https://afrjmis.net

Performance Evaluation of C4.5, Random Forest and Naive Bayes Classifiers in Employee Performance and Promotion Prediction

Daniel E. Asuquo[1], Uduak A. Umoh[2], Francis B. Osang[3] & Edikan W. Okokon[4]

[1, 2, 4]Department of Computer Science, Faculty of Science, University of Uyo, Nigeria
Email: danielasuquo@uniuyo.edu.ng, uduakumoh@uniuyo.edu.ng

[3]Department of Computer Science, Faculty of Sciences, National Open University of Nigeria, Abuja
Email: bukie3osang@yahoo.com, fosang@noun.edu.ng

ABSTRACT

Conventional methods of staff performance characterization associated with large volumes of employees' dataset, nepotism, bias, human preference, and subjectivity, has made it difficult to obtain accurate, timely, and cost-effective evaluation of staff performance. These cause predicting their chances of getting promoted cumbersome, thereby impeding workplace productivity, innovation, and effective reward system. This paper develops a framework that evaluates the performance of C4.5, Random Forest and Naïve Bayes classifiers in mining employees' data, from a secondary dataset of 110 employees between 2018-2020, of Royalty Hotel in Eket, Akwa Ibom State, Nigeria to obtain hidden patterns and interactions between attributes for informed decision making. WEKA toolkit was used during each model's training, testing and prediction. To improve competency and job retention, staff predicted as 'not promoted' were recommended by the system for professional training and development support or for normal salary increment. Results from the confusion matrix and performance metrics indicate that random forest has the best generalization capability, having outperformed the other two models in staff performance and promotion prediction, with highest prediction accuracy and F-measure value of 98.70% and 0.988, respectively. The receiver operating characteristic curve plot also shows that it yields the highest probability of prediction in the case of true positive than true negative and achieves a good measure of separability. This model can be used to predict the target class for unseen data with acceptable accuracy.

Keywords: *Staff Performance Indicators, Promotion Prediction, Classifiers, C4.5, Random Forest, Naïve Bayes*

Reference Format:

Asuquo, Daniel; Umoh, Uduak A.; Okokon, Edikan W. and Osang, Francis B. (2020), Performance Evaluation of C4.5, Random Forest and Naïve Bayes Classifiers in Employee Performance and Promotion Prediction, *Afr. J. MIS*, Vol. 2, Issue 4, pp. 41 - 55 .

1. INTRODUCTION

With the rise in market competition and unpredictable business environment, many organizations attempt to attain acceptable standards by improving their staff performance in order to enhance overall workplace productivity. Human resource (HR) managers are challenged with engraved policies, procedures and strategic plans that ensure the right people are hired, promoted and trained for the right job at the right time in order to increase organization productivity and throughputs. With routine employee performance evaluation, management can discover relevant training and development needs of staff including promotion to the right position to meet set organization's goals. Evaluation of the employee with high performance on capability, knowledge, skill, behaviour, and other abilities plays significant roles in the success of an organization (Aksakala, *et al.*, 2013). However, most organizations are faced with the challenge of providing quality service delivery, innovation, productivity; and recommending staff for promotion using HR managerial process. The reason is that most HR approaches are costly, time-consuming, and sometimes difficult to achieve desired outcome because the appraisal process is often flawed with bias, nepotism and god-fatherism. Traditional methods for evaluating employee's performance for promotion such as critical incident, narrative essays, graphic rating scales, ranking (Kateřina *et al.*, 2013) and the so-called modern methods like 360 degree and management by objectives are not devoid of uncertainty and subjectivity due to their dependent on conflicting factors, human experience, preferences and judgments (Aggarwal & Thakur, 2013).

With recent advancement in computing technology, performance classification in HR management can be implemented more effectively through data mining and knowledge discovery in database approach. Databases are rich with hidden knowledge that can be used in decision making process in order to produce intelligent decision. Clustering, classification and prediction techniques are well known machine learning (ML) methods that support intelligent decision making (Pandey & Sharma, 2013; Inyang *et al.*, 2019), after discovering hidden information, patterns and relationships from large amount of data. While clustering uses algorithms to group similar kinds of elements, classification which is a supervised ML technique learns a target function (classification model) that maps each attribute set to one of the predefined class labels (Jantan *et al.*, 2010a). The task of classification has diverse applications and according to Nasr *et al.* (2019), three phases are normally considered during the process. The first phase is the learning process where the training data are analyzed by the selected classification algorithm. The learned model or classifier is obtained as classification rules or patterns. In the second phase, the model is evaluated to estimate and validate the accuracy of the classifier through testing dataset while in the third phase, if the estimated accuracy is considered acceptable, the rules can be applied to new unseen or untrained data to give prediction about unknown class label.

Figure 1 shows the general steps taken during the prediction process.

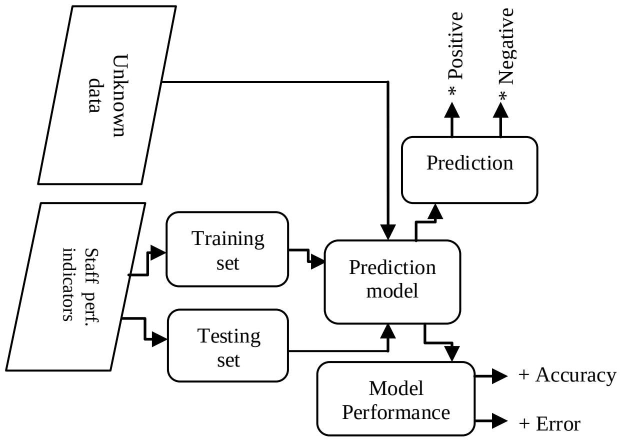


Figure 1: General Step of the Prediction Process

The importance of professional training and its effect on employees' performance has recently aroused interest among researchers. Jantan *et al.* (2014) developed a classification model for achieving academic performance with sequential minimal optimization (SMO) and support vector machine algorithms. Their results indicate acceptable prediction accuracy but require further enhancement. The study was not on staff performance evaluation and could not make recommendations. Kirimi & Moturi (2016) built a model for predicting employees' performance from previous appraisal records using only decision tree algorithm. Tupe *et al.* (2017) identified the most relevant attribute with highest information gain for predicting employees' performance using ID3 algorithm. Improvement in performance and quality of the management system was still required. Lamarca & Ambat (2018) designed a framework for developing a performance appraisal prototype using decision tree (J48) and fuzzy logic techniques. J48 generated IF-THEN rules in conjunction with fuzzy logic controller to predict individual or institutional faculty performance. The framework did not make recommendations. Nasr *et al.* (2019) proposed the use of decision tree, Naïve Bayes, and support vector machine to build a classification model for predicting employees' performance using a dataset collected from the Ministry of Egyptian Civil Aviation (MOCA). They seek to identify the most influential factors that positively impact on staff performance. To get a highly accurate model, several experiments were performed in WEKA tool to enable decision makers and HR professionals predict and enhance the performance of their employees. From the reviewed literature, most studies could not make recommendations to address workforce development challenges and results obtained from classification model require enhancements.

Consequently, this work develops a data mining classification framework as an efficient analytical tool for classifying and predicting employees' work performance in order to help HR managers decide whether or not a staff should be promoted. The framework accepts staff performance indicators and compares the performance of decision tree (C4.5), random forest and Naïve Bayes for mining employees' data in order to discover useful information for predicting their chances of getting promotion. The result from C4.5 algorithm generates a rule-set with a pruned decision tree, useful for predicting employees' future performance. The goal is to achieve acceptable classification outcome with minimal number of decisions, thereby saving cost and time. Furthermore, specific recommendations for staff training and development supports are made for employees predicted as 'not promoted'.

The rest of the paper is structured as follows. Section two reviews related works on classification models for staff performance evaluation and in other application areas. In section three, a flow diagram illustrating the major tasks performed in this study is presented. This is followed by the proposed framework for evaluating the performance accuracy of C4.5, random forest and Naïve Bayes models as well as choosing the optimal classifier for staff promotion prediction and recommendation. Section four discusses the results obtained from experiments in WEKA tool while section five concludes the paper with direction for future works.

2. RELATED WORKS

A. Staff performnace appraisal

HR managers in the past normally choose from among a number of traditional appraisal methods to classify employees. Numerous methods exist (Kateřina *et al*., 2013; Aggarwal & Thakur, 2013; Shaout & Yousif, 2014) and the type of performance appraisal system used depends on its purpose and organization's goal. While the traditional methods like rating scales seem appropriate when the major emphasis is on selecting people for promotion and merit pay increases, the collaborative methods that include input from the employees themselves may prove to be more suitable for employees human capacity development. Notwithstanding, none of these methods are devoid of uncertainty and subjectivity, thereby lacking the needed objectivity in the analysis process for timely and accurate decision making.

According to Duraisingam & Skinner (2005), the core factors often considered in evaluation criteria for staff appraisal include performance factors - knowledge, skill, productivity; behavioural traits factors - dependability, creativity, interpersonal relationship, punctuality and adaptability; results/outcomes - quantifiable results, measurable outcomes and achievements, objective attained; and supervisory factors - leadership, personnel management, planning and organization. Dessler (2011) and Aggarwal & Thakur (2013) explained that five to nine scale points can be used to rate employees by identifying the score that best describes his/her level of performance for each trait. Peter *et al.* (2019) modeled key performance indicators for academic staff advancement in higher institutions of learning using fuzzy logic. A number of attributes considered in the model includes research publications, higher degree certificates, service to the community, administrative role, student supervision role and teaching load. They emphasized that academic staff can be motivated for workplace productivity when rewarded with promotion and other forms of advancement such as special training and administrative role as and when due.

Recently, data mining and ML techniques have been considered as an efficient method for performance classification due to their ability to produce more accurate and easily interpreted results (Jantan *et al.*, 2010a; Karaolis *et al.*, 2010; Neogi *et al.*, 2011; Su *et al.*, 2012; Shang & Barnes, 2012; Lamarca & Ambat, 2018). Sarda *et al.* (2014) built models using decision tree and Naïve Bayes algorithms to rank applicants for a job profile based on their resumé and social media presence. The match making system presents companies with a list of ranked applicants where selection can be made using information retrieval technique like two-way relevance ranking but did not categorize them into classes.

Valle *et al.* (2012) used Naïve Bayes classifier to predict job performance in a call center with the aim of knowing what level of attributes signify individuals who perform well. They predicted future performance of sales agents and achieved satisfactory results with operational records. Ancheta *et al.*, (2012) used rule-based classification technique to extract knowledge significant for predicting training needs of newly-hired faculty members in order to devise the necessary development programs. They used the Cross Industry Standard Process for Data Mining (CRISP-DM) to discover significant models needed for predictive analysis and also revealed required professional trainings needed to prepare faculty members for effective tasks performance. Al-Radaideh & Al-Nagi (2012) developed a classification model for predicting employee's performance based on CRISP and decision tree. Their tool was validated through experiments with real data collected from several companies. In Jantan *et al.* (2010b), potential human talent was predicted where the pattern of talent performance was identified using C4.5 classification algorithm. Hazra & Sanyal (2016) used iterative dichotomizer 3 (ID3) algorithms for recruitment prediction of candidates within an organization. Entropy and information gain were computed to resolve the splitting aspect and establish a pruned decision tree.

Fashoto *et al.* (2018) proposed a framework for dynamic human resource information system (HRIS) to curb the problem of loss of documents, often resulting in promotion delay and career stagnation on the part of staff members. Primary data was collected through questionnaire and analyzed using Statistical Package for Social Sciences in order to determine the readiness of academic staff of Kampala International University to adopt the dynamic HRIS and its implementation. Findings showed that unavailability of Information and Communication Technology (ICT) services, poor ICT skills, and absence of organizational competition are the most significant factors that could militate against effective usage of the HRIS if adopted. Góes & de Oliveira (2020) proposed a process for employee performance evaluation using computational intelligence techniques in order to provide a fair evaluation process and minimize common problems caused by simple objective or subjective approaches. Their work combined fuzzy logic, text sentiment analysis and classification techniques, such as multi layer perceptron (MLP) artificial neural network, decision tree algorithms and Naïve Bayes into ensemble classifiers. Research data was obtained from several evaluations applied in two Brazilian institutions. Results showed consistency on the data generated by the proposed process, indicating its suitability for applications on companies of most business areas.

B. Classifiers for performance evaluation in other fields

Subasi *et al.* (2018) used ensemble ML classifiers based on Adaboost and random forest algorithm to improve the performance of automated human activity recognition. With wireless network technology, intelligent monitoring of daily human activity was implemented through wearable body sensors with wide application in healthcare, terrorist detection, emergency help, cognitive assistance and home monitoring. Baati and Mohsil (2020) suggested a real-time online shopper behaviour prediction system to predict a visitor's shopping intent immediately the website is visited. They used session and visitor information and investigated the performance of Naïve Bayes classifier, C4.5 decision tree and random forest. Results obtained with oversampling showed that random forest produced significantly higher accuracy and F1 Score than the rest. Zhang *et al.* (2020) used random forest as well as classification and regression tree (CART) algorithms to analyze students' physical education information, course exam results, student learning data and relevant feature attributes from the online teaching platform. This was aimed at determining factors that influence students' physical education performance in order to improve teaching quality thereby helping management and teachers improve teaching methods and adjust teaching strategies. The generated decision trees and classification rules significantly identified rules for correlating student learning factors and adjusting some pedagogical strategies.

Ak (2020) applied data visualization and ML techniques including logistic regression, k-nearest neighbors (k-NN), support vector machine, Naïve Bayes, decision tree, random forest, and rotation forest on a dataset for breast cancer detection and diagnosis using R, Minitab, and Python as implementation tools. Results indicate that logistic regression model has the highest classification accuracy (98.1%), and the proposed approach has the potential to improve accuracy performance as well as open new opportunities in the detection of breast cancer. Razali *et al*. (2020) applied Naive Bayes, C4.5 decision tree, k-NN, SMO, random forest, MLP neural network and simple logistic regression to classify a dataset for diagnosis of cervical cancer and determination of its risk factors. Results indicate that random forest yields the best classification accuracy of 96.40% and outperform other algorithms in cervical cancer diagnostic.

C. Description of selected classifiers

Anuradha &Velmurugan (2015) present data mining as the process of discovering significant patterns in large quantities of data. Various data mining methods and algorithms have been applied to discover and extract hidden patterns of stored data (Abdulsalam *et al*. 2014; Shruthi & Chaitra, 2016). The aim is to predict patterns, future trends and behaviors thereby allowing businesses to make effectively proactive, knowledge-driven decisions.

Decision tree classifier is a hierarchical, tree structure that follows a top-down approach. It consists of nodes and directed edges where each node could be a root, internal, or leaf node. The root node is the topmost decision node which corresponds to the best predictor while every internal node has two or more children nodes and contains splits, which test the value of an expression of the attributes (Olaniyi *et al.* 2017). Arcs from an internal node to its children are labeled with distinct outcomes of the test while each leaf node has a class label associated with it. The training set for constructing a decision tree consists of data tuples, described by a set of attributes and a class label. Attributes can have discrete or continuous values. The generated rule set can be used for future prediction of unknown class label of data tuples.

The central problem is how to choose the "best" splitting attribute. Typical among decision tree algorithms useful for prediction, classification as well as detection of interaction between variables includes ID3 algorithm, C4.5 (J48) algorithm, Classified and Regression Trees (CART) algorithm, and CHi-squared Automatic Interaction Detection (CHAID) algorithm. Both ID3, developed by Quinlan (Quinlan, 1987), and C4.5 (J48) employ the entropy measure as their splitting function while CART, developed by Brieman *et al.* (1984), uses the Gini index as its splitting function. CHAID uses the statistical X^2 test to determine the best split during tree-growing process. Other widely used techniques for classification tasks can be found in (Islam *et al.*, 2010; Abdulsalam *et al.*, 2014; Mohammed *et al.*, 2017).

(i) C4.5 algorithm

Ideally, ID3 algorithm uses two concepts - entropy and information gain, to create a tree following a top-down greedy search through the given dataset to test each attribute at every node (Fong & Weber-Jhanke, 2012). The entropy is the degree of randomness of data, used to calculate homogeneity of data attribute. If entropy is zero then sample is totally homogeneous and if one, then sample is completely uncertain. On the other hand, information gain is the decrease in entropy and the attribute with the highest information gain is selected as best splitting attribute.

Although, both ID3 and C4.5 (J48) algorithms employ the entropy measure as the splitting function, the latter uses a new selection criterion called gain ratio Quinlan (1993), which is less biased to generate decision tree with its rule sets. Secondly, C4.5 algorithm can handle numeric attributes, in contrast to ID3 algorithm from which C4.5 evolved. As an enhancement to ID3, it can also handle missing value attributes as well as continuous input attributes without need to discretize them. Lastly, it can avoid over-fitting of decision tree by providing the facility of pre and post pruning. The justification for choice of C4.5 is to achieve acceptable classification with minimal number of errors, since finding optimal tree is an NP-hard problem.

(ii) Random forest classifier

Another method introduced by Breiman (2001) to reduce the danger of over-fitting by constructing an ensemble of trees is called random forests. A random forest classifier consists of a collection of decision trees where each tree is constructed by applying an algorithm, A, on the training set, S, and an additional random vector θ, where θ is an independent and identically distributed sample from some distribution. Random forest prediction is formed by a majority vote over the predictions of the individual trees.

(iii) Naïve Bayes classifier

Naïve Bayes classifier can be utilized for predicting a target class and is based on Bayes conditional probability. The classifier assumes the predictors are statistically independent, which makes it an effective classification tool that is easy to interpret. It is best employed when faced with the problem of 'curse of dimensionality', that is, when the number of predictors is very high.

3. METHODOLOGY

The flow diagram in Figure 2 illustrates the overall methodology followed in this study. The following sections describe the major steps of this methodology.

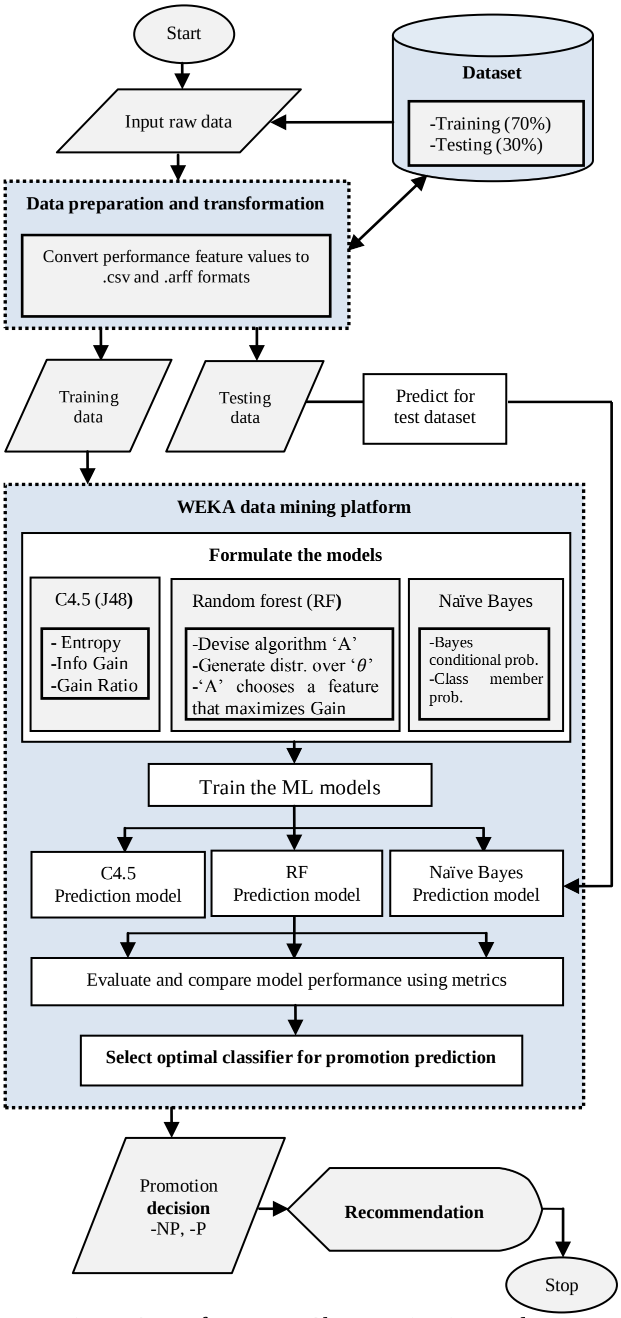


Figure 2: Performance Characterization and Promotion Prediction Approach

A. Data preparation and transformation

A dataset of 110 instances was divided in the ratio of 7:3 for model training and testing, respectively. Ten (10) selected attributes along with their CLASS prediction were extracted from the dataset, exported from Excel application and converted into comma separated (.csv) file format, and then stored as attribute relation file format (.arff) accepted in WEKA software by adding relation, attribute, and data tags (Witten *et al.* 2011). This was used in building the model for employees' promotion prediction. Three classification techniques, namely C4.5, Random Forest and Naïve Bayes, were applied on the dataset and their performance compared to determine the best model for the study and to get the most effective variables that affect and predict the employees' performance.

The 10 variables selected as staff performance predictors were expressed as numeric values between 1 and 10, where 1 indicates high level of dissatisfaction and 10 indicates high level of satisfaction. They are Team Work (TW), Problem Solving (PS), Planning and Organizing Work (POW), Technical Knowledge of Job (TKJ), Customer Service Skills (CSS), Result Oriented (RO), Attendance and Punctuality (AP), Loyalty and Dedication to Work (LDW), Going Extra Miles (GEM) and communication skills (CS). The prediction or target output - CLASS is either P for Promoted or NP for Not Promoted.

B. Formulation of the models

The entropy, information gain, gain ratio and splitting information can be evaluated using equations (1) - (4). We exploit the algorithmic illustrations in Witten *et al.* (2011) and Anuradha and Velmurugan (2015) to demonstrate the formulation of these measures as follows: We select one example at random from a set T of training examples and identify that it belongs to class, $C_i, i = 1, \dots, k$. This has a probability of $\frac{|C_i|}{|T|}$, where $|c_i|$ is the number of examples that belongs to C_i. Then, given a probability distribution, $P\left(p_1 = \right.$ *C1T, p2= C2T, …, pk=CkT* , the information contained in this distribution, called the entropy, is expressed as:

$$I(P) = -\sum_{i=1}^{k} p_i * \log(p_i) \quad (1)$$

When applied to the set $T$of training examples, $I(P)$ measures the average amount of information (bits) needed to identify the class of an example of T. Again, we consider a similar measurement after T has been partitioned in accordance with the n outcomes of a test on the feature X (X can be a test on POW, TW, CS, etc.). The expected information required is obtained as a weighted sum over the subsets as $Info\,(X,T) = \sum_{i=1}^{n} \frac{|T_i|}{|T|} Info(T_i)$, where $T_1, T_2, \dots, T_m$is the partition of T induced by the value of X. The information gain denoted by the quantity, $Gain\,(X,T)$ is defined as:

$$Gain\,(X,T) = Info\,(T) - Info\,(X,T) \quad (2)$$

Equation (2) represents the gain in information due to attribute X, which is the difference between the

https://afrjmis.net

information needed to identify an element of T and the information needed to identify an element of T after the value of attribute X has been obtained. The $GainRatio\,(X,T)$ defined as the proportion of information generated by the split that is useful for the classification is expressed in equation (3), where $SplitInfo(X,T)$is as given in equation (4).

$$GainRatio\,(X,T) = \frac{Gain(X,T)}{SplitInfo\,(X,T)} \qquad (3)$$

$$SplitInfo\,(X,T) = -\sum_{i=1}^{n} \frac{|T_i|}{|T|} \log_2 \frac{|T_i|}{|T|} \qquad (4)$$

Ideally, random forests are a combination of tree predictors where every tree depends on the values of a random vector, θ sampled independently and with the same distribution for all trees in the forest. To specify a particular random forest, we use the detailed explanation in Baati *et al.* (2020) to define the algorithm, 'A' and the distribution over θ. There are many ways to do this but here we generate θ as follows. First, we take a random subsample from S with replacements; namely, we sample a new training set S' of size m' using the uniform distribution over S. Second, we construct a sequence I_1, $I_2, \ldots$, where each I_t is a subset [d] of size k, which is generated by sampling uniformly at random elements from [d]. All these random variables form the vector, θ. Then, the algorithm, A grows a decision tree (e.g., using the ID3 algorithm) based on the sample S', where at each splitting stage of the algorithm, the algorithm is restricted to choosing a feature that maximizes gain from the set I_t. Intuitively, if k is small, this restriction may prevent over-fitting.

Naïve Bayes classifier (V_{NB}) assumes that given the target value of the instance v_i, the probability of observing the conjunction $a_1, a_2 \ldots a_n$ is just the product of the probabilities for the individual attributes (Islam *et al.*, 2010). That is:

$$P(a_1, a_2 \ldots a_n | v_i) = \text{argmax}_{v_i \in V} \prod_i P(a_i|v_i) \qquad (5)$$

$$V_{NB} = \text{argmax}_{v_i \in V} P(v_i) \prod_i P(a_i|v_i) \qquad (6)$$

Basically, this classifier ignores the possible correlation dependencies among the inputs (Ak, 2020) and reduces a multivariate problem to a group of univariate problems. Thus, the number of distinct $P(a_i|v_i)$ terms that must be estimated from the training data is just the number of distinct attribute values times the number of distinct target values.

C. Promotion prediction based on optimal classifier

The proposed data mining framework for staff performance evaluation and promotion prediction has a user interface with a dashboard that enables user input and display of prediction outcome as well as recommendations. The prediction system contains the organization's policy which constitutes the knowledge base and a set of rules for applying the HR experience to each particular situation.

Data used for feature analysis, visualization, pre-processing, classification, and prediction were taken from a secondary dataset of 110 instances between 2018-2020, of Royalty Hotel employees in Eket, Akwa Ibom State, Nigeria. The prediction accuracy of C4.5 (J48), random forest and Naïve Bayes classifiers was evaluated based on the framework using Waikato Environment for Knowledge Analysis (WEKA) toolkit. WEKA is a collection of ML algorithms for DM tasks and the algorithms can either be applied directly to a dataset or invoked from a Java code. It provides a unified package at only one application, which enables users to access the modern updated technologies in DM and ML environment. It contains several tasks such as pre-processing, classification, clustering, association, attributes selection and visualization. The tools supported by the WEKA workbench are based on statistical evaluations of the models (algorithms). Accordingly, the user can easily make comparisons among the prediction accuracy results of the applied DM algorithms for a given dataset in order to detect the most suitable algorithm for the dataset.

4. RESULTS AND DISCUSSION

A. Feature Evaluation

Using WEKA toolkit, 70% of the dataset was used for model training, while the remainder was for testing. Table 1 presents the sample dataset used for the study. A 10-fold cross validation was applied on each of the classification models and the visual interpretation of complex relationships in the dataset is presented. After data preprocessing in WEKA, a view of all attributes and class prediction indicates that 41 instances belong to CLASS 'promoted' (P) while 36 belong to 'not promoted' (NP), as shown in Figure 3. It further shows that all staff in the promoted class had POW values greater than 6. The WEKA-generated outputs for information gain and gain ratio of each predicting attribute with respect to class are shown in Figures 4-5 and tabulated in Table 2.

Highly ranked attributes have bigger values. In characterizing staff performance indicators, Table 2 further shows that POW attribute has the highest gain value, followed by TW.

Table 1: Sample Dataset of Attributes and Class for Promotion Prediction

TW	PS	POW	TKJ	CSS	RO	AP	LDW	GEM	CS	CLASS
5	6	6	7	6	7	8	8	7	7	NP
5	4	5	6	7	8	8	8	6	7	NP
6	7	8	9	8	7	7	6	8	8	P
7	7	6	7	7	8	8	7	7	5	NP
3	4	2	5	5	3	4	5	6	7	NP
4	5	5	3	2	1	5	6	7	6	NP
9	7	8	8	9	9	6	6	8	8	P
7	9	9	8	9	7	7	8	9	9	P
5	10	8	9	8	7	10	10	7	8	P
5	6	7	6	5	4	4	3	6	6	NP
9	4	7	8	8	7	9	9	5	5	P

Rank	Attribute	Information Gain	Gain Ratio
1	POW	0.6764	0.431
2	TW	0.3638	0.418
3	CS	0.257	0.285
4	PS	0.2361	0.271
5	CSS	0.1741	0.175
6	LDW	0.1634	0.17
7	AP	0.1331	0.152 (GEM)
8	GEM	0.1322	0.136 (AP)
9	TKJ	0.1182	0.122
10	RO	0.0993	0.119

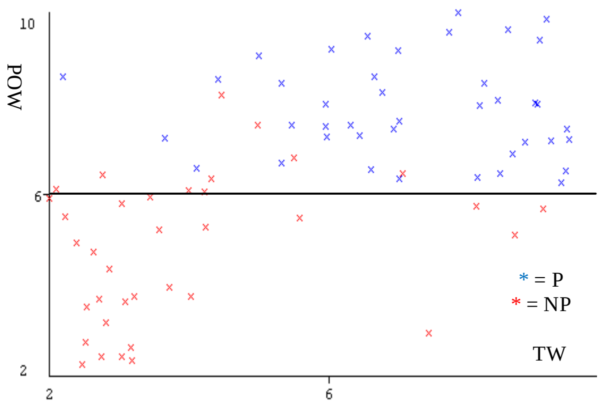
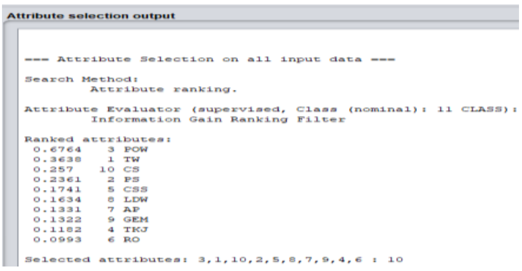


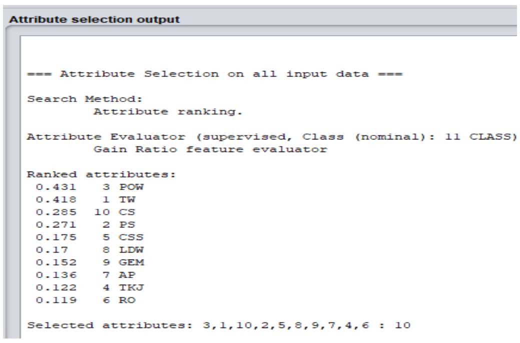
Figure 3: Proportion of Prediction Classes



Attribute selection output

```
=== Attribute Selection on all input data ===

Search Method:
        Attribute ranking.

Attribute Evaluator (supervised, Class (nominal): 11 CLASS):
        Information Gain Ranking Filter

Ranked attributes:
 0.6764    3 POW
 0.3638    1 TW
 0.257    10 CS
 0.2361    2 PS
 0.1741    5 CSS
 0.1634    8 LDW
 0.1331    7 AP
 0.1322    9 GEM
 0.1182    4 TKJ
 0.0993    6 RO

Selected attributes: 3,1,10,2,5,8,7,9,4,6 : 10
```

Figure 4: WEKA-generated Information Gain



Attribute selection output

```
=== Attribute Selection on all input data ===

Search Method:
        Attribute ranking.

Attribute Evaluator (supervised, Class (nominal): 11 CLASS)
        Gain Ratio feature evaluator

Ranked attributes:
 0.431    3 POW
 0.418    1 TW
 0.285   10 CS
 0.271    2 PS
 0.175    5 CSS
 0.17     8 LDW
 0.152    9 GEM
 0.136    7 AP
 0.122    4 TKJ
 0.119    6 RO

Selected attributes: 3,1,10,2,5,8,9,7,4,6 : 10
```

Figure 5: WEKA-generated Gain Ratio

Table 2: Attributes Information Gain and Gain Ratio

B. Classification rules

The classifier model is a pruned decision tree in textual form that was produced on the full training data. IF-THEN rules were used for the classification rule set and the pruning technique executed by removing nodes with less desired number of objects. This resulted in a tree of size 9 with 5 leaves. The first split is on the 'POW' attribute while at the second level, the splits are on 'LDW' and 'TW'. Table 3 shows the classification rules generated from the C4.5 algorithm while Figure 6 shows the visualized pruned tree.

Table 3: Classification Rules Generated by C4.5

The generated rule set indicates that a staff rated between 1 and 6 in POW has poor performance and cannot be promoted. It further revealed that a staff that is rated 7 or less in POW and 5 or less in TW also has poor performance and cannot be promoted.

On the other hand, any staff rated 7 or more in POW and 6 or more in LDW is said to have good performance and recommended for promotion. Furthermore, any staff rated 7 or less in POW and 5 or more in TW is recommended for promotion. Finally, results indicate that any staff rated 7 or more in POW is said to have good performance and is automatically recommended for promotion. This result can be implemented as an application, devoid of nepotism or bias and can be deployed for timely and accurate evaluation of staff performance for promotion, reward, recommendation and training purposes by the HR unit of the establishment. Figure 7 shows the promotion prediction and recommendation system where staff who are not due for promotion to higher ranks could be recommended for a training to acquire more skills so as to improve job competency. In visualizing classifier errors, a plot of predicted class (y-axis) vs. actual class (x-axis) as shown in Figure 8 indicates that 5 instances of class 'NP' has been incorrectly assigned to class 'P' by Naïve Bayes while 8 instances of class 'P' has been incorrectly assigned to class 'NP'. C4.5 and random forest algorithms were observed to have no error in classification.

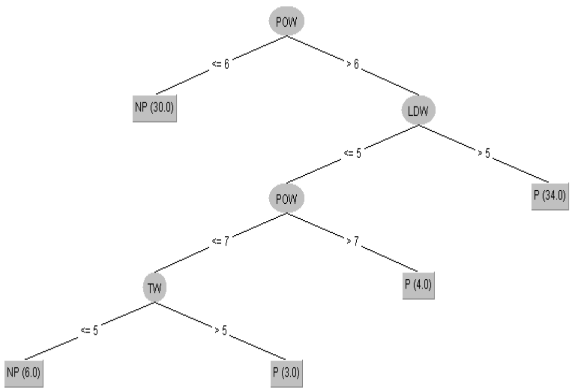


Figure 6: Pruned decision tree generated by C4.5

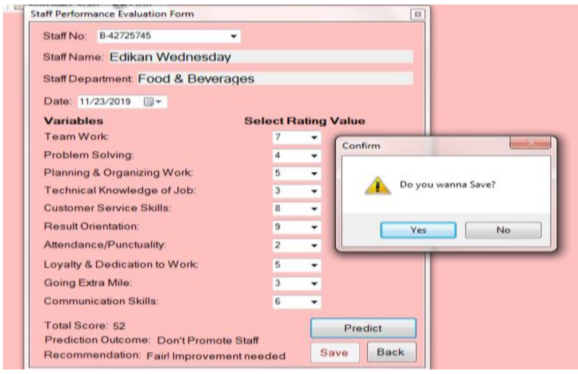


Figure 7: Promotion prediction outcome and recommendation

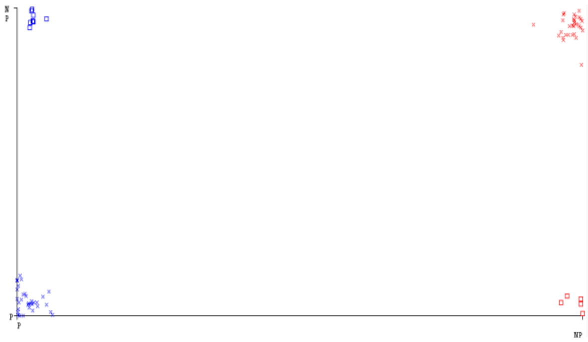


Figure 8: Naïve Bayes classifier errors

C. Performnace Evaluation of Selected Models

Each classifier in a systematic way, generated classification models of the input dataset, by employing a learning algorithm to identify the model that best fits the relationship between the attributes and the class label. One key objective the learning algorithm ought to achieve is to build models with

Rule #	Rule Antecedent	Promotion Decision	Number of Instances	Percent (%)
1	IF (POW <= 6) THEN	NP	30	38.96
2	IF (POW > 6) AND (LDW > 5) THEN	P	34	44.16
3	IF (LDW <= 5) AND (POW > 7) THEN	P	4	5.19
4	IF (POW <= 7) AND (TW > 5) THEN	P	3	3.90
5	IF (POW <= 7) AND (TW <= 5) THEN	NP	6	7.79
	TOTAL		77	100

good generalization capability to accurately predict the class labels of previous unknown records.

Model performance evaluation based on counts of test records correctly and incorrectly predicted are presented, in Figures 9-11, as binary classification in the confusion matrix for both training and testing dataset with 10-fold cross validation. Figure 9 indicates that both C4.5 and random forest have 100% correctly classified instances while Naïve Bayes model has 13 incorrectly classified instances. Furthermore, Figure 10 indicates that only 1 instance of class 'P' has been incorrectly predicted as class 'NP' by both C4.5 and random forest classifiers whereas Naïve Bayes has 4 instances of that nature. Nevertheless, after 10-fold cross validation as a pruning algorithm for tree size optimization; Figure 11 indicates that random forest outperforms the other models in the number of correct predictions. This emphasizes the fact that tree-structured classification techniques are extremely resistant to outliers.

The WEKA-generated outputs for each model based on training and testing datasets, along with 10-fold cross validation are presented in Figures 12-20 and summarized in Table 4.

a	b	← classified as
41	0	a = P
0	36	b = NP

(a) C4.5 algorithm

a	b	← classified as
41	0	a = P
0	36	b = NP

(b) Random Forest

a	b	← classified as
33	8	a = P
5	31	b = NP

(c) Naïve Bayes

Figure 9: Confusion Matrix of Classification Models (Training dataset)

a	b	← classified as
29	1	a = P
0	3	b = NP

(a) C4.5 algorithm

a	b	← classified as
29	1	a = P
0	3	b = NP

(b) Random Forest

a	b	← classif
26	4	a = P
0	3	b = NP

(c) Naïve Bayes

Figure 10: Confusion Matrix of Classification Models (Testing dataset)

a	b	← classified
38	3	a = P
1	35	b = NP

(a) C4.5 algorithm

a	b	← classified a
40	1	a = P
0	36	b = NP

(b) Random Forest

a	b	← classified as
36	5	a = P
3	33	b = NP

(c) Naïve Bayes

Figure 11: Confusion Matrix of Classification Models after 10-fold Cross Validation

```
Classifier output
=== Evaluation on training set ===

Time taken to test model on training data: 0 seconds

=== Summary ===

Correctly Classified Instances          77              100      %
Incorrectly Classified Instances         0                0      %
Kappa statistic                          1
Mean absolute error                      0
Root mean squared error                  0
Relative absolute error                  0      %
Root relative squared error              0      %
Total Number of Instances               77

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     P
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     NP
Weighted Avg.    1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000

=== Confusion Matrix ===

  a  b   <-- classified as
 41  0 |  a = P
  0 36 |  b = NP
```

Figure 12: C4.5 Training Output

```
Classifier output
=== Evaluation on training set ===

Time taken to test model on training data: 0.02 seconds

=== Summary ===

Correctly Classified Instances          77              100      %
Incorrectly Classified Instances         0                0      %
Kappa statistic                          1
Mean absolute error                      0.0355
Root mean squared error                  0.0615
Relative absolute error                  7.1202 %
Root relative squared error             12.3316 %
Total Number of Instances               77

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     P
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     NP
Weighted Avg.    1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000

=== Confusion Matrix ===

  a  b   <-- classified as
 41  0 |  a = P
  0 36 |  b = NP
```

Figure 13: Random Forest Training Output

```
Classifier output
=== Evaluation on training set ===

Time taken to test model on training data: 0.03 seconds

=== Summary ===

Correctly Classified Instances          64               83.1169 %
Incorrectly Classified Instances        13               16.8831 %
Kappa statistic                          0.6626
Mean absolute error                      0.1517
Root mean squared error                  0.3463
Relative absolute error                 30.4727 %
Root relative squared error             69.4065 %
Total Number of Instances               77

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Clas
                 0.805    0.139    0.868      0.805    0.835      0.665    0.957     0.962     P
                 0.861    0.195    0.795      0.861    0.827      0.665    0.957     0.959     NP
Weighted Avg.    0.831    0.165    0.834      0.831    0.831      0.665    0.957     0.961

=== Confusion Matrix ===

  a  b   <-- classified as
 33  8 |  a = P
  5 31 |  b = NP
```

Figure 14: Naïve Bayes Training Output

```
Classifier output
=== Evaluation on test set ===

Time taken to test model on supplied test set: 0.02 seconds

=== Summary ===

Correctly Classified Instances          32               96.9697 %
Incorrectly Classified Instances         1                3.0303 %
Kappa statistic                          0.8406
Mean absolute error                      0.0303
Root mean squared error                  0.1741
Relative absolute error                  6.3916 %
Root relative squared error             36.6899 %
Total Number of Instances               33

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.967    0.000    1.000      0.967    0.983      0.851    0.983     0.997     P
                 1.000    0.033    0.750      1.000    0.857      0.851    0.983     0.750     NP
Weighted Avg.    0.970    0.003    0.977      0.970    0.972      0.851    0.983     0.975

=== Confusion Matrix ===

  a  b   <-- classified as
 29  1 |  a = P
  0  3 |  b = NP
```

https://afrjmis.net

Figure 15: C4.5 Testing Output

```
Classifier output
=== Evaluation on test set ===

Time taken to test model on supplied test set: 0 seconds

=== Summary ===

Correctly Classified Instances          32               96.9697 %
Incorrectly Classified Instances         1                3.0303 %
Kappa statistic                          0.8406
Mean absolute error                      0.0991
Root mean squared error                  0.1776
Relative absolute error                 20.9005 %
Root relative squared error             37.4398 %
Total Number of Instances               33

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Clas
                 0.967    0.000    1.000      0.967    0.983      0.851    0.978     0.998     P
                 1.000    0.033    0.750      1.000    0.857      0.851    0.978     0.806     NP
Weighted Avg.    0.970    0.003    0.977      0.970    0.972      0.851    0.978     0.980

=== Confusion Matrix ===

  a  b   <-- classified as
 29  1 |  a = P
  0  3 |  b = NP
```

Figure 16: Random Forest Testing Output

```
Classifier output
=== Evaluation on test set ===

Time taken to test model on supplied test set: 0.01 seconds

=== Summary ===

Correctly Classified Instances          29               87.8788 %
Incorrectly Classified Instances         4               12.1212 %
Kappa statistic                          0.5417
Mean absolute error                      0.1388
Root mean squared error                  0.2783
Relative absolute error                 29.2759 %
Root relative squared error             58.6577 %
Total Number of Instances               33

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.867    0.000    1.000      0.867    0.929      0.609    0.989     0.999     P
                 1.000    0.133    0.429      1.000    0.600      0.609    0.989     0.917     NP
Weighted Avg.    0.879    0.012    0.948      0.879    0.899      0.609    0.989     0.991

=== Confusion Matrix ===

  a  b   <-- classified as
 26  4 |  a = P
  0  3 |  b = NP
```

Figure 17: Naïve Bayes Testing Output

```
Classifier output

Time taken to build model: 0.01 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances          73               94.8052 %
Incorrectly Classified Instances         4                5.1948 %
Kappa statistic                          0.896
Mean absolute error                      0.0575
Root mean squared error                  0.2152
Relative absolute error                 11.528  %
Root relative squared error             43.0843 %
Total Number of Instances               77

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.927    0.028    0.974      0.927    0.950      0.897    0.967     0.959     P
                 0.972    0.073    0.921      0.972    0.946      0.897    0.967     0.949     NP
Weighted Avg.    0.948    0.049    0.949      0.948    0.948      0.897    0.967     0.955

=== Confusion Matrix ===

  a  b   <-- classified as
 38  3 |  a = P
  1 35 |  b = NP
```

Figure 18: 10-fold Cross Validation on C4.5

```
Classifier output

Time taken to build model: 0.05 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances          76               98.7013 %
Incorrectly Classified Instances         1                1.2987 %
Kappa statistic                          0.974
Mean absolute error                      0.0917
Root mean squared error                  0.1554
Relative absolute error                 18.3977 %
Root relative squared error             31.1151 %
Total Number of Instances               77

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.976    0.000    1.000      0.976    0.988      0.974    0.999     0.999     P
                 1.000    0.024    0.973      1.000    0.986      0.974    0.999     0.998     NP
Weighted Avg.    0.987    0.011    0.987      0.987    0.987      0.974    0.999     0.999

=== Confusion Matrix ===

  a  b   <-- classified as
 40  1 |  a = P
  0 36 |  b = NP
```

Figure 19: 10-fold Cross Validation on Random Forest

```
Classifier output

Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances          69               89.6104 %
Incorrectly Classified Instances         8               10.3896 %
Kappa statistic                          0.792
Mean absolute error                      0.1172
Root mean squared error                  0.2838
Relative absolute error                 23.5164 %
Root relative squared error             56.8252 %
Total Number of Instances               77

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.878    0.083    0.923      0.878    0.900      0.793    0.969     0.972     P
                 0.917    0.122    0.868      0.917    0.892      0.793    0.969     0.969     NP
Weighted Avg.    0.896    0.101    0.898      0.896    0.896      0.793    0.969     0.971

=== Confusion Matrix ===

  a  b   <-- classified as
 36  5 |  a = P
  3 33 |  b = NP
```

Figure 20: 10-fold Cross Validation on Naïve Bayes

D. Evaluation Metrics

Though, a confusion matrix provides information needed to determine how well a classification model performs, this information when summarized with a single number, becomes much easier when comparatively analyzing these different models' performance. The estimates of the predictive performance, summarizing how accurately the classifiers were able to predict the true class of the instances under the chosen test module, were generated by WEKA's evaluation module.

Advanced evaluation metrics like accuracy, error rate, precision, recall, F-measure and specificity were used to achieve this along with true positive, true negative, false positive, and false negative values. Accuracy measures the ratio of correct predictions over the total number of instances evaluated while precision is used to measure the positive patterns that are correctly predicted from the total predicted patterns in a positive class. Moreover, recall or sensitivity measures the fraction of positive patterns that are correctly classified, while F-measure represents the harmonic mean between recall and precision.

Finally, error rate is a misclassification error that measures the ratio of incorrect predictions over the total number of instances evaluated, while the specificity is used to measure the fraction of negative patterns that are correctly classified. Whereas mean absolute error (MAE) measures the average of the absolute value of the difference between the predicted and actual values, the root mean square error (RMSE) is the square root of the quadratic loss of the respective classifiers, where p_i is the predicted value, a_i is the actual value, and $\hat{p}$ is the modeled value. Another well-known performance evaluator called receiver operating characteristic (ROC) was also computed and the ROC area values obtained. The

ROC curve simply takes the false positive rate (1-specificity) which indicates the level of misalignment in the positive category as *x* axis and the true positive rate (sensitivity) which indicates the level of correct classification in the positive category as *y* axis. Table 4 presents a comparison of the classification models based on these measures and the expressions in equations (7) – (14):

$$\text{Accuracy} = \frac{\text{TP + TN}}{\text{TP + TN + FP + FN}} \quad (7)$$

$$\text{Error rate} = \frac{\text{FP+FN}}{\text{TP + TN + FP + FN}} \quad (8)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP + TN}} \quad (9)$$

$$\text{Recall (Sensitivity)} = \frac{\text{TP}}{\text{TP + FN}} \quad (10)$$

$$\text{F} - \text{Measure} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision + Recall}} \quad (11)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN+FP}} \quad (12)$$

$$\text{MAE} = \sum_{i=1}^{n} \frac{|p_i - a_i|}{n} \quad (13)$$

$$\text{RMSE} = \frac{1}{n}\sqrt{\sum_{i=1}^{n}(p_i - \hat{p})^2} \quad (14)$$

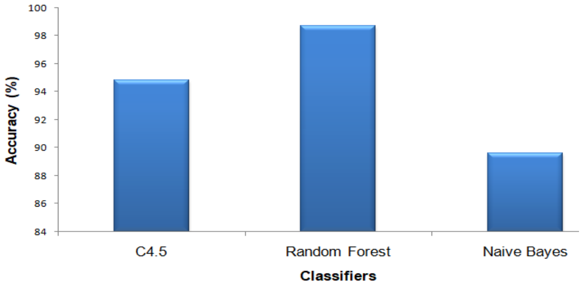


Figure 21: Accuracy Performance of the Classifiers

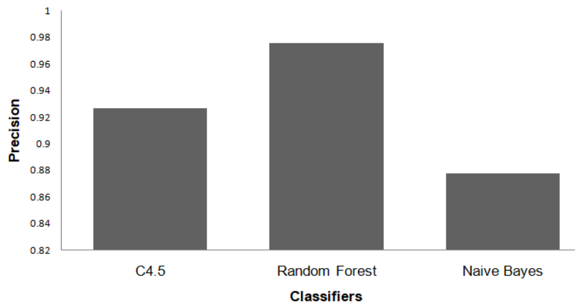


Figure 22: Precision Performance of the Classifiers

Table 4: Performance Comparison of the Classification Models

Metric	C4.5 Test/Validate	RF Test/Validate	NB Test/Validate
Accuracy	96.97; 94.81 (%)	96.97; 98.70 (%)	87.88; 89.61 (%)
Error Rate	3.03; 5.19 (%)	3.03; 1.3 (%)	12.12; 10.39 (%)
MAE	0.0303; 0.0575	0.0991; 0.0917	0.1388; 0.1172
RMSE	0.1741; 0.2152	0.1776; 0.1554	0.2783; 0.238
Correctly Classified Instances	32; 73	32; 76	29; 69
Incorrectly classified instances	1; 4	1; 1	4; 8
Precision	0.967; 0.927	0.967; 0.976	0.867; 0.878
F-Measure	0.983; 0.953	0.983; 0.988	0.929; 0.900
Recall (Sensitivity)	1.000; 0.974	1.000; 1.000	1.000; 0.923
Specificity	0.750; 0.921	0.750; 0.973	0.429; 0.868
ROC Area	0.983; 0.967	0.978; 0.999	0.989; 0.969
Time Taken to Build Model	0s	0.02s	0.03s

Table 4 shows that C4.5 and random forest algorithms have almost the same performance while the later takes more time to build the classifier model. Also, the training instances have been correctly predicted by C4.5 and random forest algorithms each by 96.97% while Naïve Bayes prediction accuracy is 87.88%. This indicates that the results obtained from the training data are optimistic compared with what might be obtained from the independent test set from the same source. More specifically, the MAE and the RMSE of the respective classifiers are not 1 or 0 indicating that not all training instances are classified correctly.

Results from a 10-fold cross validation indicate that random forest classifier yields the best classification accuracy of 98.70% and outperform other algorithms in staff performance and promotion prediction. It further shows that random forest classifier produced significantly higher precision, recall and F-measure values than the other models, confirming the results in Razali *et al.* (2020) and Baati and Mohsil (2020). In addition, Figures 21-23 is a graphical representation of the prediction accuracy, precision and specificity of the models. Finally, the area under the ROC curve in Figure 24 shows that the classifier gives a higher probability of prediction in the case of true positive than in the case of a true negative and achieves a good measure of separability.

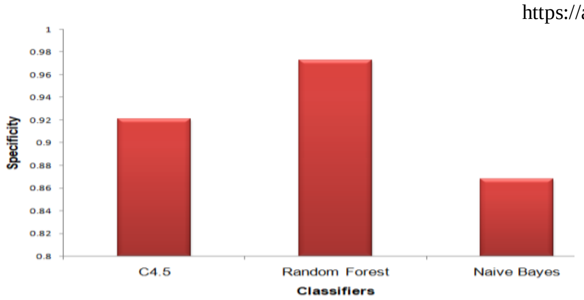


Figure 23: Specificity Performance of the Classifiers

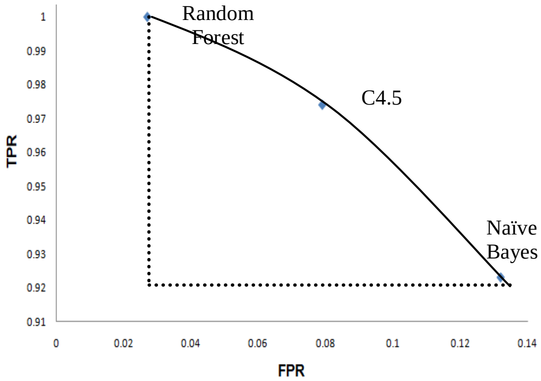


Figure 24: ROC Curve Plot

5. CONCLUSION AND FUTURE WORKS

The employee performance analysis outcome indicates that all the models have acceptable prediction accuracy but random forest is the best in terms of higher prediction accuracy and lower error rate. The generated models can be utilized by decision makers and HR managers for predicting the performance of employees and recommending them for promotion, reward or training. A 10-fold cross validation was carried out for pruning and tree size optimization which generated a rule set and pruned tree with C4.5 algorithm. The attribute with the highest gain ratio is planning and organizing work, followed by team work and communication skills.

From the developed framework, those promoted will automatically have grade level change with increased salary while those not promoted could either be rewarded with salary adjustment as normal increment or recommended for professional training opportunities for improved competency and retention. Future work shall consider the use of dataset with a greater number of employees along with comparison of predictive model accuracy and robustness with other techniques like support vector machine and K-nearest neighbors.

ACKNOWLEDGEMENT

The authors are grateful for the constructive criticism on the manuscript during the review process.

REFERENCES

[1] Aksakala, Erdem; Dağdeviren, Metin; Eraslanb, Ergün and Yukselc, İhsan (2013). 'Personnel Selection based on Talent Management', *Social and Behavioral Sciences*, Vol. 73, pp. 68-72.

[2] Kateřina, Venclová; Andrea, Šalková and Gabriela, Koláčková (2013). 'Identification of Employee Performance Appraisal Methods in Agricultural Organizations', *Journal of Competitiveness*. Vol. 5, Issue 2, pp. 20-36.

[3] Aggarwal, Ashima and Thakur, Gour S. M. (2013). 'Techniques of Performance Appraisal – A Review', *International Journal of Engineering and Advanced Technology*, Vol. 2, Issue 3, pp. 617-621.

[4] Pandey, Mrinal and Sharma, Vivek K. (2013). 'A Decision Tree Algorithm Pertaining to the Student Performance Analysis and Prediction', *International Journal of Computer Applications,* Vol. 61, No. 13, pp. 1-5.

[5] Inyang, Udoinyang G.; Umoh, Uduak A; Nnaemeka, Ifeoma and Robinson, Samuel A. (2019). 'Unsupervised Characterization and Visualization of Students' Academic Performance Features', *Computer and Information Science*, Vol. 12, No. 2, pp. 103-116.

[6] Jantan, H., Hamdan, A. R., Othman, Z. A & Puteh M. (2010a). Applying Data Mining Classification Techniques for Employee's Performance Prediction. 2010 5th International Conference on Knowledge Management (KMICe2010), Kuala Terengganu-Malaysia, pp 645-652.

[7] Nasr, Mona; Shaaban, Essam and Samir, Ahmed (2019). 'A Proposed Model for Predicting Employees' Performance using Data Mining Techniques: Egyptian Case Study', *International Journal of Computer Science and Information Security,* Vol. 17, No. 1, pp. 31-40.

[8] Jantan, H.; Yusoff, N. M. & Noh, M. R. (2014). Towards Applying Support Vector Machine Algorithm in Employee Achievement Classification. *2014 International Conference on Data Mining, Internet Computing, and Big Data,* Kuala Lumpur-Malaysia, pp. 12-21.

[9] Kirimi, John M. and Moturi, Christopher A. (2016). 'Application of Data Mining Classification in Employee Performance Prediction', *International Journal of Computer Applications*, Vol. 146, No. 7, pp. 28-35.

[10] Tupe, Shubham; Mahajan, Chetan; Uplenchwar, Dnyaneshwar and Deo, Pratik (2017). 'Employee

Performance Evaluation System Using ID3 Algorithm', *International Journal of Innovative Research in Computer and Communication Engineering*, Vol. 5, Issue 2, pp. 1757-1761.

[11] Lamarca, Bryan I. J. and Ambat, Shaneth C. (2018). 'The Development of a Performance Appraisal System using Decision Tree Analysis and Fuzzy Logic', *International Journal of Intelligent Engineering and Systems, Vol.* 11, No. 4, pp. 11-19.

[12] Shaout, Adnan and Yousif, Mohamed K. (2014). 'Performance Evaluation - Methods and Techniques Survey', *International Journal of Computer and Information Technology*, Vol. 3, Issue 5, pp. 966-979.

[13] Duraisingam, V. and Skinner, N. (2005). Performance Appraisal. In N. Skinner, A. M. Roche, J. O'Connor, Y. Pollard, and C. Todd (Eds.), Workforce Development TIPS (Theory Into Practice Strategies): A Resource Kit for the Alcohol and Other Drugs Field, National Centre for Education and Training on Addiction (NCETA), Flinders University, Adelaide-Australia.

[14] Dessler, G. (2011). A Framework for Human Resource Management (10th ed.). New York: Prentice Hall.

[15] Peter, R. E; Ajiboye, A. R. and Kareyo, M. (2019). 'Modelling of Key Performance Indicators for Staff Advancement in Higher Institutions of Learning - A Fuzzy Logic Approach', *African Journal of Management Information System*, Vol. 1, Issue 3, pp. 76 -90.

[16] Karaolis, Minas A.; Moutiris, Joseph A.; Hadjipanayi, Demetra and Pattichis, Constantinos S. (2010). 'Assessment of the Risk Factors of Coronary Heart Events based on Data Mining with Decision Trees', *IEEE Transactions on Information Technology in Biomedicine*, Vol. 14, No. 3, pp. 559-566.

[17] Neogi, Amartya; Mondal, Abhoy C. and Mandal, Soumitra K. (2011). 'A Cascaded Fuzzy Inference System for University Non-Teaching Staff Performance Appraisal', *Journal of Information Processing Systems*, Vol. 7, No. 4, pp. 595-612.

[18] Su X; Zhan Z-R. & Li, C-P. Forecasting Regional Talent Demand Based on Support Vector Machine. 2012 International Conference on Machine Learning and Cybernetics (ICMLC), Xian, pp. 355-359.

[19] Shang, C. & Barnes, D. (2012). Support Vector Machine-based Classification of Rock Texture Images Aided by Efficient Feature Selection. 2012 IEEE World Congress on Computational Intelligence (WCCI), Brisbane-Australia, pp. 1-8.

[20] Sarda, Vasudha; Sakaria, Prasham and Nair, Sindhu (2014). 'Relevance Ranking Algorithm for Job Portals', *International Journal of Current Engineering and Technology*, Vol. 4, No. 5, pp. 3157-3160.

[21] Valle, Mauricio A.; Varas, Samuel and Ruz, Gonzalo A. (2012). 'Job Performance Prediction in a Call Center using a Naïve Bayes Classifier', *Expert Systems with Applications*, Vol. 39, Issue 11, pp. 9939-9945.

[22] Ancheta, R. A; Cabauatan, R. J. M; Lorena, B. T. T and Rabago, W. (2012). 'Predicting Faculty Development Trainings and Performance Using Rule-based Classification Algorithm', *Asian Journal of Computer Science and Information Technology*, Vol. 2, No. 7, pp. 203-209.

[23] Al-Radaideh, Qasem A. and Al-Nagi, Eman (2012). 'Using Data Mining Techniques to Build a Classification Model for Predicting Employees Performance', *International Journal of Advanced Computer Science and Applications*, Vol. 3, No. 2, pp. 144-151.

[24] Jantan, Hamidah, Hamdan, Abdul R. and Othman, Zulaiha A. (2010b). 'Human Talent Prediction in HRM using C4.5 Classification Algorithm', *International Journal on Computer Science and Engineering*, Vol. 2, No. 8, pp. 2526-2534.

[25] Hazra, Souvik and Sanyal, Satyaki (2016). 'Recruitment Prediction using ID3 Decision Tree', *International Journal of Advance Engineering and Research Development*, Vol. 3, Issue 10, pp. 48-52.

[26] Fashoto, Stephen G; Ajiboye, Priscilla; Owolabi, Olumide; Metfula, Andile and Fashoto, Olubunmi Y. (2018). 'Assessment of the Readiness of Academic Staff of a Tertiary Institution for Performance Evaluation using a Dynamic Human Resource Information System', *African Journal of Computing & ICT*, Vol. 11, No. 2, pp. 53-75.

[27] Góes, Anderson S. and de Oliveira, Roberto C. L. (2020). 'A Process for Human Resource Performance Evaluation Using Computational Intelligence: An Approach Using a Combination of Rule-based Classifiers and Supervised Learning Algorithms', *IEEE Access*, Vol. 8, pp. 39403-39419.

[28] Subasi, Abdulhamit; Dammas, Dalia H; Alghamdi Rahaf D; Makawi Raghad A; Albiety Eman A; Brahimi, Tayeb and Sarirete, Akila (2018).

'Sensor-based Human Activity Recognition Using Adaboost Ensemble Classifier', *Computer Science*, Vol. 140, pp. 104-111.

[29] Baati, Karim and Mohsi, Mouad (2020). Real-Time Prediction of Online Shoppers' Purchasing Intention Using Random Forest, In I. Maglogiannis et al. (eds.): IFIP International Federation for Information Processing, pp. 43-51.

[30] Zhang Zhifei; Zhao, Zijian and Yeom, Doo-Seoung (2020). 'Decision Tree Algorithm-based Model and Computer Simulation for Evaluating the Effectiveness of Physical Education in Universities', *Complexity*, Vol. 2020, pp. 1-11.

[31] Ak, Muhammet F. (2020). 'A Comparative Analysis of Breast Cancer Detection and Diagnosis Using Data Visualization and Machine Learning Applications', *Healthcare*, Vol. 8, No. 111, pp. 1-23.

[32] Razali, Nazim; Mostafa, Salama A.; Mustapha Aida; Wahab, Mohd H. A. and Ibrahim, Nurul A. (2020). 'Risk Factors of Cervical Cancer using Classification in Data Mining', *Journal of Physics: Conference Series*, Vol. 1529, pp. 1-8.

[33] Anuradha, C. and Velmurugan T. (2015). 'A Comparative Analysis on the Evaluation of Classification Algorithms in the Prediction of Students Performance', *Indian Journal of Science and Technology*, Vol. 8, Issue 15, pp. 1-12.

[34] Abdulsalam, S. O.; Adewole, K. S.; Akintola, A. G. and Hambali, M. A. (2014). 'Data Mining in Market Basket Transaction: An Association Rule Mining Approach', *International Journal of Applied Information Systems*, Vol. 7, No. 10, pp. 15-20.

[35] Shruthi, P. and Chaitra, B. P. (2016). 'Student Performance Prediction in Education Sector using Data Mining', *International Journal of Advanced Research in Computer Science and Software Engineering*, Vol. 6, Issue 3, pp. 211-218.

[36] Olaniyi, Abdulsalam S; Kayode, Saheed Y; Abiola Hambali M; Tosin Salau-Ibrahim T and Babatunde, Akinbowale N. (2017). 'Student's Performance Analysis using Decision Tree Algorithms', *Annals of Computer Science Series*, Vol. 15, No. 1, pp. 55-62.

[37] Quinlan, J. Ross (1987). 'Simplifying Decision Trees', *International Journal of Man-Machine Studies*, Vol. 27, pp. 221-234.

[38] Breiman, L.; Friedman, J. H.; Olshen, R. A. & Stone, C. J. (1984). Classification and Regression Trees. New York: Chapman and Hall.

[39] Islam, Mohammed J.; Wu, Jonathan Q. M.; Ahmadi, Majid and Sid-Ahmed, Maher A. (2010). 'Investigating the Performance of Naïve Bayes Classifiers and K-Nearest Neighbor Classifiers', *Journal of Convergence Information Technology*. Vol. 5, No. 2, pp. 133-137.

[40] Mohammed M.; Khan M. B. & Bashier E. B. M. (2017). Machine Learning: Algorithms and Applications. London: CRC Press.

[41] Fong, Pui K and Weber-Jhanke, J. H. (2012). 'Privacy Preserving Decision Tree Learning using Unrealized Data Sets', *IEEE Transactions on Knowledge and Data Engineering*, Vol. 24, No. 2, pp. 353-364.

[42] Quinlan, J. Ross (1993). C4.5: Programs for Machine Learning. San Mateo, CA: Morgan-Kaufmann Publishers.

[43] Breiman, Leo (2001). 'Random Forests', *Machine Learning*, Vol. 45, pp. 5-32.

[44] Witten, I., Frank, B. & Hall, M. (2011). Data Mining: Practical Machine Learning Tools and Techniques (3rd ed.). New York: Morgan Kaufman Publishers.