Journal of Open Innovation: Technology, Market, and Complexity 10 (2024) 100243

Contents lists available at ScienceDirect

Journal of Open Innovation: Technology, Market, and Complexity

journal homepage: www.sciencedirect.com/journal/journal-of-open-innovation-technology-market-and-complexity

Unbiased employee performance evaluation using machine learning

Zannatul Nayem [1], Md. Aftab Uddin [*,2]

Department of Human Resource Management, University of Chittagong, Chittagong 4331, Bangladesh

ARTICLE INFO

Keywords:
Machine learning
Employee performance
Environmental/physical factors
Social/behavioral factors
Economic factors

ABSTRACT

Most of the companies' sustainability and growth depend on how well its employees perform. However, the measurement of employees' performance until now is inconclusive and inexhaustive. To accurately assess and predict an employee's performance, numerous external factors (physical/environmental, social, and economic) related to an employee's life have been taken into account in this work. The purpose of this research is to explore an unbiased AI algorithmic solution to predict future employee performance considering physical, social, and economic environmental factors that affect employee performance. We collected data of 1109 employees from the 'For-Profit Organization' in Bangladesh from both employers and employees to cover all the factors that justified the unbiased outcome. We utilized a few machine learning tools in this study including the Logistic Regression classifier, the Gaussian Naive Bayes, the Decision Tree classifier, the K-Nearest Neighbors (K-NN), the SVM classification, etc., in order to predict the employee performance evaluation. Then, we compared the effectiveness of those machine learning models by analyzing their precisions, recall, F1-score, and accuracy. This work can be utilized to obtain bias-free employee performance reviews. This fair employee performance assessment can aid decision-makers in making moral choices regarding employee promotions, career advancement, and training needs, among other things. The study also describes notes for future researchers.

1. Introduction

Performance evaluation, which includes assessing current performance, identifying good and poor performers, and providing feedback to staff, is one of the most challenging components of Human Resource Management (HRM) (Cherian et al., 2021; Stone et al., 2015; Fogoros et al., 2020). Employee performance evaluations are not practiced systematically by numerous organizations. As a result, the evaluation method becomes erratic and ineffective. A systematic approach should be adopted in order to evaluate employees at the planning stage on a regular basis (Ahmed et al., 2013). Employees with these attributes—skills, dedication, attitudes, and knowledge—are valued as assets by the company (Al-Tit et al., 2022; Li et al., 2008; Yang and Lin, 2009). By creating new knowledge at firms' level, an organization's human resources can the firms to innovate (Terán-Bustamante et al., 2021). The accurate assessment of the employees' performance contributes to the mission of the company with the maximum satisfaction of the employees (Pap et al., 2022). Since company progress depends on employee advancement, numerous executives search for efficient ways to improve performance drastically (Abbas and Yaqoob, 2009; Salam and Rahmat, 2021). To boost performance, employers first need to know the performance condition of the employee in any organization. An article in the Harvard Business Review (Antonio, 2018) claims that essential functions of the organizations, such as prediction, upselling, cross-selling, and performance management can be remarkably influenced by AI technologies (Ledro et al., 2023). In the future, businesses, communities, and nations will be significantly impacted by big data, automation, and machine learning (Lada et al., 2023; Tao, Gandomkar, Li, Brennan, and Reed, 2023). In this regard, AI has started to play a role in business, particularly in HRM, concerning the prediction and decision-making (Nilashi et al., 2023; Qureshi et al., 2023).

Moreover, AI becomes very helpful in predicting staff turnover, future performance, and employee satisfaction. The performance of an employee is influenced by a wide range of external elements relevant to their lives. These elements should also be considered in the analysis to effectively assess and anticipate an employee's performance (Sasikumar et al., 2021). Numerous dynamic factors went unacknowledged or unaccounted for in the majority of recent studies, and thus raises concerns

* Corresponding author.
E-mail address: mdaftabuddin@cu.ac.bd (Md.A. Uddin).
[1] *ORCID: 0009-0004-5488-2202*
[2] *ORCID: 0000-0002-9101-7451*

https://doi.org/10.1016/j.joitmc.2024.100243
Received 14 January 2024; Received in revised form 11 February 2024; Accepted 20 February 2024
Available online 23 February 2024

about the fairness of employee performance reviews (Patel et al., 2022; Arasi and Babu, 2019; Sujatha and Dhivya, 2022). However, the study by Rodgers et al. (2023) provides a framework that addresses some of the problems with decision-making systems in HRM by including social, economic, and physical factors in the framework. When a business uses multiple factors manually to evaluate an employee's performance, the process gets complicated since different rules are followed and preferences vary for each criterion which is why trying to account for all elements becomes a tiring and arduous procedure (Ahmed et al., 2013). An organization can quickly obtain unbiased staff performance reviews with the help of machine learning by utilizing a framework for prediction and decision-making that incorporates the physical, social/-behavioral, and economic factors impacting employee performance. In this study we aimed to provide a framework or solution mechanism for making predictions and decisions with the use of machine learning algorithms.

This study contributes in numerous ways. Firstly, it provides a particular avenue to unbiasedly predict employee performance reviews with the help of machine learning. Secondly, it covers the natural, social, and economic contextual aspects that can affect employee performance to close the gap left by earlier studies and develop an employee performance evaluation that is free of bias. Finally, this evaluation of employee performance can help decision-makers make accurate decisions regarding, among other things, employee promotions, career progression, and training requirements.

2. Literature review

Assessment of an employee's achievement within a given timeframe is the primary goal of employee performance reviews. In order to provide benchmarks that the employee can accomplish the organizational strategic goal, it identifies the areas of achievement and failure (Abdullah, 2014; Salam and Rahmat, 2021). In this study, we included three sets of factors that affect the performance of an employee. Those three different factors are **physical or environmental factors, social or behavioral factors, and economic factors.**

Physical or environmental factors are the surrounding environmental factors of an employee and his or her workplace, which include cleanliness of the workplace, noise distraction at the workplace, and uncomfortable set up of work area, which can cause any sickness such as eyesight problems, back pain and headache, adequate amount of light and fresh air, a daily journey of reaching office from home or coming home from the office. According to Platis et al. (2015), employee attitudes, behaviors, happiness, and productivity are significantly impacted by the indoor environments of the workplace. An easy working environment is essential for workers to concentrate and perform their jobs flawlessly. Hafee et al. (2019) studied whether workplace environment affects employee job performance, and found out that one unit of variation in physical environment factors results in a 35% change in employee health. When one unit of change happens in employee health, employee performance rises by 80%. Physical environmental factors are having a good impact on both employee health and employee performance. To predict future performance of the employee, physical environmental aspects should be taken into account. Naharuddin and Sadegi (2013) found that the workplace environment is a significant factor affecting employees' performance. They polled 139 employees, and the findings revealed that a physically ordered workplace and other perks like different task aids have a greater influence. Gunaseelan and Ollukkaran (2012) found that different workplace aspects have an impact on performance of the employee of the manufacturing sector in their study.

Some of the **social or behavioral factors** are - employee work-life balance, employee welfare, supervisor support, safety, teamwork, security, training facilities, good relationship with colleagues, development of the employee, etc. these factors that are most important as these are highly related with the employee job satisfaction and employee job performance. Employee contentment or happiness has a significant influence on how well a firm operates since it helps employees develop the abilities and skills they need to perform better at work (Alshurideh et al., 2023). Also team effectiveness is affected by how the team is built, team tasks, employees' potentials and relationships among the team members (Trzeciak and Banasik, 2022). In Hafee et al. (2019), they investigated if an employee's work environment has an impact on their performance. They conducted a study and discovered that one unit of variation in the behavioral environment factors affects employee health by 33%. Employee health and performance are both positively impacted by behavioral environmental factors.

According to paper Gunaseelan and Ollukkaran (2012), many workplace factors had effected participants' performance in the manufacturing industry. The results demonstrated that workers' performance was positively impacted by stable employment, training facilities, and secured jobs. Patel et al. (2022) used a few social factors and compared them separately with the performance rating to investigate the relationship between them. When comparing the performance grade, factors such as tenure with the company, employee work-life balance, training facility, and distance to home were taken into consideration. Arasi and Babu (2019) used three types of employee categories to distinguish between high- and low-performing personnel since those categories impact employee performance. They used achievement category, leadership category, and behavioral category. With regard to the achievement category, authors consider an employee's attendance, discipline, and work output attributes by analyzing how and whether the person achieved those objectives. These guidance and cooperation skills are determined for the leadership category and focus on a worker's capacity for anticipating the need for change, weighing risk, expressing ideas, and responding appropriately. In terms of the behavior category, dependability, integrity traits, behavior, and employee reaction to work procedures are recognized. Shahzad (2014) investigated the connection between corporate culture and employee performance using a survey-based research study, and they looked into a variety of Pakistani software firms. Study's findings demonstrated a positive link between work culture and performance. Results also showed a positive relationship between corporate culture and performance of the worker. They also found that employee commitment and participation boost organizational performance.

In this study, we included three different **economic factors**. Those are overtime payments, bonus based on performance, and salary increment rate. According to the result of Gunaseelan and Ollukkaran (2012), incentives and pay had a positive effect on workers' performance. Aiyetan and Olotuah (2006) determined a few motivational factors that increase employee performance. According to their study, the most commonly used motivational factor for higher performance was overtime since the employees were getting money for doing extra work in the organization. According to Hameed et al. (2014), bonuses or other forms of pay may be offered to employees who perform well. Evidence suggests that merit pay or performance-based pay is somewhat associated with improved performance, even if some studies could not demonstrate a clear impact of this system on performance. Increased compensation is projected to aid sustainable workforce, achieve the vision and mission and work objectives (Ldama and Nasiru, 2020).

Notably, a few researchers have lately begun to use AI to address this crucial HRM issue. However, the algorithmic design of the studies and solutions suggested by those academics may still be influenced by prejudice and lack many dynamic environmental aspects connected to employee performance evaluation. The purpose of the article by Patel et al. (2022) was to assist firms in obtaining employee performance evaluations based on factors that influence them, which began by contrasting various AI-based algorithms, including XGBoost, decision trees, random forests, and artificial neural networks. However, it suggested an ensemble strategy called RanKer that incorporated the aforementioned strategies. Even though there are crucial environmental factors that affect employee performance, a few of them, such as rewards, workplace safety, supervisor support, and recognition, were left out. Arasi and Babu (2019) used a neuro-fuzzy system to distinguish between high- and low-performing personnel and the results showed that the objective function in the employee quality evaluation can be optimized by using the neuro-fuzzy profiling system. This AI method identifies those who require additional training facilities and employee career development in the leadership, achievement, and behavior categories only.

Moreover, these categories do not cover all the areas that influence performance. Sujatha and Dhivya (2022) proposed method that uses XGBoost and gradient boosting to predict employee performance in an MNC organization. They ran statistical tests but did not compute a correlation matrix. Obiedat and Toubasi (2022) proposed a combination method based on ensemble machine learning for predicting employee productivity that also lacks statistical analysis. Fallucchi et al. (2020) used a variety of ML approaches to predict staff attrition, and an industry-standard real-time dataset (provided by IBM) was used to train and test the models. According to the study, most recent studies failed to acknowledge or take into account the various dynamic physical, social, and economic environmental factors, which questions the bias free employee performance evaluation.

Moreover, some of them used minimal volume of data to analyze by using machine learning tools. By incorporating social, economic, and natural environmental elements into the framework, the paper by Rodgers et al. (2023) offered a theoretical framework to help decision-making systems. However, the model was not applied, and there was no statistical analysis using any algorithmic approach. In order to fill up the gap in the prior research and produce an employee performance review that is free of bias, our study incorporated natural, social, and economic environmental factors that can affect the decision maker's perception and the information. Table 1 displays the conclusions of a few analyses of existing research.

Table 1
Analysis of existing literature.

Author	Year	Target and contributions	Research Gap
(Patel et al., 2022)	2022	A proposed ensemble learning technique for rating-based employee performance classification was built using DT, ANN, RF, and XGBoost as some of the component algorithms.	Some of the essential factors that affect employee performance were ignored.
(Arasi and Babu, 2019)	2019	This neuro-fuzzy system enables the algorithm the ability to distinguish between high-performing and low-performing employees, optimizes the objective function in the evaluation of employee abilities, and pinpoints the additional requirements of the training facilities and employee career development in the leadership, achievement, and behavior categories only.	• Target variable contains high and low performing employees only. • Performance evaluation was based on employer perspectives and did not consider the employee's surroundings and some critical factors that affect performance.
(Sujatha and Dhivya, 2022)	2022	The method of this study employs gradient boosting and XGBoost to forecast employee performance in an MNC business.	Statistical analyses were performed, but no correlation matrix was produced.
(Obiedat and Toubasi, 2022)	2022	An ensemble machine learning-based combination method for forecasting employee productivity.	There has been no statistical analysis in this study.
(Fallucchi et al., 2020)	2020	To predict staff attrition, several ML techniques were applied. The models were trained and tested using an industry-recognized real-time dataset (provided by IBM).	They used ML techniques to predict employee attrition.
(Rodgers et al., 2023)	2023	The research [6] provides a TP model that addresses some of the problems with decision-making systems by combining social, economic, and environmental factors into the framework.	• Only a theoretical model • There has been no statistical analysis using any algorithmic approach in this study.

3. Materials and methods

3.1. Research design and research approach

In our study, data were collected through survey questionnaire. Sveral features that affect employee performance will be weighted, including gender, education background, employee job involvement, age, over-time, natural, social and economic environmental features, *etc.* Quantitative and qualitative approaches were used in this research since the qualitative data were transformed into quantitative data. Performance rating was taken as dependent variable while other factors such as physical or environmental factors, employee qualification and experience, employee job satisfaction, social and economic factors etc. were taken as independent variables. Using machine learning, the future performance of employees can be predicted by inserting some of the employee data that demonstrate those crucial factors.

In this scenario, various classification methods, including Gradient Boosting classifier, Gaussian Naive Bayes, K- nearest neighbors (K-NN), Decision Tree, SVM, Logistic Regression, Random Forest etc. have been used to forecast employee performance. The best classifier was also chosen after illustrating the accuracy of those machine learning models. The overall research was accomplished according to the following technical map, as shown in Fig. 1.

3.2. Feature selection

We selected 28 features that cover the three types of factors mentioned in the above section including some of the employee's data such as age, gender, academic qualification *etc.*, and we believe that by collecting data regarding those features, we were able to process those data to get a bias free performance evaluation. In order to cover all the factors that affect employee performance, we went through some of the articles stated in the literature review and collected those features. Here, the dataset's employee performance ratings are provided on a scale from 1 (one) to 3 (three), with 1 denoting poor performance, 2 (two) denoting medium performance and 3 denoting exceptional performance. We took the performance rating as dependent variable while other factors or features such as physical or environmental factors, employee qualification and experience, employee job satisfaction, social factors, and economic factors etc., as independent variables. We employed several supervised machine learning algorithms to classify or categorize the data in order to predict future employee performance because we had both independent and dependent variables. When a problem involves both input and output data, supervised machine learning algorithms are used. In order to perform classification in any problem, classifiers are employed. In order to categorize employee performance, we employed multiple classifiers.

Some of the-

- Physical/environmental features are - noise as distraction, cleanliness, ventilation, comfortable working environment, *etc.*
- Social/behavioral features are – workplace safety, employee work life balance, employee welfare, and training facilities *etc.*
- Economic features are – incentives, overtime pay, and salary increments.

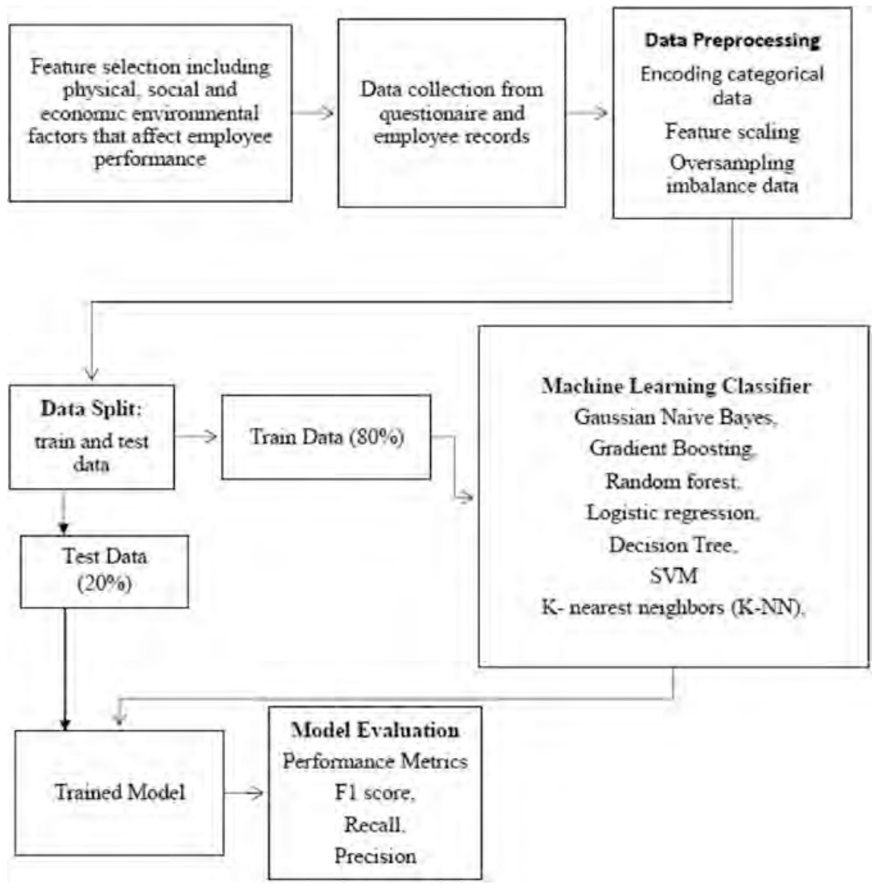

Fig. 1. Technical Map of the study.

Selected independent features for our study showed in the Table 2.

3.3. Data collection and questionnaire design

In order to get data to utilize them, some of the 'For-Profit Organization', such as BSRM, Confidence Salt Ltd., Abul Khair Group, Berger Paints Bangladesh, etc. were targeted since these are the prominent parts of the industrial sector in Bangladesh. We contacted the Deputy General Manager, Deputy Manager, Managing Director, and HR Manager of those companies for the data collection. We collected 1109 sample data from 1109 employees through a questionnaire and from the HR manager through collecting employee records. Thus, we collected the data subjectively and objectively, which made our outcome unbiased. In this survey questionnaire we used dichotomous scale (Yes/No), linear numeric scale (1/2/3) and Likert scale with options including strongly disagree, disagree, neutral, agree and strongly agree. Employee performance ratings for this dataset are given on a scale of 1–3, with 1 representing poor performance, 2 denoting average performance, and 3 denoting excellent performance. Those 28 features that affect employee performance were weighted including gender, education background, employee job involvement, age, overtime, physical, social and economic environmental features, etc.

3.4. Data preprocessing

3.4.1. Encoding categorical data

Models should be provided with the data in integer format to make better predictions. Categorical data must be encoded and transformed into integer format (Verma, 2021). In this study, a Label Encoder was used to encode ordinal categorical data.

3.4.2. Feature scaling

In the case of building machine learning models, the feature scaling is an essential step in the data preprocessing process. It should be done to make the machine learning model stronger or perform better. Only numbers can be seen by the machine learning algorithm. When there is a substantial difference between the data ranges, such as when some data values are in the thousands and some are in the tens, numbers or data with that greater range become more significant and play a more possessive role in training the model (Roy, n.d., 2020). So, feature scaling is necessary to reduce the data range and to give all the data equal importance in training the model. In this study, we used a standard scaler for feature scaling. When scaling the data, the Standard Scaler takes into account that, each characteristic's data are distributed uniformly, and the data are scaled such that the distribution center is 0 and the standard deviation is 1. By computing the relevant statistics on the samples in the training set, the Standard Scaler applies centering and scaling to each feature independently (Roy, n.d., 2020).

$$xnew = \frac{x - \mu}{\sigma}$$

The above equation was used to carry out the usual scaler approach of feature scaling. Here, x is the individual value of the chosen feature, x_{new} is the chosen feature's new value after scaling, μ is the chosen feature's mean, and σ is its standard deviation.

3.5. Training models and evaluation

80% of the data were used to train these machine learning models and the rest 20% were used to test the accuracy level of those models. The accuracy level shows how much the actual performance rating of those 20% data match with the performance rating that is predicted or

Table 2
Features considered for the classification of employee performance.

Serial No.	Feature type	Feature name	Description
1	General Information	Gender	Male or Female
2		Age	Current age of the employee
3		Highest Education	• SSC • HSC • Bachelor/Honors/BBA • Masters/MBA
4	Physical factors	Noise distractions	Have quiet work environment without any noise distractions.
5		Cleanliness	The workplace is dusty and not cleaned properly.
6		Fresh air and natural light	The workplace had complete fresh air and needed natural light.
7		Amount of space	Satisfaction with office space for storage and getting necessary materials on the office desk.
8		Sickness/health problem	Suffer any sickness/ health problem such as Headache, Back pain, Eye side problems, etc., during the employment.
9		Distance from home	Satisfaction with my home-to-office or office-to-home journey.
10	Behavioral/ Social factors	Work-stress	Work under high tensions due to the job.
11		Nervousness before meetings	I feel nervous before attending meetings at my workplace.
12		Work-life balance	I often take my job home with me (Do office work at home).
13		Family-time	I can spend quality time with my family (every day or weekend).
14		Job satisfaction	Would the employees choose the same profession if he or she is given a second chance?
15		Company experience	The number of companies where the employee has work experience.
16		Work experience	Total years of work experience.
17		Training facility	Number of times the employee has completed training.
18		Work experience in current role	Number of years the employee spent in his or her current role.
19		Promotion	Number of years since the employee got his or her promotion.
20		Work experience with current supervisor	Number of years worked under the current supervisor.
21		Work experience at current company	Number of years the employee has work experience at the current company.
22		Interpersonal relationship	Have good relationships with colleagues.
23		Supervisor support	Have satisfactory supervisor support and relationship.
24		Teamwork	I feel good working with the team.
25		Acceptance to change	Employee does not hesitate to accept any change in his/her workplace, such as technological change, functional change, etc.
26	Economic factors	Overtime	Doing overtime or not doing overtime.
27		Salary increment	Last percentage increase in salary (in %).
28		Performance-based bonus	Get a bonus based on performance on the job.

categorized by the machine learning models based on those 80% data. The more the accuracy the more the model is suited for solving the problem or making prediction. Training the machine learning models and evaluating them were discussed in the results section.

4. Results

4.1. Data analysis

Through machine learning and other statistical analysis, we compared the relationship among the data of the selected features that have a significant impact on employee performance.

4.1.1. Correlation matrix

Fig. 2 shows the correlation matrix of the features. Employee performance rating was used as the dependent variable in this study. In contrast, other variables, including work satisfaction, employee qualifications and experience, job-related experience, social and environmental factors, and economic factors, were used as independent variables. Among those variables, age, volume of space, work-life balance, family time, work stress, experience, training facility, acceptance to change, and salary increment features showed a significant relationship with employee performance rating which were near to one in correlation matrix.

There were no neutral relationships (which means the relationship is zero) among the variables in the correlation matrix. We used all 28 of the independent variables in this study since they demonstrated a relationship—positive or negative (not zero)—between performance rating (dependent variable) and them. Figure 3 shows some of the essential features in this study. These features showed a higher and positive relationship with the dependent variable performance rating which are near to 1 in the correlation matrix. The relationship of the data on those important features with the performance rating of those employee are shown in Fig. 4, Fig. 5, Fig. 6, Fig. 7, Fig. 8, and Fig. 9 through statistical analysis. From the analysis of the data between employee performance rating and training facility factor in Figure 4, we can see that employees who have received the most training facilities performed better than those who have received fewer training facilities. The relationship between age and performance rating in Figure 5 shows that employee with age between 30 and 45 performed better than employee with other age. Looking at the relationship between salary increment and performance rating in Figure 6, we can see that the percentage of low performer decreased when the employees received higher salary increment. Relationship between work experience and performance rating in Figure 7 depicts that employees who are working 11 years to 21 years performed better. Figure 8 shows that employees who received earlier promotions performed better. Figure 9 depicts that employees who have been working at the current company for 8 years to 23 years performed better.

4.2. Training ML models and evaluation

In this study, the performance ratings of the employees were encoded into three classes. Performance rating one (Low), two (medium) and three (high) were encoded into respectively class 0, class 1 and class 2.

4.2.1. Gaussian Naive bayes

In this study, gaussian naive Bayes was employed. Although changing the hyperparameter in Gaussian naive Bayes is probably not necessary, we used hyperparameter tuning so that the model performs better. We used priors=None and var_smoothing=1e-9 to get the highest model accuracy. The priors are not altered based on the data unless expressly stated. In variable smoothing, the variances make up a part of the largest variance of all parameters for calculation stableness. The accuracy score that we achieved by using this model is 61.4%. In this model other performance metrics are: Precision, Recall, and F1 score which are

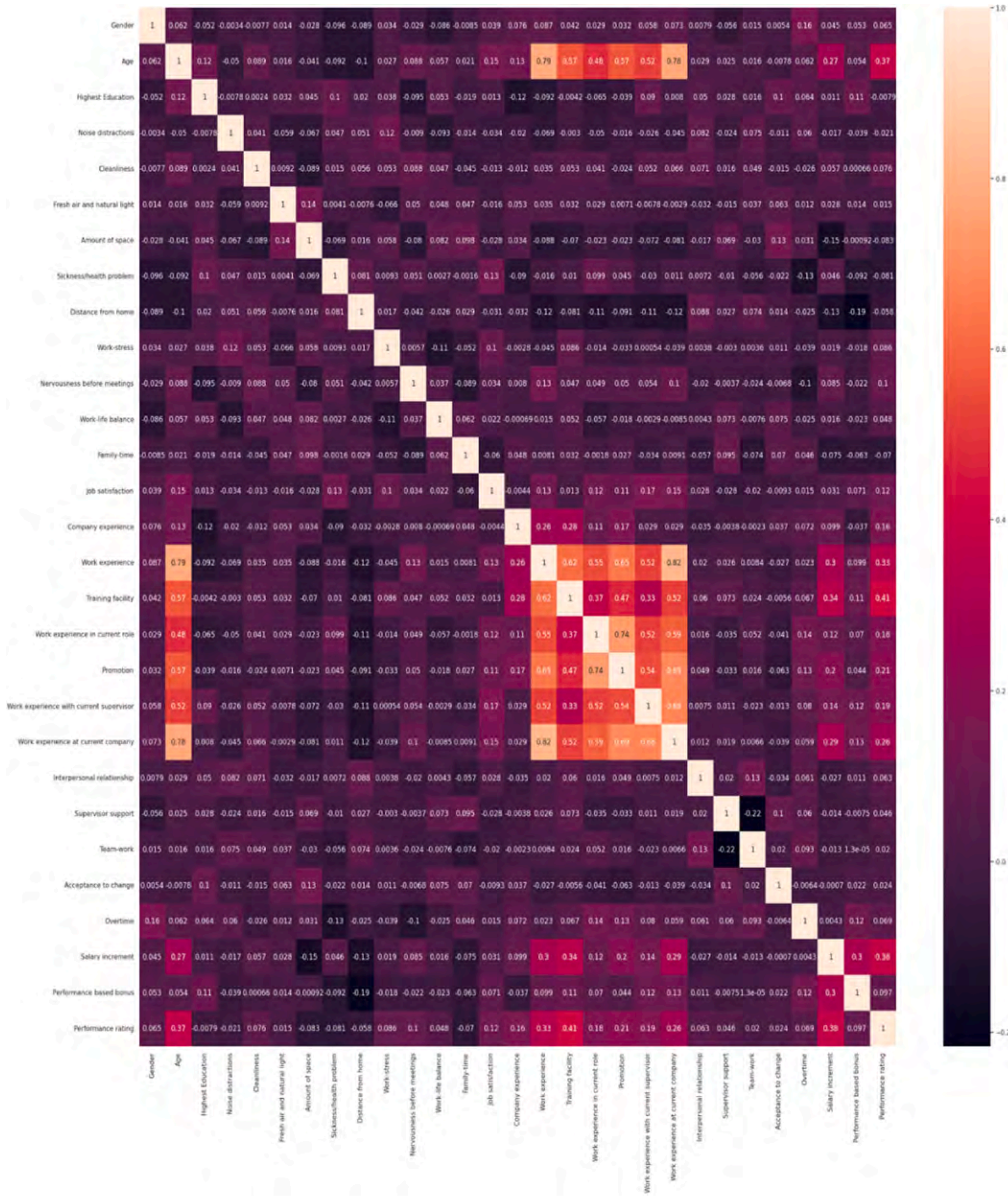

Fig. 2. Correlation Matrix.

shown in Table 3. Precision for class 0 is 60%, class 1 is 54%, and class 2 is 80%. Recall for class 0 is 95%, for class 1 is 47% and for class 2 is 42%. F1 score for class 0 is 73%, for class 1 is 50% and for class 2 is 55%. The Confusion matrix of this model is shown in the following Fig. 10.

4.2.2. Random forest

Random forest is known as a multipurpose and easy-to-use machine learning algorithm. It can consistently obtain outstanding results, even in the absence of hyper-parameter tuning. People use it more frequently since it is simple and diverse (Donges, 2023). Hyperparameter

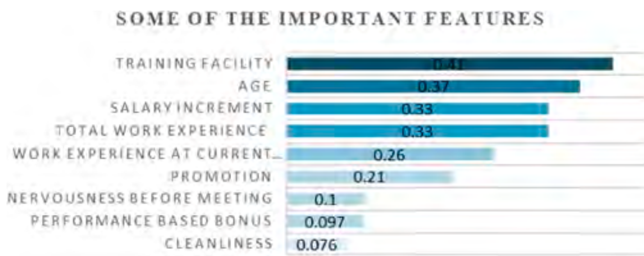

Fig. 3. List of some features in order of importance that are highly related with performance rating.

Fig. 4. Distribution of performance rating by Training Facility.

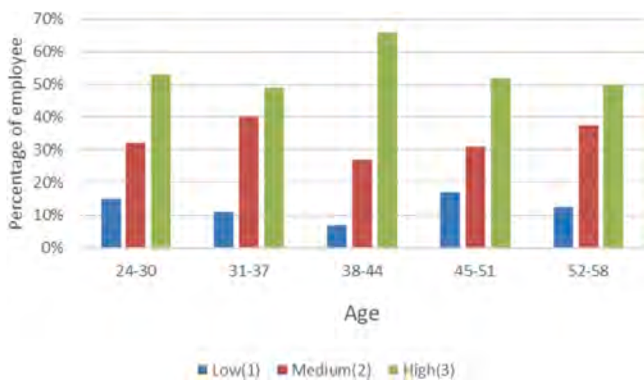

Fig. 5. Distribution of performance rating by age.

Fig. 6. Distribution of performance rating by salary increment.

adjustment was not necessary, since it has given the best performance. We found a 98.2% accuracy score by training this model. In Table 3, other evaluation metrics are displayed. Class 0 has a precision of 100%, class 1 of 99%, and class 2 of 96%. Recall for classes 0, 1, and 2 is 100%, 96%, and 99%, respectively. Class 0 has an F1 score of 100%, class 1 of

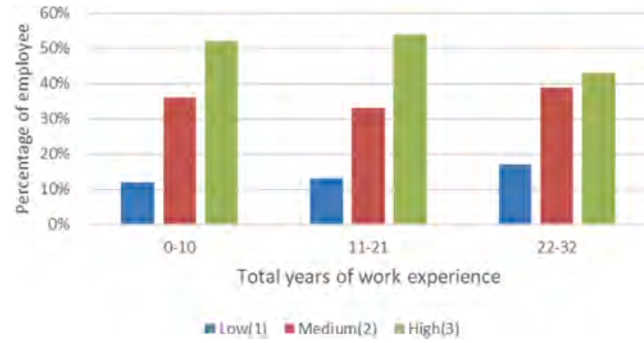

Fig. 7. Distribution of performance rating by total work experience.

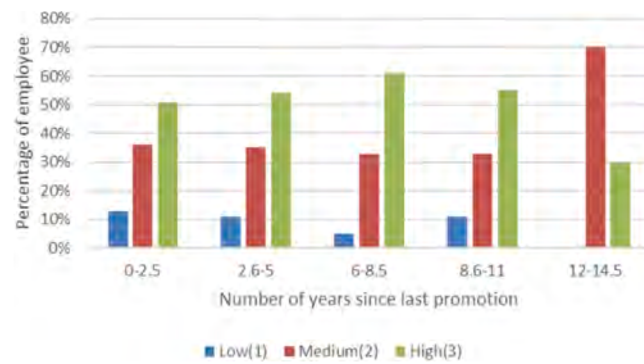

Fig. 8. Distribution of performance rating by promotion.

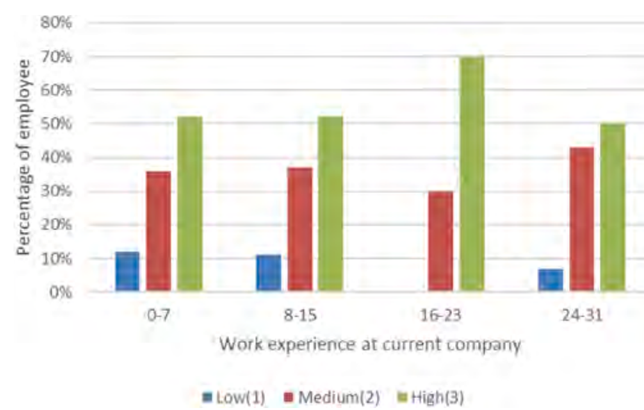

Fig. 9. Distribution of performance rating by work experience at current company.

97%, and class 2 of 97%. The confusion matrix for our model is displayed in Fig. 11.

4.2.3. *Decision tree*

We used max_depth and random_state hyperparameter to tune in decision tree classifier. A stopping condition that restricts the number of splits that can be carried out in a decision tree is what the maximum depth parameter does: it sets a maximum depth. A random_state is used to recreate the same results when using a random procedure. We took max_depth=20, random_state=30 to improve the performance of the model. In this study, we achieved an 89.8% accuracy score while training this model. In Table 3, additional evaluation metrics are displayed. Class 0 precision is 97%, Class 1 precision is 87%, and Class 2 precision is 85%. Recall rates are 95%, 85%, and 89% for classes 0, 1, and 2, respectively. F1 scores for classes 0, 1, and 2 are 96%, 86%, and 87%, respectively. The confusion matrix for this model is displayed in

Table 3
Evaluation metrics for ML models used in this study.

Classifier	classes	Precision	Recall	F1-score	Accuracy score
Gaussian Naive Bayes	0	60%	95%	73%	61.4%
	1	54%	47%	50%	
	2	80%	42%	55%	
Random Forest	0	100%	100%	100%	98.2%
	1	99%	96%	97%	
	2	96%	99%	97%	
Decision Tree	0	97%	95%	96%	89.8%
	1	87%	85%	86%	
	2	85%	89%	87%	
K-nearest neighbors (K-NN)	0	84%	100%	91%	88.1%
	1	87%	86%	87%	
	2	97%	78%	86%	
Gradient Boosting	0	97%	99%	98%	92.4%
	1	90%	89%	89%	
	2	91%	89%	90%	
Logistic Regression	0	75%	89%	81%	67.8%
	1	60%	46%	52%	
	2	66%	69%	67%	
SVM	0	73%	88%	80%	68.4%
	1	60%	53%	57%	
	2	71%	63%	67%	

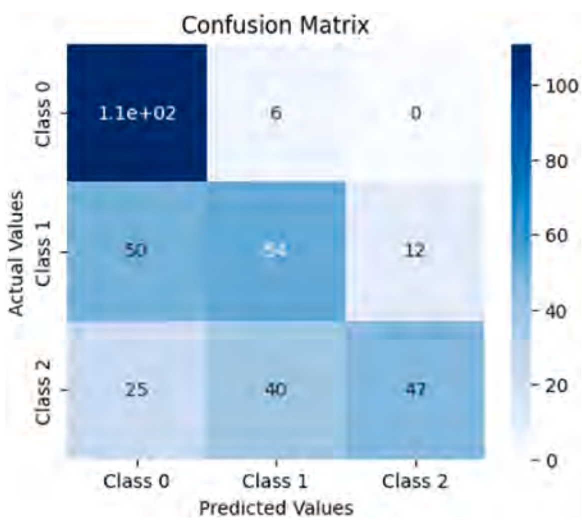

Fig. 10. Confusion matrix of Gaussian naive Bayes.

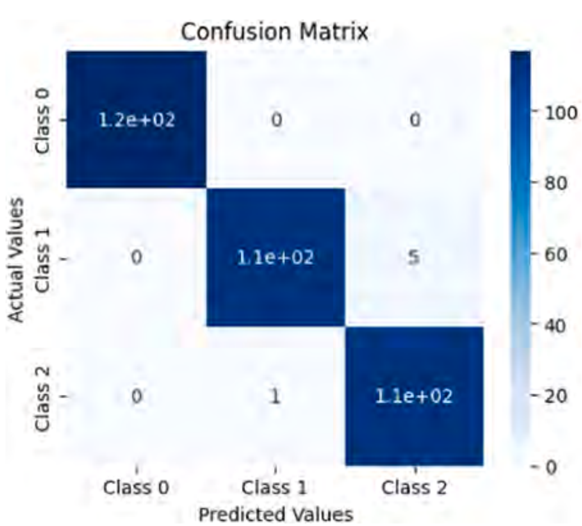

Fig. 11. Confusion matrix of Random forest.

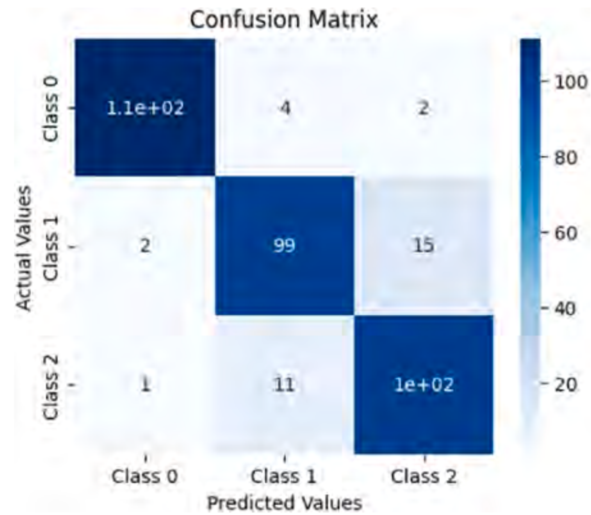

Fig. 12. Confusion matrix of Decision Tree.

Fig. 12 below.

4.2.4. K-nearest neighbors (K-NN)

We tune the hyperparameter by taking n_neighbors=3 in K-nearest neighbors classifier. Here, the number of neighbors that will vote for the target point's class is indicated by the hyperparameter "n_neighbors." We trained this model in our research, and the accuracy score was 88.1%. In Table 3, other evaluation metrics are displayed. For class 0, class 1, and class 2, the model achieved a precision value of 84%, 87%, and 97%, respectively. Additionally, for the class-0 rating, class-1 rating, and class-2 rating, the model obtained recall values of 100%, 86%, and 78%, respectively. F1-Score values for classes 0, 1, and 2 are 91%, 87%, and 86%, respectively. Fig. 13 shows the confusion matrix for our model.

4.2.5. Gradient boosting

The hyperparameter that was used in the Gradient Boosting classifier was n_estimators. The number of trees the algorithm builds before averaging forecasts is indicated by the term "n_estimators." We used n_estimators=45 in this analysis since it performs the best. We trained this model to an accuracy score of 92.4%. In Table 3, additional evaluation metrics are displayed. The model's precision value for classes 0, 1, and 2 was 97%, 90%, and 91%, respectively. Additionally, the model

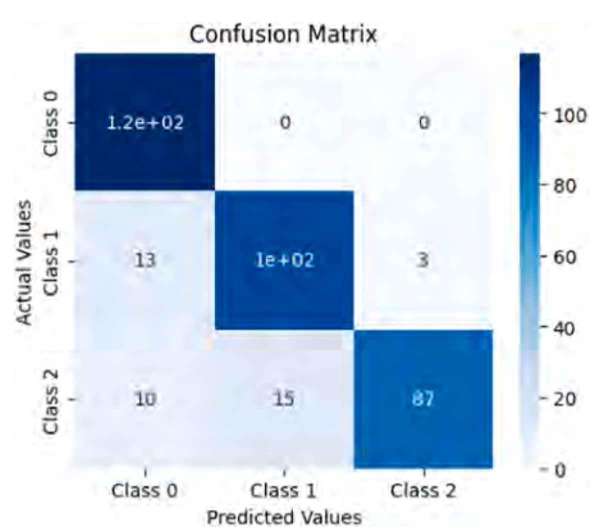

Fig. 13. Confusion matrix of K-nearest neighbors (K-NN).

obtained recall values of 99%, 89%, and 89% for the class-0 rating, class-1 rating, and class-2 rating, respectively. Classes 0, 1, and 2 have F1-Score values of 98%, 89%, and 90%, respectively. In Fig. 14 below, the confusion matrix for this model is shown.

4.2.6. Logistic regression

Although it is probably not necessary to modify the hyperparameter in Logistic Regression, we used hyperparameter C=5 and random_state=0, where the penalty strength is controlled by the C parameter, which might also be advantageous and to replicate the same outcomes while utilizing a random process, a random_state is helpful. When trained on the standard dataset, this model provides 67.8% accuracy. Precision for class 0 is 75%, class 1 is 60% and class 2 is 66%. Recall for class 0 is 89%, for class 1 is 46% and for class 2 is 69%. F1 score for class 0 is 81%, for class 1 is 52% and for class 2 is 67%. The Confusion matrix of this model is shown in the following Fig. 15.

4.2.7. SVM

To tune the SVM classifier, we employed the Kernel and C hyperparameters. The C parameter regulates the penalty strength, and Kernel is used to transform the data into desired form. To train the model, we took kernel= linear and C=40.

Using this model, we were able to attain a 68.4% accuracy score. In Table 3, additional evaluation metrics are displayed. For class-0 ratings, precision is 73%, for class-1 ratings, 60%, and for class-2 ratings, 71%. Class-0 recall is 88%, class-1 recall is 53%, and class-2 recall is 63%. For class-0 rating, class-1 rating, and class-2 rating, respectively, the F1-Score is 80%, 57%, and 67%. The confusion matrix for this model is displayed in Fig. 16 below.

4.3. Performance evaluation and comparison of ML models

This section describes the model's performance used in this study using a variety of performance metrics, including accuracy, recall, precision, and F1-score.

Model that gave us the highest accuracy among other used ML models is Random Forest which is 98.2%. Moreover, the least accuracy score was given by model Gaussian naive Bayes which is 61.4%. Other models gave us satisfactory accuracy scores above 80% and below 100%. The second highest accuracy score was 92.4%, which we achieved by training Gradient Boosting. From the following Fig. 17, we can say that K-NN, Random Forest, Gradient Boosting and Decision Tree models performed better than other models that were trained in this study.

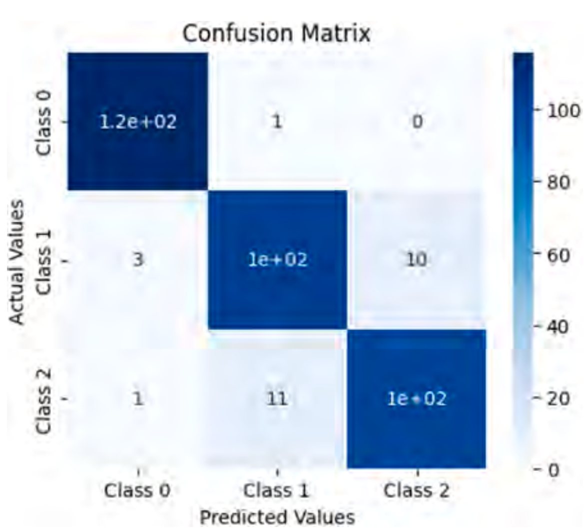

Fig. 14. Confusion matrix of Gradient Boosting.

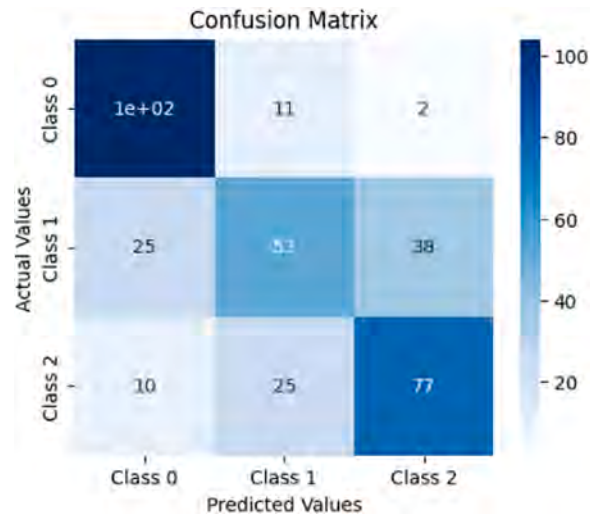

Fig. 15. Confusion matrix of Logistic Regression.

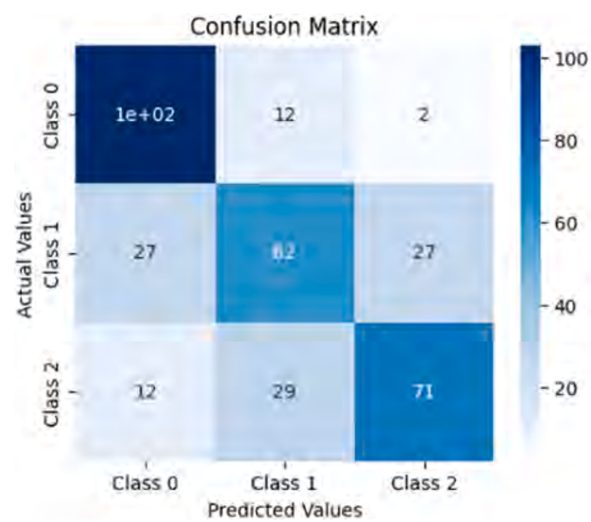

Fig. 16. Confusion matrix of SVM.

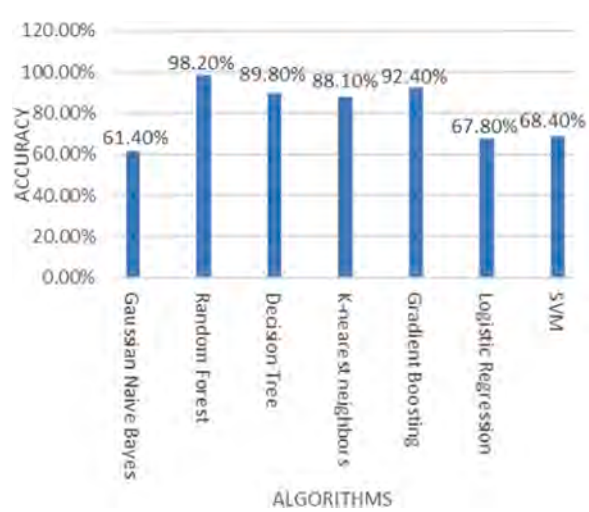

Fig. 17. Performance comparison among proposed models of this study.

5. Discussion

The purpose of this study was to develop a framework for predicting future employee performance in an unbiased way. Seven machine learning models were applied. All the physical, social or behavioral, and economic factors that affect job performance were considered by these models.

The following models were trained in this study: Gaussian Naive Bayes, K-Nearest Neighbors (K-NN), Random Forest, Gradient Boosting, Logistic Regression, SVM, and Decision Tree. Compared to other models, Decision Tree, Random Forest, Gradient Boosting, and K-Nearest Neighbors (K-NN) performed better. We obtained a respectable accuracy score from other models. The Gaussian naïve Bayes model had the lowest accuracy score (61.4%), while the Random Forest model had the best accuracy score (98.2%).

Patel et al. (2022) applied some of the traditional ML approach and a proposed ensemble ranker approach for predicting employee performance where the ranker approach provided the highest accuracy score in that study which was 96.25% and lowest score was given by Artificial Neural Network which was 87.08%. The highest accuracy score of Obiedat and Toubasi (2022) for predicting employee productivity was given by Random Forest, which was 98.3%.

In this study, the dependent variable was employee performance ratings. In contrast, the independent variables were work satisfaction, employee credentials and experience, job-related experience, social and environmental factors, and economic considerations. Among these factors, the relationships between age, space, work-life balance, family time, stress at work, work experience, training facility, openness to change, and salary increment features with employee performance ratings were stronger.

With the aid of machine learning, this study showed how to predict employee performance reviews in an unbiased way. In order to fill the gap created by past studies and provide a non-biased employee performance rating, it included the environmental, social, and economic contextual factors that can influence employee performance.

5.1. Implication of the study

This assessment of employee performance can assist decision-makers in reaching sound decisions about employee promotions, career advancement, training needs, and other issues in HRM. Any firm can use machine learning to forecast future employee performance by including some of the employee data that exemplifies those critical factors used in this study. Organizations can use this framework to get unbiased staff performance reviews very easily with the help of machine learning. This study assumes that its outcome can assist any organization by introducing a fresh avenue to get bias-free employee performance reviews that can guide HR to make neutral HR decisions regarding promotion, employee career development, training requirements, etc. These decisions are essential factors for an organization and also for a country in order to get skilled human resources.

5.2. Strengths of the study

To create an unbiased method of predicting an employee's future performance, this study gathered data from both the employer and the employee via a questionnaire and the employee's record within the company. Moreover, natural, social, and economic factors that impact an employee's performance were also covered. In today's world, predictions and decision-making in any advanced corporation are made with the help of AI to make them more accurate and perfect. Thus, the prediction of employee future job performance was made using machine learning models in this study.

5.3. Scope for future research

In this study we considered 'For Profit organization' and collected the information of the employees of those organization. In the future, we can target any specific industry to apply this research as it was done by targeting all types of 'For Profit organizations'. In this research we used machine learning algorithms, however, we can use viable approach or apply deep learning approach to do the research. We can also suggest a strategy in the future to help any firm protect its employee data from hackers and other intruders.

6. Conclusion

The performance of the company's personnel determines a large portion of its sustainability and growth. Many external aspects (physical/environmental, social, and economic) relevant to an employee's life have been included in this work in order to measure and anticipate an employee's performance effectively. The goal of this project is to create an AI algorithmic-based ethical decision-making framework that takes into account the various environmental factors—physical, social, and economic - that have an impact on worker performance. Our results were impartial since we gathered information from a few "For-Profit Organizations" in Bangladesh, both objectively and subjectively. In this study, we used a variety of machine learning methods, and we were able to acquire a reasonable accuracy score. The Random Forest model obtained the highest accuracy score, and the Gaussian naïve Bayes model had the lowest. An impartial employee performance review can be obtained by using this work. Decision-makers can use this equitable employee performance evaluation to help them make moral decisions about training requirements, career advancement, and employee promotions, among other things.

Funding acknowledgement

This research receives no grants from funding agencies in the public, commercial, or not-for-profit sectors.

Ethical Statement

In this research data were collected through questionnaire from employee and employee records from employer in the organization with full consent of those employees and employer of the target organization. There is no sensitive information in the data. So, ethical approval is not needed in this research.

CRediT authorship contribution statement

Zannatul Nayem: Conceptualization, Data curation, Formal analysis, Methodology, Validation, Software, Writing- original draft preparation. **Md. Aftab Uddin:** Supervision, Writing- review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Authors are grateful to Ms. Mansura Nusrat (PhD fellow in the University of Texas at El Paso), and Ms. Salma Begum, Mr. Mohammed Rafiqul Islam and Mr. Mohammad Tamzid Hossain (Lecturers of Human Resource Management, University of Chittagong) for their voluntary assistance to the development of this paper.

References

Abbas, Q., Yaqoob, S., 2009. Effect of leadership development on employee performance in Pakistan. *Pakistan*. Econ. Soc. Rev. 47 (2), 269–292.

Ahmed, Imtiaz, Sultana, Ineen, Paul, Sanjoy, Azeem, Abdullahil, 2013. Employee performance evaluation: a fuzzy approach. Int. J. Product. Perform. Manag. 62 (7), 718–735. https://doi.org/10.1108/IJPPM-01-2013-0013.

Aiyetan, A.O., Olotuah, A.O., 2006. Impact of motivation on workers' productivity in the Nigerian construction industry, 4-6 September 2006. In: Procs 22nd Annual ARCOM Conference. Association of Researchers in Construction Management, Birmingham, UK, pp. 239–248, 4-6 September 2006.

Alshurideh, M.T., Al Kurdi, B., Alzoubi, H.M., Akour, I., Obeidat, Z.M., Hamadneh, S., 2023. Factors affecting employee social relations and happiness: SM-PLUS approach. J. Open Innov. Technol. Mark. Complex. *9* (2) https://doi.org/10.1016/j.joitmc.2023.100033.

Al-Tit, A.A., Al-Ayed, S., Alhammadi, A., Hunitie, M., Alsarayreh, A., Albassam, W., 2022. The impact of employee development practices on human capital and social capital: the mediating contribution of knowledge management. J. Open Innov. Technol. Mark. Complex. *8* (4) https://doi.org/10.3390/joitmc8040218.

Antonio, V., 2018. How AI is changing sales. *Harvard Business Review*, 30. Retrieved [October 2023], from: ⟨https://hbr.org/2018/07/how-ai-is-changing-sales⟩.

Arasi, M.A., & Babu, S. (2019). *International Journal of Advanced Trends in Computer Science and Engineering Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse39852019.pdf* . *8*(February 2020), 231–237.

Cherian, J., Gaikar, V., Paul, R., Pech, R., 2021. Corporate culture and its impact on employees' attitude, performance, productivity, and behavior: An investigative analysis from selected organizations of the United Arab Emirates (UAE). J. Open Innov.: Technol., Mark., Complex. *7* (1), 1–28. https://doi.org/10.3390/joitmc7010045.

Donges, N. (2023). *Random Forest: A Complete Guide for Machine Learning.* https://builtin.com/data-science/random-forest-algorithm.

Fallucchi, F., Coladangelo, M., Giuliano, R., De Luca, E.W., 2020. Predicting employee attrition using machine learning techniques. Computers *9* (4), 1–17. https://doi.org/10.3390/computers9040086.

Fogoroș, Teodora, Maftei, Mihaela, Bițan, Gabriela, Kurth, Bastian, 2020. Study on methods for evaluating employees performance in the context of digitization. Proc. Int. Conf. Bus. Excell. 14 (1), 878–892. https://doi.org/10.2478/picbe-2020-0084.

Gunaseelan, R., Ollukkaran, B.A., 2012. A study on the impact of work environment on employee performance. Namex Int. J. Manag. Res. 71, 1–16.

Hafee, I., Yingjun, Z., Hafeez, S., Mansoor, R., Rehman, K.U., 2019. Impact of workplace environment on employee performance: mediating role of employee health. Bus., Manag. Educ. *17* (2), 173–193. https://doi.org/10.3846/bme.2019.10379.

Hameed, A., Ramzan, M., Hafiz, M., Kashif Zubair, M., Ali, G., Arslan, M., 2014. Impact of compensation on employee performance. Int. J. Bus. Soc. Sci. *5* (2), 302–309. https://doi.org/10.9790/0837-2509011722.

Lada, S., Chekima, B., Karim, M.R.A., Fabeil, N.F., Ayub, M.S., Amirul, S.M., Ansar, R., Bouteraa, M., Fook, L.M., Zaki, H.O., 2023. Determining factors related to artificial intelligence (AI) adoption among Malaysia's small and medium-sized businesses. J. Open Innov.: Technol., Mark., Complex. *9* (4) https://doi.org/10.1016/j.joitmc.2023.100144.

Ldama, J., Nasiru, M., 2020. Salary increase and its impacts on employee performance in adamawa state university. Peer-Rev., Referee, Index. J. IC *87* (6), 86.

Ledro, C., Nosella, A., Dalla Pozza, I., 2023. Integration of AI in CRM: challenges and guidelines. J. Open Innov.: Technol., Mark., Complex. *9* (4) https://doi.org/10.1016/j.joitmc.2023.100151.

Li, J., Pike, R., Haniffa, R., 2008. Intellectual capital disclosure and corporate governance structure in UK firms. Account. Bus. Res. *38* (2), 137–159. https://doi.org/10.1080/00014788.2008.9663326.

Naharuddin, N.M., Sadegi, M., 2013. Factors of workplace environment that affect employees performance: a case study of Miyazu Malaysia. Int. J. Indep. … *2* (2), 66–78 http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2290214.

Nilashi, M., Ali Abumalloh, R., Samad, S., Minaei-Bidgoli, B., Hang Thi, H., Alghamdi, O. A., Ahmadi, H., 2023. The impact of multi-criteria ratings in social networking sites on the performance of online recommendation agents. Telemat. Inform. *76*, 101919 https://doi.org/10.1016/j.tele.2022.101919.

Obiedat, R., Toubasi, S., 2022. A Combined Approach for Predicting Employees' Productivity based on Ensemble Machine Learning Methods. Inform. (Slov.) *46* (5), 49–58. https://doi.org/10.31449/inf.v46i5.3839.

Pap, J., Mako, C., Illessy, M., Kis, N., Mosavi, A., 2022. Modeling Organizational Performance with Machine Learning. J. Open Innov.: Technol., Mark., Complex. *8* (4), 177 https://doi.org/10.3390/joitmc8040177.

Patel, K., Sheth, K., Mehta, D., Tanwar, S., Florea, B.C., Taralunga, D.D., Altameem, A., Altameem, T., Sharma, R., 2022. RanKer: An AI-based employee-performance classification scheme to rank and identify low performers. Mathematics *10* (19), 1–21. https://doi.org/10.3390/math10193714.

Platis, C., Reklitis, P., Zimeras, S., 2015. Relation between Job Satisfaction and Job Performance in Healthcare Services. Procedia - Soc. Behav. Sci. *175*, 480–487. https://doi.org/10.1016/j.sbspro.2015.01.1226.

Qureshi, A.A., Ahmad, M., Ullah, S., Yasir, M.N., Rustam, F., Ashraf, I., 2023. Performance evaluation of machine learning models on large dataset of android applications reviews. Multimed. Tools Appl. *82* (24), 37197–37219. https://doi.org/10.1007/s11042-023-14713-6.

Rodgers, W., Murray, J.M., Stefanidis, A., Degbey, W.Y., Tarba, S.Y., 2023. An artificial intelligence algorithmic approach to ethical decision-making in human resource management processes. Hum. Resour. Manag. Rev. *33* (1), 100925 https://doi.org/10.1016/j.hrmr.2022.100925.

Roy, B. (n.d.). (2020). All about Feature Scaling. *Towards Data Science*. https://towardsdatascience.com/all-about-feature-scaling-bcc0ad75cb35.

Salam, Rahmat. (2021). The Importance Performance Assessment And Its Impact On Improving Performance Of Public Service Organizations In South Tangerang City. *Sosiohumaniora*. 23.226.10.24198/*sosiohumaniora*.v23i2.31963.

Sasikumar, R., Alekhya, B., Harshita, K., Sree, O.H., Pravallika, I.K., 2021. Employee performance evaluation using sentiment analysis. Rev. Geintec Gest. Inov. Tecnol. 11, 2086–2095.

Shahzad, F., 2014. Impact of organizational culture on employees' job performance: an empirical study of software houses in Pakistan. Int. J. Commer. Manag. *24* (3), 219–227. https://doi.org/10.1108/IJCoMA-07-2012-0046.

Stone, D.L., Deadrick, D.L., Lukaszewski, K.M., Johnson, R., 2015. The influence of technology on the future of human resource management. Hum. Resour. Manag. Rev. *25* (2), 216–231.

Sujatha, P., Dhivya, R.S., 2022. Ensemble Learning Framework to Predict the Employee Performance, 01-03 March 2022 2022 Second International Conference on Power, Control and Computing Technologies (ICPC2T). IEEE, Raipur, India, pp. 1–7. https://doi.org/10.1109/ICPC2T53885.2022.9777078, 01-03 March 2022.

Tao, X., Gandomkar, Z., Li, T., Brennan, P.C., Reed, W., 2023. Using radiomics-based machine learning to create targeted test sets to improve specific mammography reader cohort performance: a feasibility study. J. Pers. Med. *13* (6), 888.

Terán-Bustamante, A., Martínez-Velasco, A., Dávila-Aragón, G., 2021. Knowledge management for open innovation: bayesian networks through machine learning. J. Open Innov. Technol. Mark. Complex. *7*, 40. https://doi.org/10.3390/joitmc.

Trzeciak, M., Banasik, P., 2022. Motivators influencing the efficiency and commitment of employees of agile teams. J. Open Innov.: Technol., Mark., Complex. *8* (4) https://doi.org/10.3390/joitmc8040176.

Verma, Y. (n.d.). (2021). A Complete Guide to Categorical Data Encoding. *Mystery Vault*. India.

Yang, C.C., Lin, C.Y.Y., 2009. Does intellectual capital mediate the relationship between HRM and organizational performance? Perspective of a healthcare industry in Taiwan. Int. J. Hum. Resour. Manag. *20* (9), 1965–1984. https://doi.org/10.1080/09585190903142415.