gives the most accurate result, that is 90.3%.(k. Sripath Roy,2019)

Mostafa et al. proposed artificial neural network technology to predict students' academic performance. In this study, the model of neural networks is created, which predictor students' GPA by using their personal information, place of resistance, and academic information. According to this model, the accuracy of prediction is 73.68%. (Mostafa,2021)

3. 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 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, and Bhakti Ratnaparkhi's paper predicted that different classifiers accurately predict a student's career (K.S. Lokesh, 2015).10

Let's look at these machine learning algorithms used to develop the classifier.

It will help students to predict careers after their graduation in engineering.

3.1. K Nearest Neighbor

The KNN algorithm is used for boat regression and classification problems. This algorithm categorises all the cases into new ones based on key neighbours' majority votes. It is most suitable for classification rather than regression. Here are the following steps that are involved in the process of K&N:

- Select some neighbours, for example, K =10;
- Now calculate the Euclidean distance, that is, the distance between two different data points;
- Now let's categorise the neighbours based on distance; for example, the four nearest neighbours are in category A; after that three nearest neighbours to be category B, and the remaining three be in category C;
- Now, this is the new data that has been prepared by using KNN.

3.2. Support Vector Machine

Supervised learning at Corrigan is used for classification and regression problems. The goal of the SVM algorithm is to mark the decision boundaries that are also called a hyperplane. This algorithm divides different training data sets into various classes based on the hyperplane. In this algorithm, the n-dimensional space is plotted where it belongs to the number of features with the value of a particular coordinate. It helps to maximise the classifiers' margin.

3.3. Stochastic Gradient Descent

The Stochastic Gradient Descent classifier is the most efficient technique for machine learning classification problems under the Linux support vector machine and logistic regression. The Stochastic Gradient Descent classifier is merely an exaggeration technique that has not corresponded to the specific dataset of machine learning classification techniques. It is the most efficient algorithm and easy to implement. Here, the linear support vector machine has been used as a classifier so that the Stochastic Gradient Descent algorithm can be applied to optimise the accurate result.

3.4. Logistic Regression

It is a regression technique of machine learning used to perform classification problems. Logistic regression is used to calculate the probability of the given class-specific data. If there is more than 50% accuracy, the value is assigned to that class, and if the result is less than 50%, it will assign the value to the other class. Thus, it is stated as a binary classifier.

3.5. Decision Tree

This technique is used to conclude some conditions and flag this as a node and a leaf node. The first decision mode tells us about the attribute to be selected, while the leaf node tells us about a class. It is a primary node, also called root node pictures selected based on two different methods:

- Gini index Method
- information Gain Method

To obtain a result, entropy is calculated, which is the measure of the volatility of the data. It is the impurity that is present in the data set.

The formula of Entropy is as follows: E = -p*log2(p) – q*log2(q)

Hence, the information gained is as follows:

Gain (T, X) = Entropy(T) – Entropy (T, X)

Where,

X belongs to the split node from T, and;
T belongs to the Parent node before the split.
Whereas the Gini index is calculated as follows:
Gini = 1 - Σ(pi)2
Where P belongs to the probability of a particular class

3.6. Neural Network

These artificial neural networks are also called simulated neural networks of machine learning techniques that involve deep learning, like the human brain, to process the data, such as detecting objects, recognising speech, and classifying objects. This machine-learning classification technique constructs the model using the hidden elements during the training process to produce multiple data layers.

3.7. Gaussian Naive Bayes

Gaussian Naive Bayes is a type of Naive Bayes that follows Gaussian normal distribution and is established on the Bayes theorem. It works on continuous data, and we explored it to calculate conditional probability.

3.8. Random Forest

It is one of the supervised learning. This algorithm is used to develop the classifier, containing several decision trees. Random forest trains itself with the help of the bagging method, which is a combination of learning models that aids the overall result.

4. Research Design and Methodology

4.1. Data Collection

Data is collected by designing a questionnaire containing 10 questions in which 8 columns are created, which will be treated as input variables, and 1 column is created for dependent variables, which will be our career option selection dependent on each input variable. A total of 330 observations were collected using the questionnaire. Questionnaire columns are-

- User name
- What is your percentage in class 10?
- What is your percentage in class 12/Diploma?
- What is your marks percentage in B.Tech to date?
- Rate your aptitude skill
- Rate your communication skill
- Rate your technical skill
- Rate your management skill
- Rate your general knowledge
- Which career option will you prefer after graduation?

4.2. Data pre-processing

Data pre-processing consists of techniques to transform unstructured data into structured data. The designed questionnaire contains raw values where we have to pre-process the raw data before using machine learning classification algorithms.

4.2.1. Feature Reduction

Obtained data contains time-stamp, E-mail and contact no., which are not required in our final model, so we will remove this column.

4.2.2. Encoding Categorical Features

Collected data contains only categorical features. First, we must encode each feature before training our data because machine learning algorithms perform well with numerical data. For independent features, we encoded the rating value of the questionnaire as Excellent - 5, Very good - 4, and good - 3. Average - 2, Poor - 1 and for dependent feature, we used a python library known as label encoder for encoding this feature. After encoding the dependent feature, the assigned number of each class is classified as - Gov Job - 0, M.Tech/ME/MS -1, MBA - 2, Others - 3, Prvt Job - 4.

4.3. Classifiers

For this study, we have selected 8 popular machine-learning classification algorithms. These are -

- KNN
- SVM (Support vector machine)
- Stochastic gradient descent
- Logistic Regression (For multiclass)
- Decision Tree
- Neural Network
- Gaussian Naive Bayes
- Random Forest

4.4. Training and Testing

Since we have 330 observations, we decided to train our model on 80% of the data and 20% for testing.

5. Results & Discussions

5.1. Accuracy

Now that we have prepared our algorithm, the obtained accuracies are shown in Table 1, where the accuracy of each algorithm is calculated with their R2 score, at which random state we get this accuracy and how much time it took for an algorithm to train itself, i.e., Execution Time Hyperparameter tuning is done on every algorithm to get the desired accuracy.

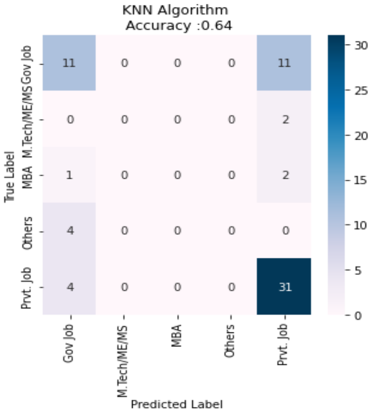
Table 1. Calculated accuracy score, Hyperparameter tuning, Random state, and execution time for each classifier

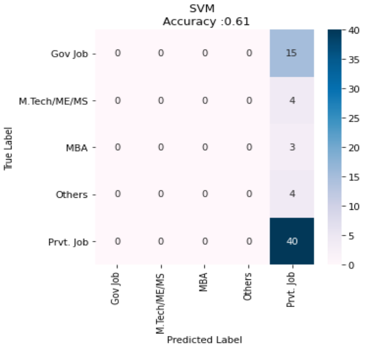
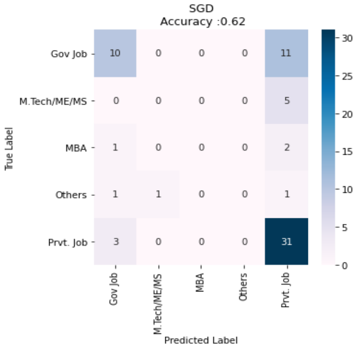
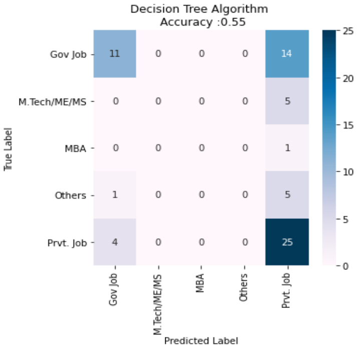
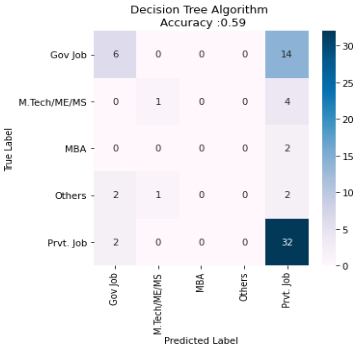
S. No.	Classification Algorithms	Hyperparameter	Execution Time (seconds)	Random State	Accuracy Score
1	K - Nearest Neighbor	n_neighbors = 18	0.00524	937	63.64%
2	Support Vector Machine	kernel='rbf', gamma=0.5, c=0.1, probability =True	0.06546	440	60.61%
3	Stochastic Gradient Descent	loss='hinge', random_state=100	0.0215	42	62.12%
4	Logistic Regression	multi_class='multinomial', solver='lbfgs'	0.05089	42	54.55%
5	Decision Tree	max_depth = 5, random_state = 200	0.00489	317	59.09%
6	Neural Network	no. of epochs = 100 batch_size = 30	3.14762	42	45.45%
7	Gaussian Naive Bayes	Default	0.00517	160	60.61%
8	Random Forest	n_estimators = 100, random_state = 42	0.21717	42	46.97%

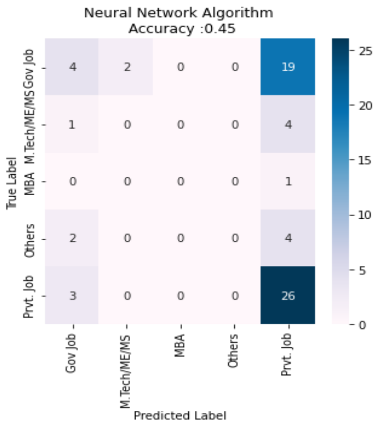
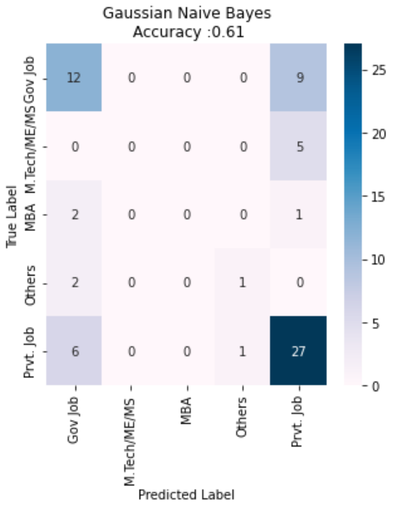
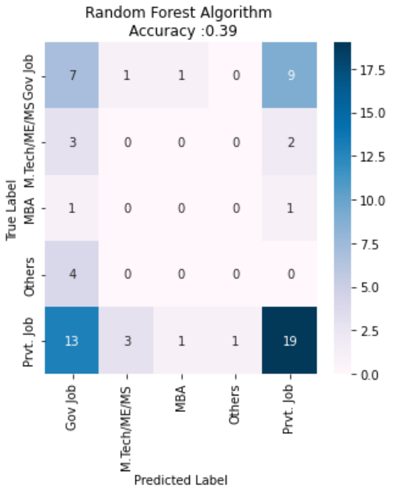
5.2. Confusion Matrix and Heatmap

The confusion matrix is a great technique to check the performance of any classifier. We have a total of 5 classes, so the size of the confusion matrix will be 5x5. A graphical representation of the confusion matrix, i.e., the heatmap, is also obtained from the confusion matrix.

Table 2. The confusion matrix and heatmap of each classifier are shown

S. No.	Classification Algorithms	Confusion Matrix	Heatmap
1	K - Nearest Neighbor	[[11 0 0 0 11] [0 0 0 0 2] [1 0 0 0 2] [4 0 0 0 0] [4 0 0 0 31]]	

2	Support Vector Machine	[[0 0 0 0 15] [0 0 0 0 4] [0 0 0 0 3] [0 0 0 0 4] [0 0 0 0 40]]	
3	Stochastic Gradient Descent	[[10 0 0 0 11] [0 0 0 0 5] [1 0 0 0 2] [1 1 0 0 1] [3 0 0 0 31]]	
4	Logistic Regression	[[11 0 0 0 14] [0 0 0 0 5] [0 0 0 0 1] [1 0 0 0 5] [4 0 0 0 25]]	
5	Decision Tree	[[6 0 0 0 14] [0 1 0 0 4] [0 0 0 0 2] [2 1 0 0 2] [2 0 0 0 32]]	

6	Neural Network	[[4, 2, 0, 0, 19], [1, 0, 0, 0, 4], [0, 0, 0, 0, 1], [2, 0, 0, 0, 4], [3, 0, 0, 0, 26]]	
7	Gaussian Naive Bayes	[[12 0 0 0 9] [0 0 0 0 5] [2 0 0 0 1] [2 0 0 1 0] [6 0 0 1 27]]	
8	Random Forest	[[7 2 1 0 8] [3 0 0 0 2] [0 0 1 0 1] [4 0 0 0 0] [9 3 1 1 23]]	

5.3. Classification Report

Now that we've got our predicted value and confusion matrix, we create a classification report for each classifier. A classification report tells the prediction quality with the help of classification metrics such as precision, recall, f1-score and support.

Table 3. Classification report

S. No.	Classification Algorithms	Classification report
1	K - Nearest Neighbor	precision-recall f1-score support Gov Job 0.55 0.50 0.52 22 M.Tech/ME/MS 0.00 0.00 0.00 2 MBA 0.00 0.00 0.00 3 Others 0.00 0.00 0.00 4 Prvt. Job 0.67 0.89 0.77 35 accuracy 0.64 66 macro avg 0.24 0.28 0.26 66 weighted avg 0.54 0.64 0.58 66
2	Support Vector Machine	precision recall f1-score support Gov Job 0.00 0.00 0.00 15 M.Tech/ME/MS 0.00 0.00 0.00 4 MBA 0.00 0.00 0.00 3 Others 0.00 0.00 0.00 4 Prvt. Job 0.61 1.00 0.75 40 accuracy 0.61 66 macro avg 0.12 0.20 0.15 66 weighted avg 0.37 0.61 0.46 66
3	Stochastic Gradient Descent	precision recall f1-score support Gov Job 0.67 0.48 0.56 21 M.Tech/ME/MS 0.00 0.00 0.00 5 MBA 0.00 0.00 0.00 3 Others 0.00 0.00 0.00 3 Prvt. Job 0.62 0.91 0.74 34 accuracy 0.62 66 macro avg 0.26 0.28 0.26 66 weighted avg 0.53 0.62 0.56 66
4	Logistic Regression	precision recall f1-score support

		Gov Job 0.69 0.44 0.54 25 M.Tech/ME/MS 0.00 0.00 0.00 5 MBA 0.00 0.00 0.00 1 Others 0.00 0.00 0.00 6 Prvt. Job 0.50 0.86 0.63 29 accuracy 0.55 66 macro avg 0.24 0.26 0.23 66 weighted avg 0.48 0.55 0.48 66
5	Decision Tree	precision recall f1-score support Gov Job 0.60 0.30 0.40 20 M.Tech/ME/MS 0.50 0.20 0.29 5 MBA 0.00 0.00 0.00 2 Others 0.00 0.00 0.00 5 Prvt. Job 0.59 0.94 0.73 34 accuracy 0.59 66 macro avg 0.34 0.29 0.28 66 weighted avg 0.52 0.59 0.52 66
6	Neural Network	precision recall f1-score support Gov Job 0.40 0.16 0.23 25 M.Tech/ME/MS 0.00 0.00 0.00 5 MBA 0.00 0.00 0.00 1 Others 0.00 0.00 0.00 6 Prvt. Job 0.48 0.90 0.63 29 accuracy 0.45 66 macro avg 0.18 0.21 0.17 66 weighted avg 0.36 0.45 0.36 66
7	Gaussian Naive Bayes	precision recall f1-score support Gov Job 0.55 0.57 0.56 21 M.Tech/ME/MS 0.00 0.00 0.00 5 MBA 0.00 0.00 0.00 3

		Others 0.50 0.33 0.40 3 Prvt. Job 0.64 0.79 0.71 34 accuracy 0.61 66 macro avg 0.34 0.34 0.33 66 weighted avg 0.53 0.61 0.56 66
8	Random Forest	precision recall f1-score support Gov Job 0.30 0.39 0.34 18 M.Tech/ME/MS 0.00 0.00 0.00 5 MBA 0.33 0.50 0.40 2 Others 0.00 0.00 0.00 4 Prvt. Job 0.68 0.62 0.65 37 accuracy 0.47 66 macro avg 0.26 0.30 0.28 66 weighted avg 0.47 0.47 0.47 66

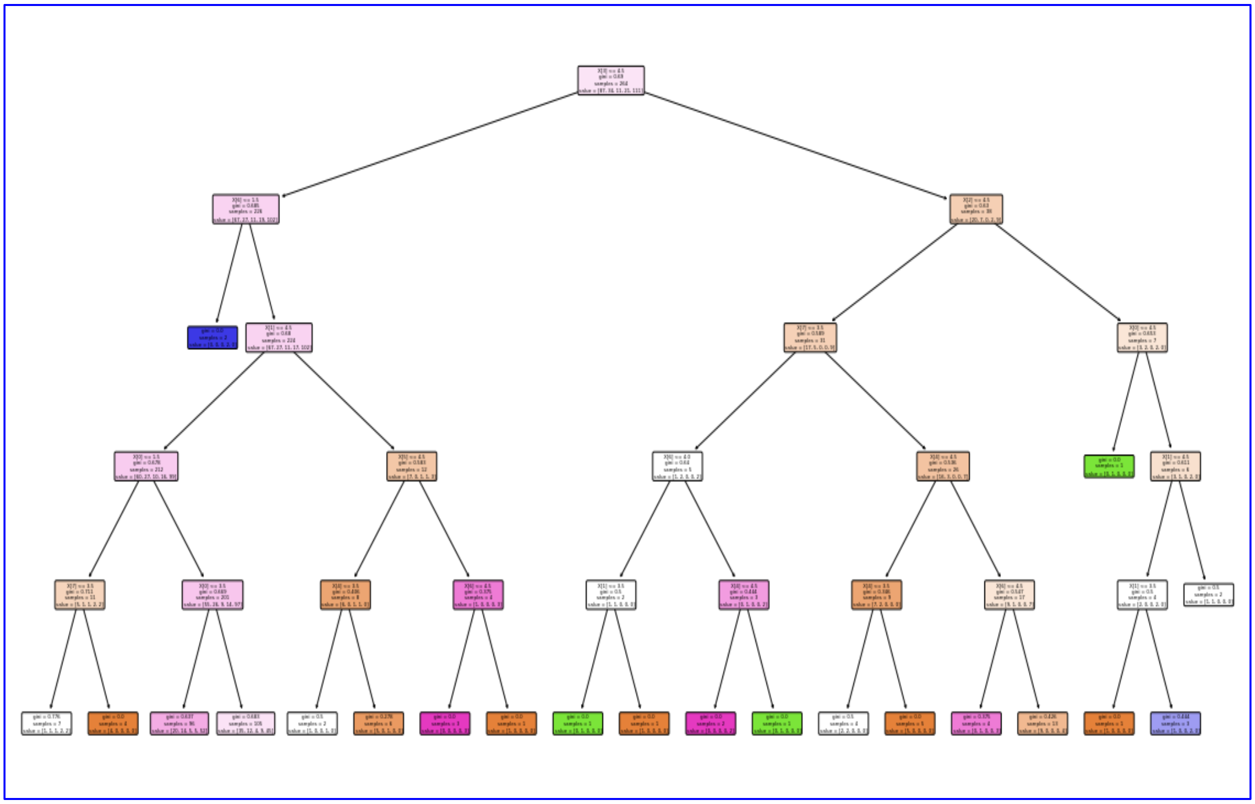


Fig. 1 Decision tree plot

5.4. Decision Tree Visualisation

The decision tree plot is shown in Fig. 1. For the Decision Tree Classifier, we chose max_dept = 5 with random state 200, and the rest default parameters are chosen, i.e., Gini impurity and entropy.

5.5. Mean Squared Error, R2 Value, Log Loss

Mean square error estimates the error between the actual value and the predicted value. The machine learning optimiser reduces this error to get the best fit line for prediction. The coefficient of determination (R2 Score) is calculated to check the performance of each algorithm.

It tells how our predicted values vary from the actual value's mean. Loss loss is evaluating the probabilities of each prediction concerning actual data. We can see from the table that the best-fit algorithms are according to these Values: for lowest mean squared error of the Support vector machine is (4.42), following the R2 value of the k-nearest neighbour that is (-0.38) and the log loss value of the neural network that is (1.2).

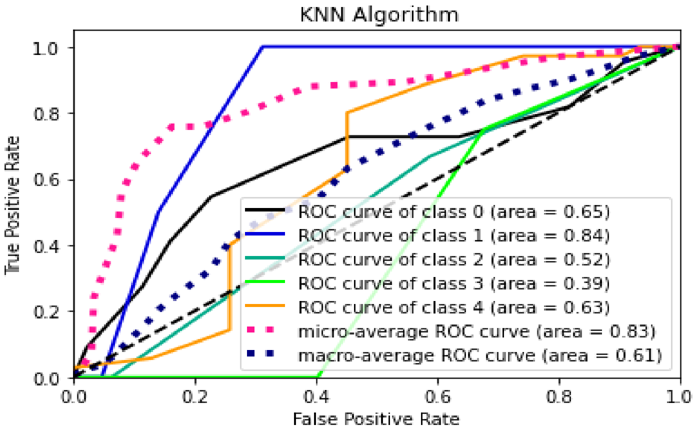
5.6. ROC Curve (Receiving operating characteristic) curve.

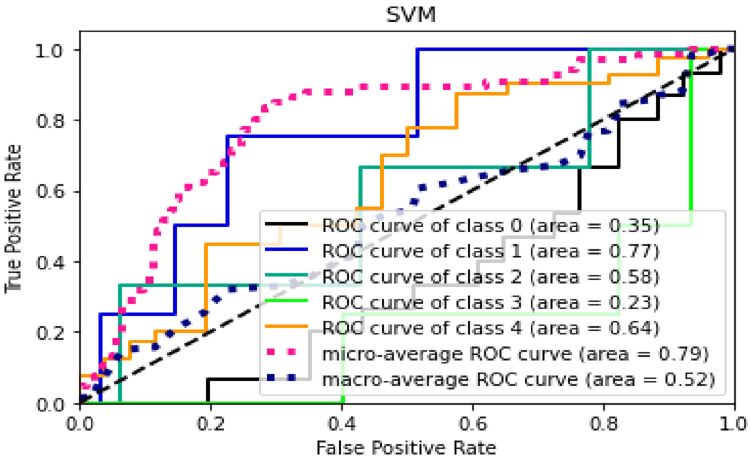
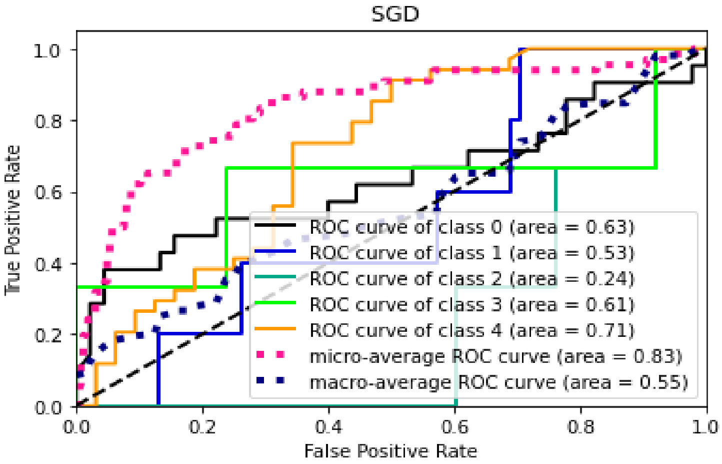
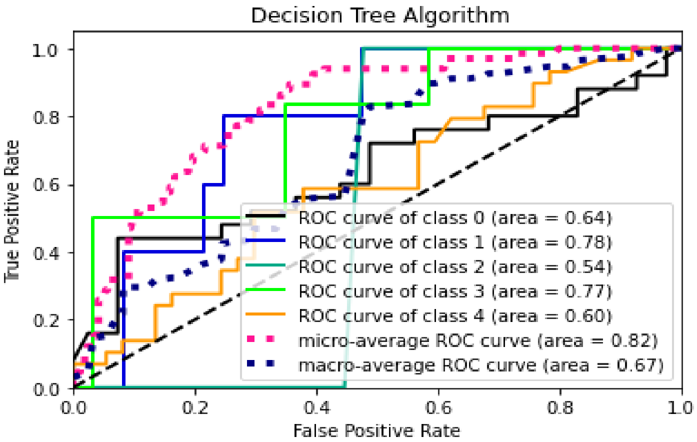
ROC curve is obtained for each classifier (see Table 4). ROC curve shows the performance of classifiers on each label, and micro and average macro ROC is obtained, which shows each class's avg and aggregate avg. The closer the ROC curve is to the top right corner of the roc space, the better the classifier's performance will be, while if the ROC curve is closer to the diagonal line, the lesser the performance will be.

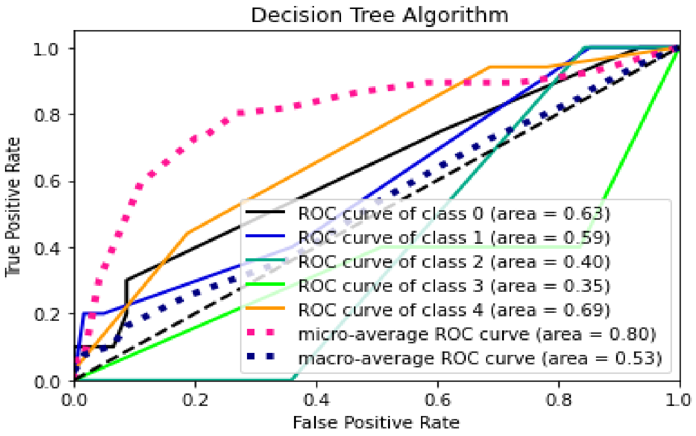
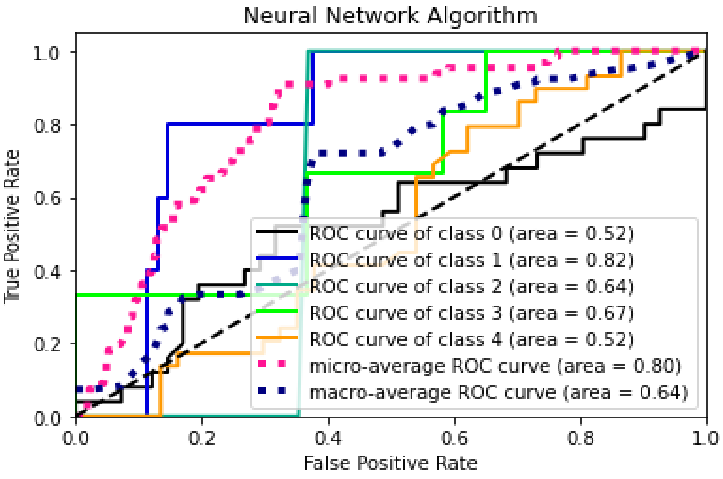
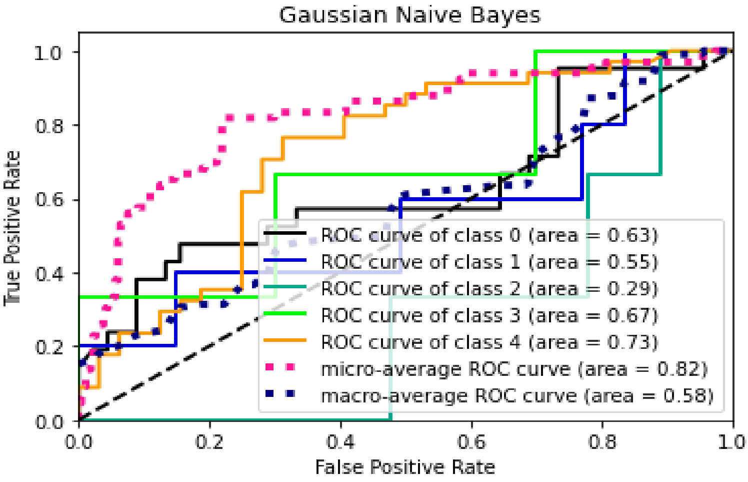
Table 4. Mean square error, R2 value, log loss

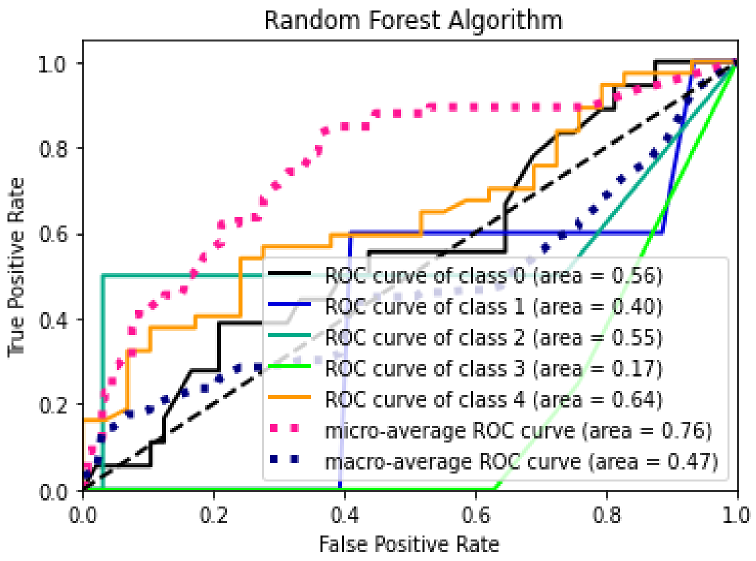
S. No	Algorithm	Mean Squared Error	R2 Value	Log Loss
1	K – Nearest Neighbor	4.64	-0.38	2.07
2	Support Vector Machine	4.42	-0.54	1.25
3	Stochastic Gradient Descent	4.47	-0.35	1.19
4	Logistic Regression	5.32	-0.55	1.17
5	Decision Tree	4.91	-0.51	3.57
6	Neural Network	6.32	-0.85	1.2
7	Gaussian Naïve Bayes	4.79	-0.44	1.22
8	Random Forest	5.62	-0.78	3.32

Table 5. ROC curve representation

S. No.	Classification Algorithms	ROC Curve
1	K - Nearest Neighbor	

2	Support Vector Machine	
3	Stochastic Gradient Descent	
4	Logistic Regression	

5	Decision Tree	
6	Neural Network	
7	Gaussian Naive Bayes	

8	Random Forest	

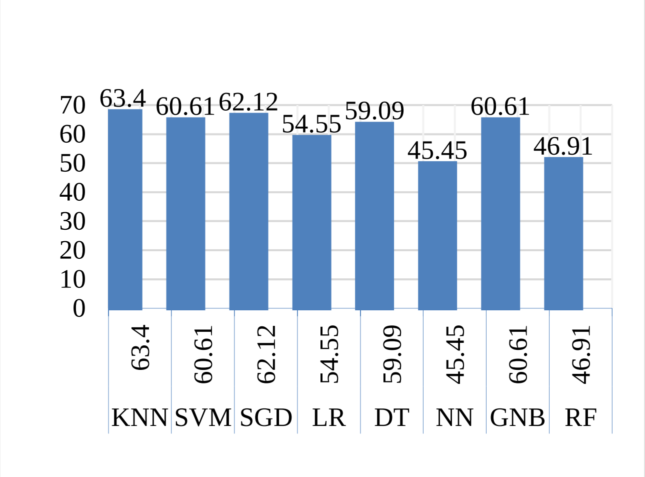


Fig. 2 Bar Graph to represent the accuracy of each algorithm

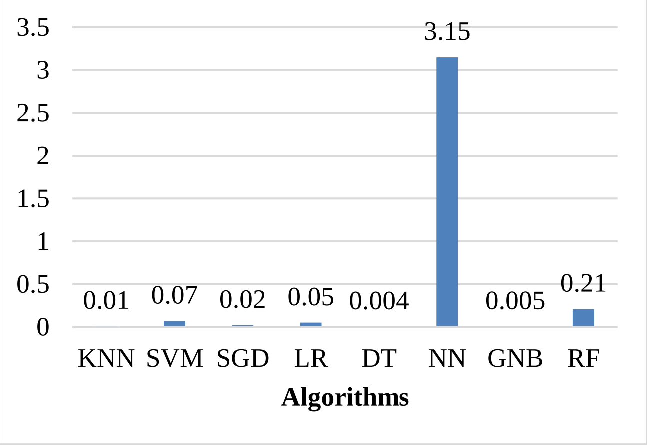


Fig. 3 Bar graph to represent execution time

6. Conclusion

In this study, we have to create a classification model to predict career options for an undergraduate student. Various input features such as the student's mark percentage in 10th, 12th, B.Tech/Diploma, skills in communication etc., are taken into consideration, and the 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, which is obvious from Fig. 2. 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 (see Fig. 3). It has been seen that neural networks take a lot of time to train the data and give very few scores, so neural networks are not a good fit for this problem.

Classification matrices are estimated and compared for each algorithm using classification metrics like precision, recall, f-1 score, and support. Then, 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

[1] Zafar Iqbal et al., *Machine Learning-Based Student Grade Prediction: A Case Study*, Arxiv Preprint Arxiv:1708.08744, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[2] G. Vadivu, and Sornalakshmi, "Applying Machine Learning Algorithms for Student Employability Prediction Using R," *International Journal of Pharmaceutical Sciences Review and Research*, vol. 43, no. 1, pp. 38-41, 2017. [Google Scholar] [Publisher Link]

[3] Byung-Hak Kim, Ethan Vizitei, and Varun Ganapathi, "Gritnet: Student Performance Prediction with Deep Learning," *Arxiv Preprint Arxiv:1804.07405*, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[4] Jie Xu et al., "A Machine Learning Approach for Tracking and Predicting Student Performance in Degree Programs," *IEEE Journal of Selected Topics In Signal Processing*, vol. 11, no. 5, pp. 742-753, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[5] M. Pojon, *Using Machine Learning to Predict Student Performance* (Master's Thesis), 2017. [Google Scholar] [Publisher Link]

[6] Mamta Singh et al., "Machine Learning Techniques for Prediction of Subject Scores: A Comparative Study," *International Journal of Computer Science and Network*, vol. 2, no. 4, pp. 77-79, 2013. [Google Scholar] [Publisher Link]

[7] Bendangnuksung, and Prabu P, "Students' Performance Prediction Using a Deep Neural Network," *International Journal of Applied Engineering Research*, vol. 13, no. 2, pp. 1171-1176, 2018. [Publisher Link]

[8] S K Pushpa et al., "Class Result Prediction Using Machine Learning," *In 2017 International Conference on Smart Technologies for Smart Nation* (*Smarttechcon*), *IEEE*, pp. 1208-1212, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[9] Gerritsen, L, *Predicting Student Performance with Neural Network*, Tilburg University, Netherlands, 2017. [Google Scholar]

[10] Martin Solis et al., "Perspectives to Predict Dropout in University Students with Machine Learning," *IEEE International Work-Conference on Bioinspired Intelligence (IWOBI*), *IEEE* pp. 1-6, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[11] Francisco García-Peñalvo et al., "Proposing a Machine Learning Approach to Analyze and Predict Employment and Its Factors," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 5, pp. 39-45, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[12] Dileep Chaudhary et al., "Student Future Prediction Using Machine Learning," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, vol. 5, no. 2, pp. 1104-1108, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[13] Lamiaa Mostafa, and Sara Beshir, "University Selection Model Using Machine Learning Techniques," *The International Conference on Artificial Intelligence and Computer Vision*, vol. 1377, pp. 680-688, 2021. [Google Scholar] [Publisher Link]

[14] Nikita Gorad et al., "Career Counselling Using Data Mining," *International Journal of Engineering Science and Computing (IJESC)*, vol. 5, no. 4, pp. 10271-10274, 2017. [Publisher Link]

[15] Roopkanth, K, and Bhavana, V, "Student Career Area Prediction using Machine Learning," *IEEE*, 2018.

[16] Zahyah Alharbi et al., "Using Data Mining Techniques to Predict Students at Risk of Poor Performance," *SAI Computing Conference (SAI), IEEE*, pp. 523-531, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[17] Ministry of Research Technology and Higher Education of Republic of Indonesia Higher Education Statistical Year Book 1-194, 2017.

[18] Flasiński M, *Introduction to Artificial Intelligence*, (Cham: Springer International Publishing) History of Artificial Intelligence, 2016.

[19] Jie Lu et al., "Recommender System Application Developments: A Survey," *Decision Support Systems*, vol. 74, pp. 12-32, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[20] D Kurniadi et al., "Estimated Software Measurement Based on the use Case for Online Admission System," *IOP Conference Series: Materials Science and Engineering*, vol. 434, pp. 012062, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[21] Arindam K. Das, and Esteban Rodriguez-Marek, "A Predictive Analytics System for Forecasting Student Academic Performance: Insights from a Pilot Project at Eastern Washington University," *2019 Joint 8th International Conference on Informatics, Electronics and Vision, (ICIEV) & 3rd International Conference on Imaging, Vision and Pattern Recognition, (IVPR)*, *IEEE*, pp. 255– 262, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[22] Hussein Altabrawee et al., "Predicting Students' Performance using Machine Learning Techniques," *Journal of University of Babylon for Pure and Applied Sciences*, vol. 27, no. 1, pp. 194–205, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[23] Mukesh Kumar, A.J. Singh, and Disha Handa, "Literature Survey on Student's Performance Prediction in Education using Data Mining Techniques," *International Journal of Education and Management Engineering*, vol. 7, no. 6, pp. 40–49, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[24] Phocenah Nyatanga, and Sophia Mukorera, "Effects of Lecture Attendance, Aptitude, Individual Heterogeneity and Pedagogic Intervention on Student Performance: a Probability Model Approach," *Innovations in Education and Teaching International*, vol. 56, no. 2, pp. 195–205, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[25] Sadia Zaman Mishu, and S. M. Rafiuddin, “Performance Analysis of Supervised Machine Learning Algorithms for Text Classification,” *In 2016 19th International Conference on Computer and Information Technology (ICCIT), Dhaka, Bangladesh, IEEE*, pp. 409–13, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[26] Radhika R Halde, “Application of Machine Learning Algorithms for Betterment In Education System,” *In 2016 International Conference on Automatic Control and Dynamic Optimisation Techniques (ICACDOT), Bangalore, India: IEEE*, pp. 1110–14, 2016. [CrossRef] [Google Scholar] [Publisher Link]