ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/351282903

Performance Evaluation of Machine Learning Predictive Analytical Model for Determining the Job Applicants Employment Status

Article *in* Malaysian Journal of Applied Sciences · April 2021

DOI: 10.37231/myjas.2021.6.1.276

CITATIONS
11

READS
672

4 authors, including:

Awujoola Joel Olalekan
Nigerian Defence Academy
12 PUBLICATIONS **42** CITATIONS
SEE PROFILE

Philip Odion
Nigerian Defence Academy
17 PUBLICATIONS **50** CITATIONS
SEE PROFILE

Martins Ekata Irhebhude
Nigerian Defence Academy
52 PUBLICATIONS **172** CITATIONS
SEE PROFILE

All content following this page was uploaded by Awujoola Joel Olalekan on 03 May 2021.

The user has requested enhancement of the downloaded file.

MALAYSIAN JOURNAL OF APPLIED SCIENCES 2021, VOL 6 (1), 67-79
E-ISSN:0127-9246 (ONLINE)
https://doi.org/10.37231/myjas.2021.6.1.276
https://journal.unisza.edu.my/myjas

ORIGINAL ARTICLE

Performance Evaluation of Machine Learning Predictive Analytical Model for Determining the Job Applicants Employment Status

***Awujoola J Olalekan[1], Philip O Odion[1], Martins E Irhebhude[1], Halima Aminu[2]**

[1]Department of Computer Science, Nigerian Defence Academy, Nigeria
[2]Kano University of Science and Technology, Wudil, Nigeria

[*]Corresponding author: ojawujoola@nda.edu.ng

Received: 04/02/2021, Accepted: 28/04/2021, Published: 30/04/2021

Abstract

Several higher institution of learning faces issue or difficulty of turning out more than 90% of their graduates who can competently satisfy and meet the requirements of the industry. However, the industry is also confronted with the difficulty of sourcing skilled tertiary institution graduates that match their needs. Failure or success of any organization depends mostly on how its workforce is recruited and retained. Therefore, the selection of an acceptable or satisfactory candidate for the job position is one of the major and vital problems of management decision-making. This work, therefore, proposes a modern, accurate and worthy machine learning classification model that can be deployed, implemented, and put to use when making predictions and assessments on job applicant's attributes from their academic performance datasets in other to meet the selection criteria for the industry. Both supervised and unsupervised machine learning classifiers were considered in this work. Naïve Bayes, Logistic Regression, support vector machine (SVM). Random Forest and Decision tree performed well, but Logistic Regression outperformed others with 93% accuracy.

Keywords: Machine learning, classification, Naïve Bayes, Logistic Regression, Support Vector Machine (SVM), Random Forest and Decision tree

Introduction

Unemployment is a contemporary matter that take place globally and thus initiate unfavourable impacts on worldwide (Othman et al., 2018). Consequently, graduate employability is one of the noteworthy elements to be spotlighted in unemployment problem. There are many other features that influence graduate employability among is excellent academic performance (cumulative grade point average, CGPA), this has been the major prepotent element while determining individual's employment status. Nevertheless, many scholars have shown that it's not only CGPA that determines the graduate employability; in reality many other factors may actually effect the graduate achievement in getting a job (Othman et al., 2018).

The failure or success of any organisation is directly linked or associated with the processes or techniques adopted by its human resource management unit when recruiting its workforce and how they retain them. Many organisations keep huge amounts of data and information on entrance processes and evaluations. This information, nevertheless, are in many cases left un-used, or rather analyzed via elementary statistical methods. Human Resource unit are required to sort or group the applicants profiles manually before they can select the appropriate candidate for the position they are recruiting. This however, required considerable amount of time, this even dependent on the number of applications.

For almost two decades now, information retrieval systems (IRS) have been adopted for used by human resources management (HRM) units to collate, arranged and classified job applicants curriculum vitae, recommendation letter from recognized individuals, transcripts, certificates and results. However, these type of systems are becoming pervasive in such a way that almost 74% of major U.S. organizations adopted some form of electronic selection system to assist in the hiring process (Stone D. L., Deadrick, Lukaszewski, & Johnson, 2015). Therefore, these technique has been able to conserve money and time required in the recruitment process.

Personnel employment prediction is a classification task that required the use of differs techniques in natural language processing and machine learning while attempting to predict a placement based on skills, knowledge, experience, qualifications, job requirements, interests, etc. Also, Selection is a management resolution whereby a company or organization choose whether to employ individuals using their score on a single assessment, e.g from interview or aptitude test or it can as well be based on a composite of many assessments.

Over the last few years, HRM turned out to be the prime focus of organization attention and have contributed immensely to the time and assets of developing companies. Failure or success of any organization is linked to the method and strategy adopted while recruiting its workforce and how they are retained (Jazani, 2000). Therefore, satisfactory applicant selection for particular job position is one of the main problem of management decision-making. Although, sometimes data gathered or acquired publicly not have value or meaningless. But when that same data is refined and processed to become part of knowledge and information, it may create great value. On that account, it is advisable that organisations should acquire very strong skills and technique in data-processing, for example machine learning and data mining (Khaef, Ahmad, Mottaghi, & MV, 2007). In the current era, new intelligent technology has created tools to aid humankind in transforming large volumes of data into knowledge and information (Khaef, Ahmad, Mottaghi, & MV, 2007).

Companies have long understood that hiring the best employees produces a competitive advantage which is very difficult for the opponents to match. Among the major important challenges in sourcing for the good applicant for an employment offer is its imprecise nature which is when it's been influenced by "feel" and not as much as by skills and talent. For couple of decades, information retrieval systems (IRS) have been utilized by department of human resource management (HRM) services to classify and filter job applicants in line with some varieties of lay down rules of weighted features collected from curriculum vitae and cover letters, recommendation letters, interviews with the applicant, and other associated materials like certificates of attendance and result transcript (Stone, Deadrick, Lukaszewski, & Johnson, 2015). This kind of models have been proved to be able to save lots of financial commitments and time in the recruiting exercise (Zielinksi, 2020).

Predictive modelling is a technique whereby a model is designed and developed to predict the outcome. However, if the outcome of the created model is categorical, then it is referred to as classification but if it's numerical, then it is regression (Nor Azziaty, Kian Lam, & Chen Kim, 2017). Clustering is under the descriptive modelling and can be refer to as the allocation of observations into clusters. Therefore, when observation are similar in the same cluster, then association rules can generate meaningful associations amongst observation.

Prediction and Classification are basically the crucial and vital tasks in data mining and machine learning. Choice of algorithm selection is also important and often depends on the type of data (ie ratio, nominal, interval or ordinal,) that will be used. Machine learning category

is provided for every data mining algorithms and different data mining algorithms designed to be used to set data based on knowledge.

In this research study, we propose a data analytics approach, that is capable of been used as a decision support tool for personnel selection in real-world scenario and to improve recruitment decisions of applicants to specific job positions. This work considered five different machine learning model with Naïve Bayes, Logistic Regression, support vector machine (SVM). Random Forest and Decision tree, out of which after performance evaluation is computed the best model will be recommended for personnel employment selection. Therefore, the output of these models is the probabilities of successful recruitment per employee and job.

Prediction and Classification Techniques

Prediction and classification techniques are few of the vital data mining and machine learning task that have been widely used by scientist in the area of prediction analysis (Nor Azziaty, Kian Lam, & Chen Kim, 2017). Several studies on employment predictions and classifications often used the following classifiers, such as Neural Network, Decision Tree, Naive Bayes, Random Forest, Logistic Regression and more. Below are some related work done on personnel employment selection.

Novel Neural Network (NN) was employed to predict the unemployment rate based on information from the web (Xu, W., Li, Z., Cheng, C., & Zheng, 2012). The study established that the efficiency and effectiveness of the proposed method served as a potential alternative for forecasting the unemployment trend.

Sequential Minimal Optimization (SMO), Bayesian, Decision Trees, Multilayer perceptrons and Ensemble Method were proposed by (Mishra, 2016) where emotional skills, socio-economic conditions and academic achievement were considered as to build predictive models in the data set. Decision tree algorithm proved to be more accurate and suitable to predict the employability of students.

Sapaat, Mustapha, Ahmad, & Chamili, (2011) developed Graduate Employability model with Bayesian and Decision tree to ascertain if graduate students can be employed or un-employed. J48 a variant of decision tree came up with the best accuracy when compared with Bayesian classifier.

Author (Tajul, Ab, & Yusof, 2016) proposed a model with five machine learning classifiers; Logistic Regression, KNearest Neighbor, Naive Bayes, Multilayer Perceptron, and J48 Decision Tree. The model is to predict if graduates will be employed in the private or public sector, un-employed or continue their education. However, Logistic Regression algorithm outperformed other classifiers.

Shahiri, Husain, & Rashid, (2015) considered five classification tasks with machine learning algorithms such as KNN, ANN, Naive Bayes, Decision Tree and SVM to predict student performance. Attributes used in this study are demography, internal and external evaluation. The results of the experiment revealed that the prediction error of Neural Network is less and has the highest classification accuracy than other tasks.

Aziz, Ismail, Ahmad, & Hassan, (2015) proposed a framework for predicting the students' performance. The predictive model utilize Naïve Bayes classifier and performed classification on six features such as income, families, university entry mode, race, gender, hometown, and CGPA. The study found that gender, hometown and family income contribute immensely to academic achievement.

The author (Jantawan, B., & Tsai, C., 2014) proposed a graduate employability model that employs Bayesian methods to determine some major and relevant attributes of graduate employability in the dataset. The work considered using the following algorithms: Naive Bayesian, Naive Bayesian Simple, Averaged One- Dependence Estimators with subsumption resolution, Naive Bayesian Updateable, Averaged One-Dependence Estimators, and Bayesian networks for accuracy computation and further compare the accuracy of each algorithm under the model. From the results 3 factors (work province, occupation type, and times find work) has a direct effect on the employability status of the graduate. The results revealed that Averaged One-Dependence Estimators with subsumption resolution (AODEsr)

algorithm outperformed other with accuracy of 98.3% followed by the AODE with 96.1% accuracy.

The author (Othman et al., 2018) proposes data mining techniques to determine factors that affects the graduate's employability status. Three classification algorithms were considered for this work Support Vector Machines, Artificial Neural Networks and Decision Tree. Their accuracy performance were being compared for the best models. The results revealed that decision tree J48 yields higher percentage of classification accuracy of 66.0651% compared to other techniques. However, after parameter tuning the result increased to 66.1824%.

Graduates Employment model was proposed by Jantawan B. &., (2013) to predict if student graduates are unemployable or in an un-determined state by combining Cross Industry Standard Process for Data Mining methodologies and Process knowledge discovery methodologies to build the model. The work determined from the dataset some attributes that can positively affects the employment status of the graduate after 12 months of graduation. The work further carried out performance evaluation between the Bayesian and Decision tree algorithm. Bayesian algorithm has the highest accuracy of 99.77%.

Aziz, M.T., & Yusof, Y. (2016) proposes a model to classify the graduate's applicant into employed, unemployed or further study using, Logistic regression, Multilayer perceptron, Naive Bayes, k-nearest neighbor and Decision tree J48. The result revealed that Logistic regression output the highest percentage accuracy of 92.5%.

The author (Masethe & Masethe, 2014) conducted a research to predict the Work Integrated Learning (WIL) placement based on the student's performance. This work selected Bayes Net, Naive Bayes, J48, REPTREE and Mobile Cart algorithm for the experiment to predict students Gender, sponsors, attendance, subjects and semester. The model performance revealed that Bayes Net and Naive Bayes algorithm resulted in higher accuracy than the other.

Machine Learning

Machine Learning is a technique where the machines are trained in such a way that it gains the ability to respond to a particular input or scenario based on the previous inputs it has learnt. Simply it the giving computers the ability to learn by using statistical techniques. Machine learning helps the computers to act without explicitly being programmed. This aims at reducing the human intervention in the machine dependable problems and scenarios. This helps in solving very complex tasks and problems very easily and without involving much human labor. Various applications of machine learning include classification, prediction, image recognition, medical diagnosis, algorithm building, self-driving cars and much more. Majority of problems in machine learning can be solved using supervised and unsupervised learning. And if the final output classes and sets are not known and it is done by identifying the similarity between data point and their characteristics and finally they are made into groups based on these characteristics then it is called un-supervised.

Machine learning approach was deployed for predicting the employability status of fresh graduate students with three classification algorithms such as Random Forest (RF), Decision Trees (DT), and Support vector machine (SVM). In other to obtain the best model, the three classifiers were measured through the performance matrix f1-score, recall and precision measures. The experiment showed that support vector machine obtained 91.22% accuracy which is better than DT with 85% and RF with 84% accuracy (Casuat, C.D., & Festijo, E.D. 2019)

The researcher (Hugo, L. 2018) developed five machine learning models with Logistic Regression, Discriminant Analysis, Decision Tree, Artificial Neural Network and Support Vector Machine to predict the employment of undergraduate students at graduation. SVM outperformed other models with 87.26% accuracy, followed by artificial neural network with 73.11% and logistic regression with accuracy of 72.17%. Decision Tree and Discriminant analysis had least performance accuracy of 71.70% and 69.81% accuracy.

Materials and Methods

The main aim of the proposed research methodology is to apply five classification algorithms namely Logistic Regression, Naïve Bayes, SVM, Random Forest and Decision Tree on the data set, compute performance evaluation on them and build a suitable, effective and efficient prediction model on the best algorithm that correctly classifies applicants employability and assessment of the attributes of the applicants dataset to meet the selection criteria of work demanded by the industry. The model methodology process flow represented in Figure 1 will be followed.

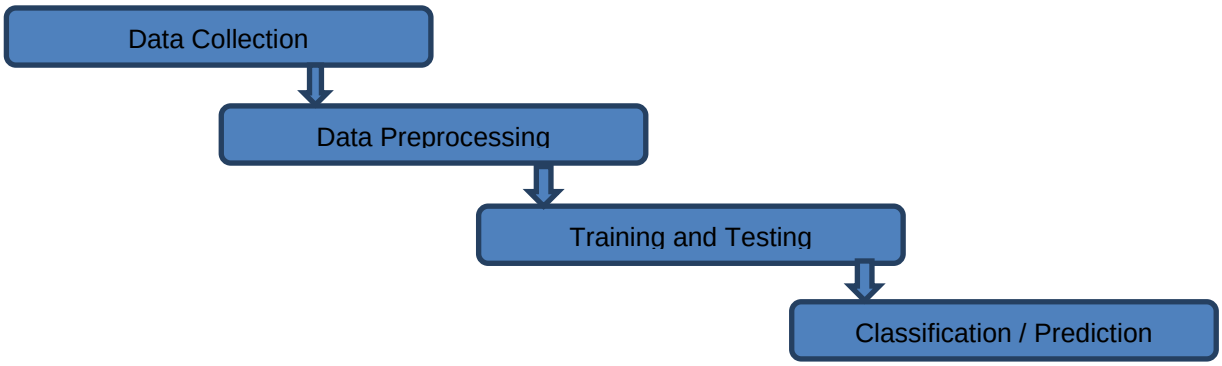

Figure 1: Methodology process flow

Proposed Model
The work propose to build an efficient model to predict personnel or applicant employment status among graduate students of the tertiary institution with five classifiers such as logistic Regression, I Bayes, Decision Tree and Support Vector Machine and Random Forest.

Naïve Bayes
A I Bayes classifier is an algorithm built on adopting simple probability Bayes theorem that requires only very small amount of training data to estimate the mean and variance of variables. It is built on Bayesian classification methods. Where the Bayes's theorem equation represent the relationship of conditional probabilities of statistical quantities.

Decision Tree
Decision Tree is a classifier that looks like tree in graphical representation and the model is used to predict a target attribute value constructed on the number of features entry from example set by creating a classification model.

Logistic Regression
Logistic regression analysis algorithm is mostly used predict or classify the conclusive consequence of one or more of the predictor variables and compute the relationship between a categorical variable with the dependent variable to change the probability score.

Support Vector Machine
The standard support vector machine (SVM) extract sets of input data and predicts or classify, for each of the given input in which one of the two possible classes comprises of the input, making the SVM a non-probabilistic binary linear classifier.

Random Forest
Random forest is a supervised learning classifier that is mostly used for both regression and classification. However, it is majorly used for prediction / classification problems. Random forest consists of trees and many trees can be refer to as a robust forest. It also creates

decision trees on data samples and then output the prediction from each of them before finally selects the best solution by means of voting.

All the five models will be tested and evaluated, therefore model with the highest accuracy will be implemented for employment status prediction among job applicants. The model will be developed and design using python programming language.

Experimental Setting and Evaluation Measure

The classification algorithms Is trained on 80% of the applicants in the dataset while trained model will be used to predict the applicant's selection status on the remaining 20% of the applicants during the test stage of the experiment. Then validate the prediction with the ground truth. Five selected machine learning classification algorithms will be consider for the experiment and the evaluated result of the prediction model will be measured on confusion matrix, percentage accuracy score, recall score and precision score.

Preprocessing Stage

For better performance, the dataset is preprocessed. Null values were thoroughly checked. It was found out that only salary column contain null value and it was dropped. However, Feature selection technique is also adopted to remove all the un-important features, in order to improve the prediction accuracy. All the categorical features were encoded accordingly by ensuring that they were all converted to numeric values. The dataset is then split into 80% train set and 20% test set. It was found out that the dataset is highly imbalanced, then resample techniques was also applied.

Datasets Features

The dataset features and their meaning are listed below:-

1. sl_no : Serial Number.
2. gender : Sex of the candidate (Male or Female).
3. ssc_p : Class 10^{th} percentage.
4. ssc_b : Class 10^{th} board of education.
5. hsc_p : Class 12^{th} percentage.
6. hsc_b : Class 12^{th} board of education.
7. hsc_s : Field of study in Class 12^{th}.
8. degree_p : Degree percentage.
9. degree_t : Undergrad field of education.
10. workex : Years of work experience.
11. etest_p : Employability test percentage.
12. 72pecialization : Field of study during Post Graduation.
13. mba_p : MBA percentage.
14. status : placed/not placed.

Results and Discussion

The experiment is divided into several stages in other to improve the prediction accuracy and to ease the result analysis. However, models accuracy performance evaluation results were considered with and without classifiers parameter tuning.

Data Imbalance

After the splitting of the datasets into 80% (172) training and 20% (43) test set, it was discovered that the data is highly imbalance. This can affect the performance accuracy of the

classifiers. Therefore resample technique is adopted. Figure 2 shows the graphical representation of the target class image before resampling while Figure 3 represents the target class after resampling technique.

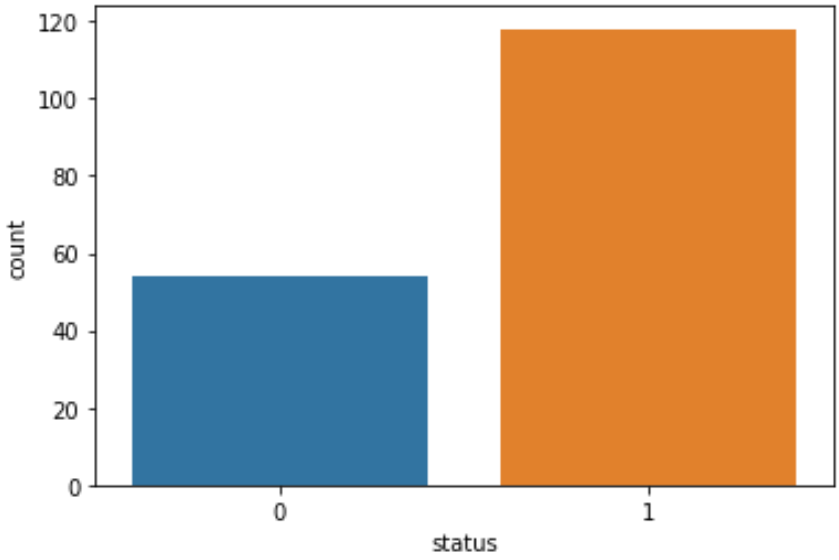

Figure 2: Imbalance target class

In Figure 2, the value 0 represent the number of applicants not-placed while 1 represent numbers of placed applicant. The total number of placed applicants are 69% (118) while those not place are 31% (54). Therefore, the class in not balance. After sampling, placed class target which is the majority class was dropped off by 11 to remain 108 while the minority class which is the not-place remain 54.

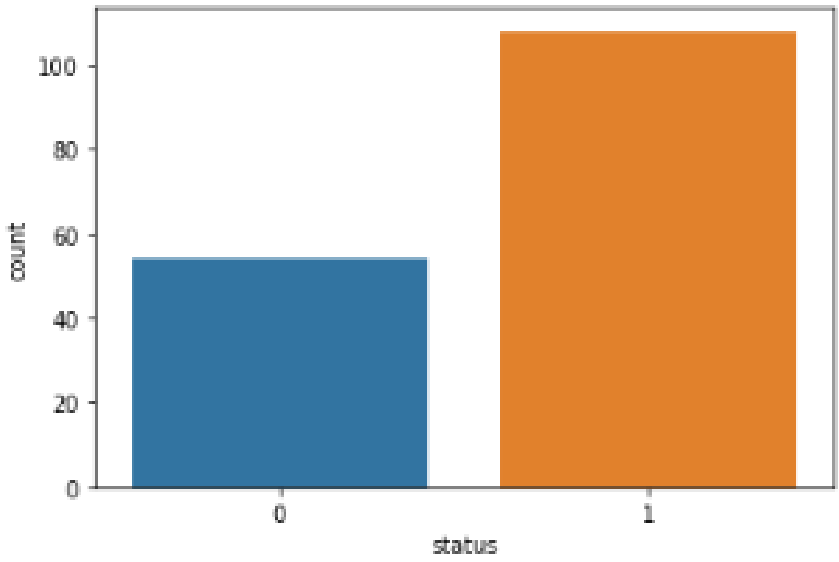

Figure 3: Target class after resample

Feature Reduction and Correlation

At this stage data visualization is required in order to choose the best features to train the model on, since there are so many columns. This will also help to free the system memory for quick computation. Figure 4 and Figure 5 shows the comparison difference of the features and correlation image of the features respectively.

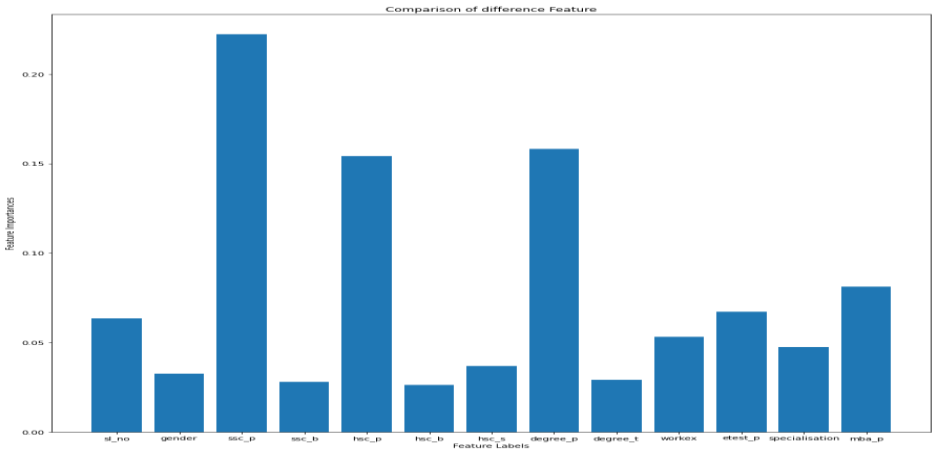

Figure 4: Feature comparison difference

As shown in Figure 4, serial number (sl_no) of the candidate in the real sense is not supposed to have effect on the selection of the applicants or not, but the model could find some pattern in the feature. Also, ssc_p, hsc_p and degree_p (Class 10 percentage) are the most important features. The following features ssc_p, hsc_p, and degree_p have very high positive correlation with each other as shown in figure 5. Therefore they were considered during the classification experiment.

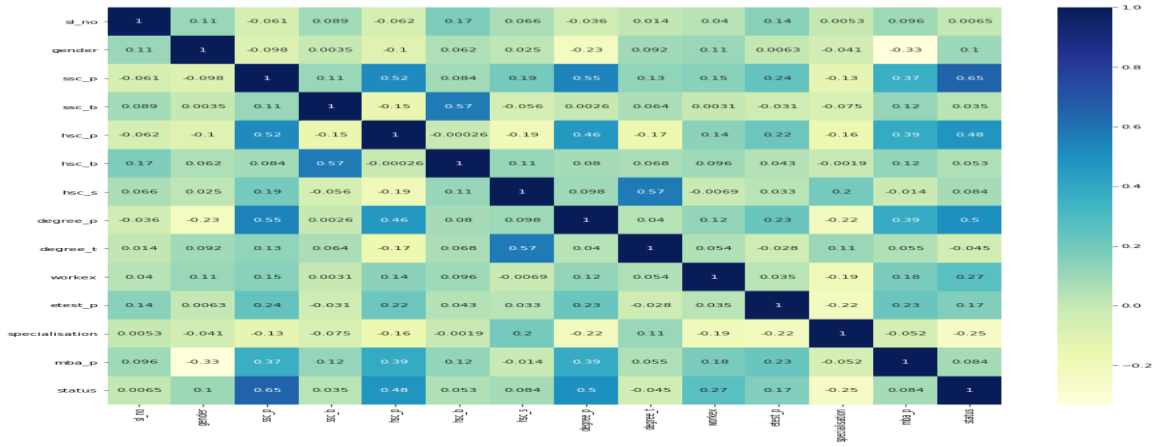

Figure 5: Features correlation

Prediction Results without Parameter Tuning

The experiment was carried out with the aim of comparing the performance of the classifiers. Therefore, algorithm with the best predictive model will be adopted for the implementation of job employment status prediction. The performance result of all the considered classifiers are shown in table1 while their pictorial performance were displayed in Figure 7 and Figure 8.

Table 1: Classifiers accuracy results without Parameter tuning

Classifiers	Accuracy	Precision	Recall	F1 Score
Logistic Regression	0.91	0.93	0.93	0.93
Decision Tree	0.88	0.9	0.93	0.92
Random Forest	0.91	0.9	0.96	0.93
NaiveBayes	0.88	0.87	0.96	0.91
Support Vector Machine	0.93	0.93	0.97	0.95

Support vector machine classifier outperformed others with 93% accuracy, followed by Random forest and logistic regression classifiers with 91% accuracy as shown in Table 1. However, NaiveBayes and Decision tree also had same accuracy of 88%. Obviously, SVM model is hereby recommended for the implementation of Job employment status prediction.

Figure 6 clearly show the pictorial representation of the performance accuracy of the classifiers used in the experiment. Confusion matrix in Figure 7 showed the numbers of True positive rate, true negative rate, false positive rate and false negative rate as it applied to each of the model performance.

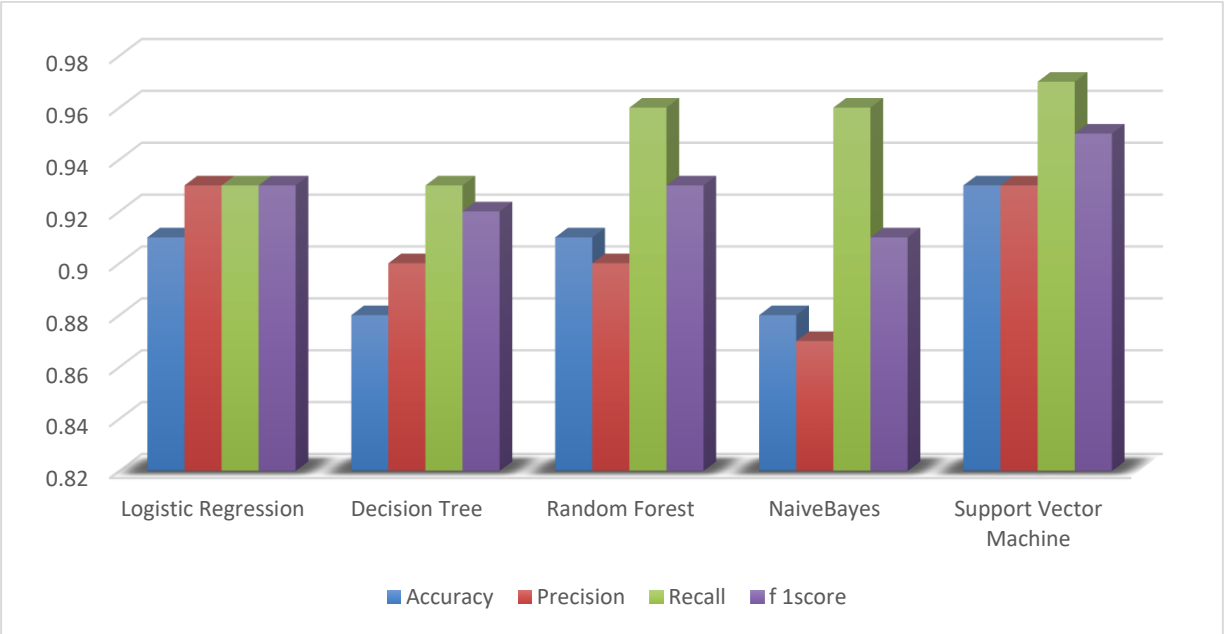

Figure 6: Classifier performance accuracy

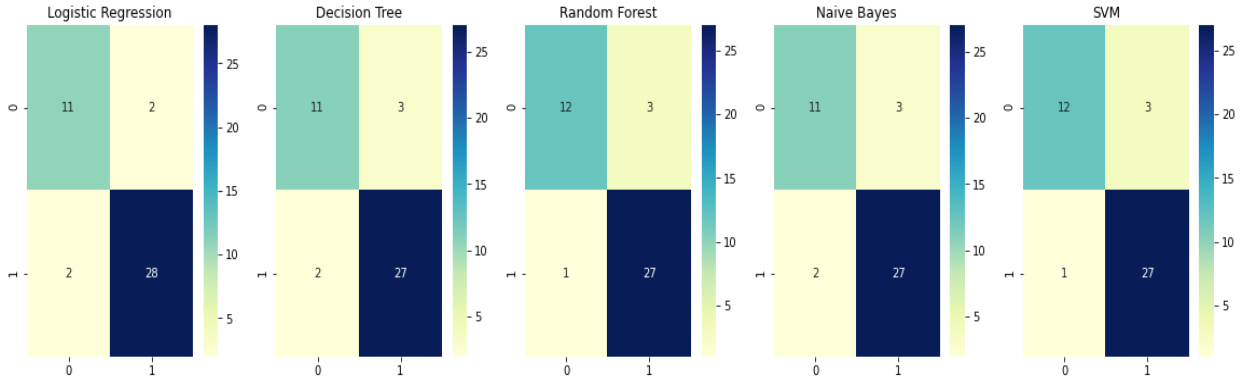

Figure 7: Classifiers model performance confusion matrix

Prediction Results with Parameter Tuning

The experiment was taking further by fine tuning the classifiers parameter to their best state and evaluate the models results. The performance results of the classifiers with parameter tuning are shown in Table 2 while their graphical representation were displayed in Figure 9 and Figure 10.

Table 2: Experiment with parameter tuning

Classifiers	Accuracy	Precision	Recall	F1 Score
Logistic Regression	0.91	0.93	0.93	0.93
Decision Tree	0.84	0.87	0.9	0.88
Random Forest	0.93	0.93	0.97	0.95
NaiveBayes	0.91	0.93	0.93	0.93
Support Vector Machine	0.93	0.93	0.97	0.95

Table 2 revealed that Logistic Regression, Decision tree and SVM showed no improvement after tuning them as they were already in their best with default parameter values. But could see some significant improvement in the Random forest and Naïve Bayes. Therefore Random forest and SVM model can be used for implementing the job employment status prediction.

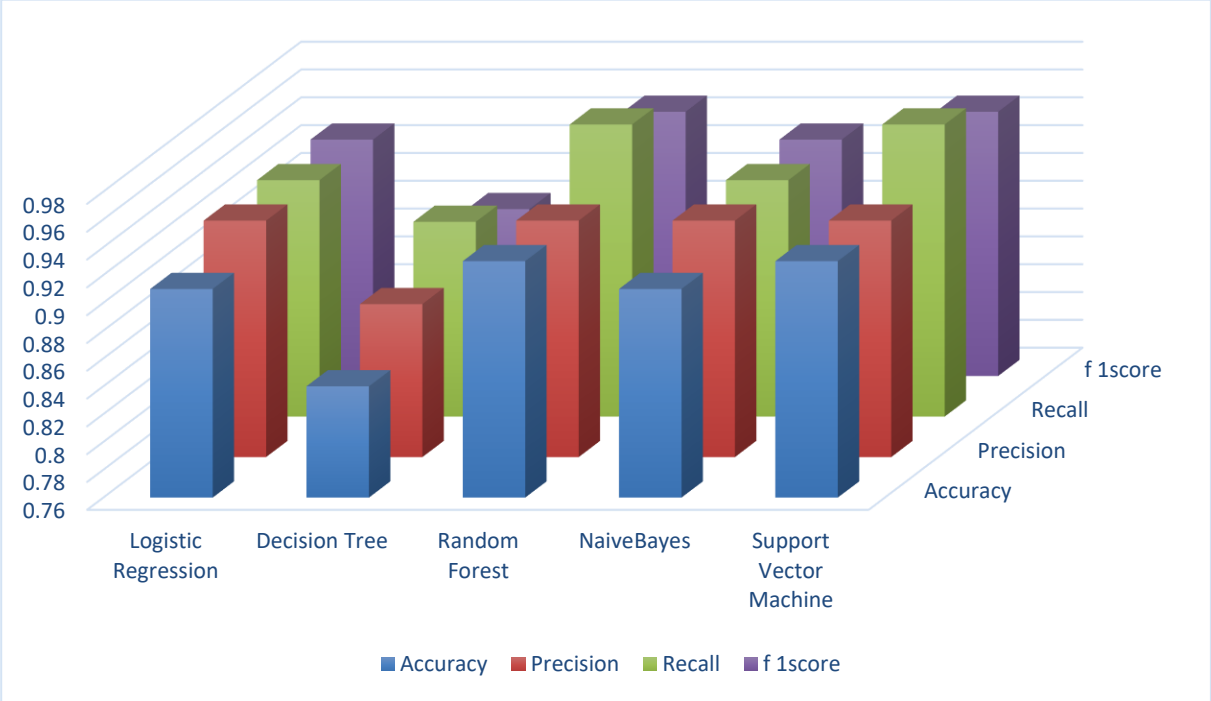

Figure 7: Classifier performance accuracy with parameter tuning

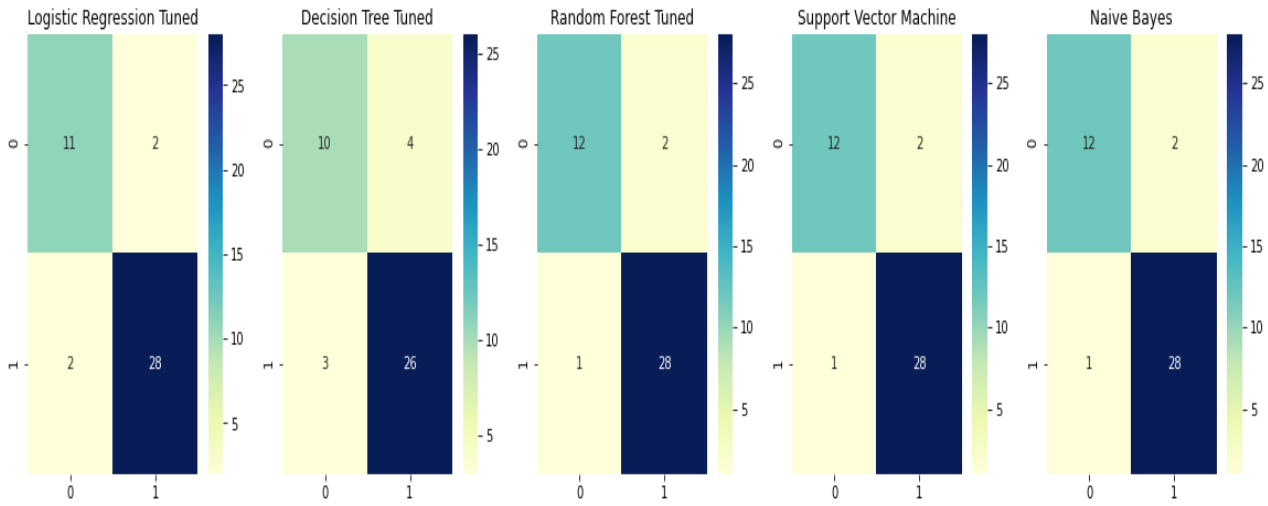

Figure 8: Classifiers results accuracy with parameter tuning in confusion matrix form

Table 3: Summary of the performance evaluation result with and without parameter tuning

	Result Accuracy without Tuning				**Result Accuracy with Tuning**			
Classifiers	Accuracy %	Precision	Recall	F1 Score	Accuracy %	Precision	Recall	F1 Score
Logistic Regression	0.91	0.93	0.93	0.93	0.91	0.93	0.97	0.97
Decision Tree	0.88	0.9	0.93	0.92	0.84	0.87	0.9	0.88
Random Forest	0.91	0.9	0.96	0.93	0.93	0.93	0.97	0.95
NaiveBayes	0.88	0.87	0.96	0.91	0.91	0.93	0.93	0.93
Support Vector Machine	0.93	0.93	0.97	0.95	0.93	0.93	0.97	0.95

Table 4: Result performance comparison

	Authors	**Classifiers**	**Results**
1	Casuat,C.D., & Festijo, E.D. (2019)	Random Forest (RF	84%
		Decision Trees (DT)	85%
		Support vector machine (SVM)	91.22%
2	KianLam et al., (2019)	Logistic Regression	58.50%
		Decision Tree	58.84%
		Naïve Bayes	57.66%
		k-Nearest Neighbor	57.40%
		Neural Network	59.07%
		Support Vector Machine (SVM)	59.01%
3	Pothuganti M & Neelam S (2019)	Random forest	86%
		Decision Tree	84%
4	Hugo, L. (2018)	Logistic Regression,	72.17%
		Discriminant Analysis	69.81%
		Decision Tree	71.70%
		Artificial Neural Network	73.11%
		Support Vector Machine	87.26%
5	Othman et al., (2018)	Support Vector Machines	66.0967
		Artificial Neural Networks	65.2937
		Decision Tree	66.0651%
6	**Proposed**	Logistic Regression	91%
		Decision Tree	88%
		Random Forest	91% and 93%
		NaiveBayes	88%
		Support Vector Machine	**93%**

Conclusion

This work proposed a machine learning model for job employment status prediction for human resource management unit of an organization to aid recruitment process. Five different models were developed and performance evaluation was carried out to determine the best model. However, Support vector machine classification model outperformed other selected classifiers with 93% accuracy but when the models parameters were optimized, Random forest accuracy was improved from 91% to 93%. Also Naïve Bayes improved from 88% to 91%.

Though, machine learning algorithms differs in terms of performance based on datasets and datasets preprocessing techniques. The technique adopted in this work will however contribute significantly in the placement process, especially for the HR placement team to analyze applicants' performance before selection. Furthermore, it will assist the higher institutions to have quality placement records of their graduates and at the same time moderate their curriculum and improve on their teaching and learning method.

Future work on employment status prediction is recommended, by adding more relevant features and increase the number of instances for better result. Also Neural network model is recommended to be used for the prediction.

References

Aziz, M.T., & Yusof, Y. (2016). Graduates employment classification using data mining approach. *AIP Conference Proceedings, 1761*: 020002.

Aziz, A. A., Ismail, N. H., Ahmad, F., & Hassan, H. (2015). A framework for students academic performance analysis using naive bayes classifier. *Jurnal Teknologi*, 75: 13–19.

Casuat, C.D., & Festijo, E.D. (2019). Predicting Students' Employability using Machine Learning Approach. *IEEE 6th International Conference on Engineering Technologies and Applied Sciences (ICETAS),* 1-5.

Hugo, L. (2018). Predicting Employment through Machine Learning. *NACE Journal 2019*. NACE National Association of Colleges and Employers, Bathelehem, PA.

Jantawan, B. &. (2013). The Application of Data Mining to Build Classification Model for Predicting Graduate Employment. *International Journal of Computer Science and Information Security*, 11(10), 1–8.

Khaef, E., Ahmad, A., Mottaghi, P., & MV, S. (2007). Examining the influence of employment model utilization based on data mining, on the employees replacement rate. *First Iran Data Mining Conference*. 11: 1-10.

Masethe, M. A., & Masethe, H. D. (2014). Prediction of Work Integrated Learning Placement Using Data Mining Algorithms. *International Journal of Innovative Technology and Exploring Engineering*, 8: 22–24.

Mishra, T. (2016). Students ' Employability Prediction Model through Data Mining. 11(4), 2275–2282.

Nor Azziaty, A. R., Kian Lam, T., & Chen Kim, L. (2017). Predictive Analysis and Data Mining among the Employment of Fresh Graduate Students in HEI. *AIP Conference Proceedings 1891*, 020007 (2017) (pp. 020007-1 - 020007-6).

Othman, Z., Shan, S.W., Yusoff, I., & Kee, C. (2018). Classification techniques for predicting graduate employability. *International Journal on Advanced Science, Engineering and Information Technology*, 8, 1712-1720.

Pothuganti M & Neelam S (2019). *International Journal of Applied Engineering Research* 14: 2188-2191.

Stone, D. L., Deadrick, D. L., Lukaszewski, K. M., & Johnson, R. (2015). The influence of technology on the future of human resource management. *Human Resource Management Review*. 25(2), 216-231.

Sapaat, M. A., Mustapha, A., Ahmad, J., & Chamili, K. (2011). A Data Mining Approach to Construct Graduates Employability Model in Malaysia. *Journal on New Computer Architectures and Their Applications 1:* 1086–1098.

Shahiri, A. M., Husain, W., & Rashid, N. A. (2015). A Review on Predicting Student's Performance Using Data Mining Techniques. *Procedia Computer Science*, 72: 414–422.

Tajul, M., Ab, R., & Yusof, Y. (2016). Graduates Employment Classification using Data Mining Approach. *AIP Conference Proceedings, 1761*: 020002.

Xu, W., Li, Z., Cheng, C., Tingting, Z (2013). Data mining for unemployment rate prediction using search engine query data. *Service Oriented Computing and Applications,* 7: 33–42.

Zielinksi, D. (2020). Recruiting Gets Smart Thanks to Artificial Intelligence. Retrieved from https://www.shrm.org/resourcesandtools/hrtopics/technology/pages/recruiting-gets-smartthanks-to-artificial-intelligence.aspx,

How to cite this paper: Awujoola J Olalekan, Philip O Odion, Martins E Irhebhude**,** Halima Aminu (2021). Performance Evaluation of Machine Learning Predictive Analytical Model for Determining the Job Applicants Employment Status . *Malaysian Journal of Applied Sciences, 6*(1), 67-79.

View publication stats