Published in AJSE, Vol:22, Issue: 3
Received on 4th April 2023
Revised on 13th September 2023
Accepted on 3rd December 2023

AIUB JOURNAL OF SCIENCE AND ENGINEERING
ISSN: 1608 – 3679 (print) 2520 – 4890 (Online)

Employee Promotion Prediction Using Improved AdaBoost Machine Learning Approach

Md Abu Jafor, Md. Anwar Hussen Wadud,
Kamruddin Nur *(Senior Member, IEEE),* and Mohammad Motiur Rahman

***Abstract*—Employee promotion is an important aspect of the human resource management process. Due to different factors, it refers to the automatic improvement among the employees in an organization. It improves their job satisfaction and motivation by providing more significant income, status, and responsibilities. By building up loyalty, promotion reduces employee attrition. But, it is difficult to accurately decide, whether an employee should or should not be promoted based on their current and past performance. So, human resource management does research about promotion prediction, because there are a limited number of research about the finding of employee promotion prediction in the existing studies. The aim of this research study is to implement an employee promotion prediction framework using machine learning. A modified AdaBoost classifier is used for automatic promotion prediction, and six machine learning techniques like Support Vector Machine (SVM), Logistic Regression (LR), Artificial Neural Network (ANN), Random Forest (RF), XGBoost (XGB), and AdaBoost are applied in performance comparison. Through a complex assessment process, the performance of these supervised machine learning algorithms for predicting employee advancement is analyzed using assessment metrics on the employees' evaluation dataset. The Artificial Neural Network (ANN) and AdaBoost model provide better results on this dataset than all traditional machine learning techniques. Finally, Our proposed modified AdaBoost approach outperformed all other methods evaluated with an accuracy of 95.30%.**

***Index Terms*—AdaBoost, ANN, Employee Attrition, Promotion Prediction.**

I. Introduction

EMPLOYEE promotion is a tedious task. One organization is not just a single entity, but it has many departments, especially in large organizations. Different sections are responsible for different tasks like, the Human Resources department looks after managerial tasks, the IT department performs the different computer-related tasks, the Accounting unit of a company calculates profit and loss, the Marketing department advertises a product and circulates to make a product familiar, Research and Development (R&D) section tries to invent a new product or to make a product superior and beneficial over other products of a different company, Production department finally produces product for launch, and so on. In any system, among these departments, the Human Resource (HR) department is one of the most important departments. The management system of an office quietly depends on this division, for example, recruiting employees, transferring employees, giving promotions to an employee, and so on [1]. Accurately assigning a promotion to an employee is very much important for the HR department as well as the employees of that company. This process does not provide just only financial benefits but also social status to the employees, and improves working motivation and vice-versa of that company. Promoting one employee position is not so easy process for the HR department. A huge number of manual paperwork is a big challenge and also required an ample time span. Because different types of employee details are stored in different datasets, especially for big companies. This promotion calculation is really a difficult job when it requires calculating among these different datasets [2]. Several factors are considered while promoting an employee. Different organizations use different factors to promote their employees. Most common key factors are:

- **Innovation (rate high)**: Innovation is one of the main factors of promotion, and sometimes it is considered the most vital point of promotion. This helps a company being a more profitable and renowned one. Every company announces rewards for its employees' innovative ideas which is encouraging. When a company provides its employees with opportunities to improve their creativity and innovation [3] for the benefit of the firm, it can be motivation for other employees by expanding employee knowledge and job experience.
- **Productivity or Efficiency**: It is another sign for an employee that s/he is really applicable for promotion or not. This is a way that employees compete for better performance [4] in the workplace. Employers can set different types of parameters of efficiency like; maximum production using minimum time, minimum resources, and minimum system loss. Because employers know that efficient employees are assets for their company. The company set different types of parameters to calculate employees' efficiency.
- **Experience or Seniority**: Real-life problems are sometimes much more critical than theoretical problems. These problems only could be solved by those employees who have gained knowledge through so many experiences [5]

Md Abu Jafor, Department of Computer Science and Engineering, Mawlana Bhashani Science and Technology University, majafor29@gmail.com.
Md. Anwar Hussen Wadud, Assistant Professor, Department of Computer Science and Engineering, Bangladesh University of Business Technology, mahwadud@gmail.com.
Kamruddin Nur, Professor, Department of Computer Science, American International University-Bangladesh (AIUB), kamruddin@aiub.edu.
Mohammad Motiur Rahman, Professor, Department of Computer Science and Engineering, Mawlana Bhashani Science and Technology University, mm73rahman@gmail.com.

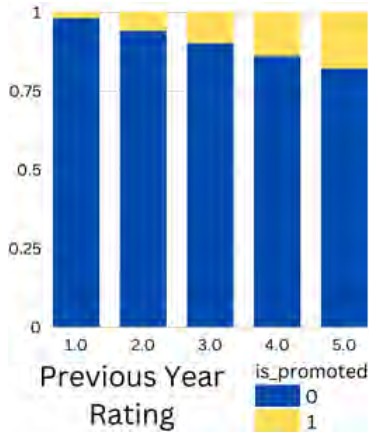

Fig. 1. Experiences in promotion prediction

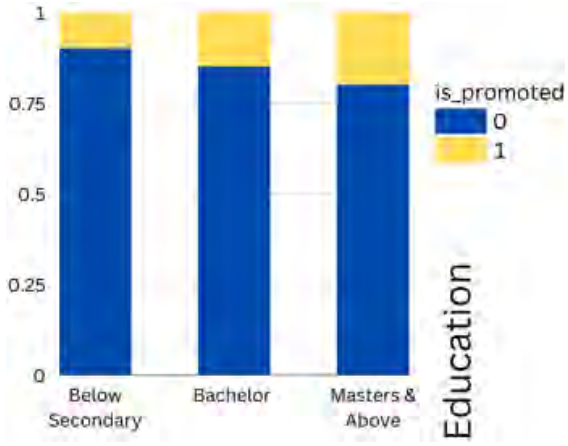

Fig. 2. Promotion distribution with educational background

as shown in Fig. 1. On the other hand, senior employee needs to be honored and given financial benefits for their long-term involvement. As a result, it can be a factor in promotion prediction for an employee.

- **Altruistic Behavior**: A good working environment depends on the people who work there, and the development of a company depends on the cooperation [6] of those employees. If all the employees have the same goal to achieve, they should be cooperative with each other. When one employee faces any obstacle to performing his task or a novice joins a team, then the experienced one should help them. Then a good working environment will be created. Otherwise, a company never achieves its goal when experienced employees do not extend their cooperative hand to others.
- **Punctuality**: Human reputation is a very important component of employee promotion. Employee punctuality, loyalty, disloyalty, and other criteria are considered in this section. These characteristics are also auto-responsible for demotion instead of promotion if the performance of an employee is not good in this section. Punctuality [7] is a habitual fact. It is developed for a long time of practice.
- **Managerial Ability**: There is another factor, which is called management skills, to be considered for promotion prediction. Good management skills can bring one company from bottom to top. Sometimes a beginner or a comparatively small company can be renowned or gain success within a very short time through a good management team [8]. They can produce a huge amount of products using a small number of resources. Systematic HR practices work their beneficial effects on innovation performance through the capacity in knowledge acquisition, sharing, and application. So, employers can give promotions to their employees based on their managerial capacity.
- **Educational Background**: To think in a different direction, one should know all other existing ways of problem-solving techniques. In this case, education is mandatory. It is considered that education is the key to success. Distribution of promotion among all of the employees with different educational backgrounds [7] is considerable. Here we can see from Fig. 2 that, most of the promoted employees have a Master's degree and above, then a bachelor's, and lastly Below secondary. The highest educational qualification is the highest chance of promotion.
- **Effect of Race**: In this study, it is examined the relationships between race [9], experience, and performance of managers from white and black people. In comparison with black officials, whites were more accepted in their own organizations, getting higher ratings from their supervisors, higher pay, and a higher level of job satisfaction. Clear and unclear effects of race exist in organizations.

Research on employee promotion is being performed by different researchers and some of them used Machine Learning [10] to predict employee promotion [4], [7], [11]. These researchers worked on promotion prediction based on different features and factors. For instance, one author focused on basic features, post features, and the correlation between attributes [7]. In this paper, the author showed that some studies exhibited promotion within a company was affected by several causes, like gender [12], [13], [14], age [15], [16], education background [17], [18], job experience [5], emotional intelligence [19] and communication patterns [20], [21]. In terms of research methods, the main approaches are qualitative analysis [22], quantitative analysis [23], and their combination. Here, the author constructs some promotion features and makes forecasts with machine learning methods based on the data from a Chinese state-owned enterprise. Author of a few papers [4], [24], [25] considered some other criteria. They showed that a long-time worker could be justified for privileges by his/her contribution, devotion, loyalty, punctuation, and so on to the company that is satisfiable for promotion or not. Furthermore, the recruitment process should be effective and free from repetition [26][27]. This recruiting process is not just a single step, it requires several steps and the involvement of many existing employees. As a result, it hampers normal official activities which could be a big loss for the company. Otherwise, prudent recruitment can save a company's time and cost in an effective way.

Many analyses of promotion prediction exist in the research field b ut n one o f t hese d emonstrated a n e ffective w ay to work with them. To make a prediction process automatic and artificially integrated, there have many challenges. To evaluate different features, to ensure high data security, to ensure the validity of the employee dataset, only the Human Resource department can change the credentials, closely monitoring the group discussion to find t heir i nvention a nd problem-solving logic yielding to boost the company, and to ensure that the system should be unbiased are the key challenges of promotion prediction systems.

All of the above are very much challenging and the main barriers to taking the right decision of promotion prediction for employees. Researchers are trying to overcome these challenges in an effective way, but all of the solutions are not being existed in a system.

In this work, one system is being tried to discover that will help us by making the system user-friendly, robust, and perfect for promotion calculation. The main focus is to make a system that will be easy, time-saving, low-cost, and have better accuracy. It will reduce ample manual paperwork that required a huge time because the dataset is from different tables. Since this work is done by a machine so it is free from the claim of bias. This work has multiple steps, after the introduction, the Related Work section presents some similar research works to gather prior knowledge and the basic theory on this topic. The later section made a scratch model of the project and implementation of the architecture with a reliable algorithm. The next section is the dataset gathering and analysis part of this paper. The result and discussion part is the last section before the conclusion of this paper. From the above discussion, The main contributions of this paper are as follows:

1) Employees' evaluation dataset collected from Kaggle was imbalanced, we have used Synthetic Minority Oversampling Technique (SMOTE) technique to balance this dataset
2) We have applied different machine-learning models to predict employee promotion automatically
3) We have proposed a modified AdaBoost technique for automatic employee promotion prediction and
4) We have performed a comparison between proposed and traditional machine learning models for better analysis

The remainder of the paper is structured as follows: Section II review the background study and related works. Then, the step-by-step process is shown in the methodology in section III where the main method and proposed method are presented. The result and discussion section is in section IV. Finally, the conclusion of this paper is shown in section V with further plans and applications.

II. Related Works

The main barrier to the development of an organization is employee turnover[1] which hinders productivity and long-term strategies. In this paper, data from Human Resource Information Systems (HRIS) was used, and checked Extreme Gradient Boosting (XGBoost) [28], [29] technique with six supervised classifiers, where XGBoost was shown as a superior algorithm based on accuracy level, low runtimes, and memory efficiency for predicting employee turnover.

Two steps to complete the prediction of an employee; building core interpersonal features and post features[7], and correlation analysis between attributes and promotion. Furthermore, Random Forest and GINI models were applied for two different purposes. The former model was applied to verify the validity of attributes and the latter one was to analyze the influence of each feature of promotion. In conclusion, it said that the Random Forest model was better than other relative models.

Experienced employees were very much important for any company. But if these employees left that company, it appeared as a very destructive matter. Leaving a company happened for different reasons like dissatisfaction, low income, social status, and so on. The management system should take proper steps to preserve the talented people [30] of that company. This paper showed a machine learning approach that can predict employee attrition by considering three main experiments. First of all, involving training the original class-imbalanced dataset with SVM, KNN, and Random Forest methods. The second try-out used the Adaptive Synthetic approach to overcome the class-imbalanced dataset, then training again on the novel dataset using previously mentioned methods. The last experiment was manual under-sampling of the data to balance between classes.

To take care of and give satisfaction to the employee requires proper nourishment. Intelligent HRM [4] is mandatory for intelligent production. The author of this paper used different big data and artificial intelligence technologies in the multiple features of the organizational position. Here these professional features were tested where versatility is long-lasting, imagination is plentiful, or practical knowledge is visible. A smart HRM department can examine an employee's above professional features. In conclusion, the author concluded that different positions in a company are good to boost the organizational dedication and job satisfaction of an employee.

One of the profitable things of a business organization is not performing the same task repeatedly like publishing a circular and then recruiting employees again and again when workers leave their jobs. The recruitment cost will be a complete loss when candidates do not commence the job after completing this entire procedure. Binary classification [26], [31] of the machine learning model is used to predict desired candidates who will not stay away from joining or backing out immediately after the recruitment process. This model considers some germane properties like physical and mental caliber.

An employee becomes experienced over time. Enterprises usually fight for talented and experienced employees. Alongside, they circulate various lucrative offers to attract talented and expert employees. As a result, employees leave their current organization and join somewhere else for a better life. If the HR department cannot identify the causes and inform the authority accurately and timely, it can cost huge business losses. To solve this problem and to predict employee resignation [32], the Logistic Regression model was used here. Two functions of this model were used for two different

purposes; the Cross-Entropy method as the objective function, and Newton's method and Regularization to optimize the model.

The complete growth of every company is depending on the performance of its employees. The contributions of these employees are assessed by some internal and external attributes like targets, and achievements. The author of this paper proposed a framework called EPP (Employee Performance Predictor) to evaluate the performance of each individual employee. The attributes of this predictor [24][4][25] are determined by data mining based on the employees' historical performances. Better performance was assessed the more priority for promotion. Those employees were selected for promotion who were considered as the highest priority.

SVM, K-Nearest Neighbor, Naïve Bayes, Decision Tree, AdaBoost, and Random Forest Classifiers were compared with the proposed feed-forward neural network along with a 10-fold cross-validation [33], [34], [35] procedure. This procedure was under a single platform for forecasting employee attrition. On the basis of performance measurement metrics, this proposed method provided better results than the above models. A fair system is very important for selecting appropriate employee for promotion. Using Data Mining techniques employee performance was predicted, and the knowledge flow model [25] of the Open-Source tool was used to measure the elements in this paper. When human beings take the responsibility to find out an employee among other employees, it could be an unfair and biased system. It was shown in this paper how a system determines a fair selection of employees for promotion.

The first and foremost step of a company's development is the process of employee recruitment. This paper discussed how a company can recruit potential employees accurately by accessing historical data. Data Mining and Fuzzy Logic [27] approaches to do this task more easily and accurately. Here, Data mining was used to classify attributes of latent employees, and then Fuzzy logic is used to order employees who will be recruited. Again, total investment could be wasted if the recruited person does not join his job. That's why we need to know the joining efficiency [26] of the employee before resume selection. Total process was to be done without any kind of risks.

A company can be developed, and earn maximum benefit from employees if they are physically fit. In this research, a survey [36] was presented which indicated the promotion of employees' health. This survey requires three steps such as 5-level architecture depicted to show the data flow, an organized study to evaluate the current studies, and finally, constraints and emerging perspectives were offered. Furthermore, this survey exhibits that both individual and group-wise addressing was important to form a smart office.

The open-source tool Weka [34] was used here to perform a comparative study on employee attrition prediction. A comparison of classification and clustering algorithms was executed based on performance. The criteria of performance were accuracy, precision, recall, F1 score, and time taken to build the model. Among all comparisons, the Naïve Bayes method showed the highest result of attrition prediction.

From the above discussions, we noticed that all papers performed the result analysis of employees' either attrition or promotion. These papers were based on some specific factors, for instance, some papers worked on HRM, some papers worked on core interpersonal features and post features, some papers worked on an employee's performance, some papers worked on the steps of employee recruitment, some papers illustrated a survey on health, and so on. The results of these papers were not so applicable to all. They proposed a model for a fixed company. In this paper, a model has been proposed which can predict employees' promotion successfully for all types of companies. Here, we showed a comparison of results among all machine learning algorithms.

III. Methodology

We have used several subsections to discuss data collection, preprocessing, and feature engineering processes. We then discuss some baseline models used to predict employee promotions. Finally, we describe our proposed model in detail. The graphical representation of the proposed model is presented in Fig. 3.

A. Dataset

Employees Evaluation for Promotion [11] dataset from the Kaggle was used in this research to perform our method and examine the dataset. For the purpose of getting necessary penetrations, the dataset of promotion and the factors of promotion of the employees were validated by a dataset analyzer. We extract major and minor features by applying a feature extraction technique to derive best-fit parameters to achieve our prediction purposes. When we analyze the dataset, we encounter that it was not properly balanced. A resampling technique was used to balance the dataset. Then a previously processed dataset was ready for building a model. We divided the whole dataset in a ratio of 3:1, where 75% of the total data was used to train the proposed Machine Learning model and the rest 25% of the total data was used to test the proposed Machine Learning model. This Machine Learning technique was adjusted by different parameters and finally, the proposed method was ready in a more generalized form to predict employee promotion prediction by putting the employee details.

B. Collecting Raw dataset

There was large and small-scale dataset available on the internet, but we used a large-scale dataset of Kaggle. The employees' Evaluation for Promotion dataset was used in our research work which was developed by the data scientists of that company. It contains huge types of employee features like role in the company, innovation, loyalty, punctuation, experience, and so on. These features were processed and examined yielding to employee promotion. There were 54808 entries of different employees with their 13 attributes in this dataset, which required 3.64 MB of memory.

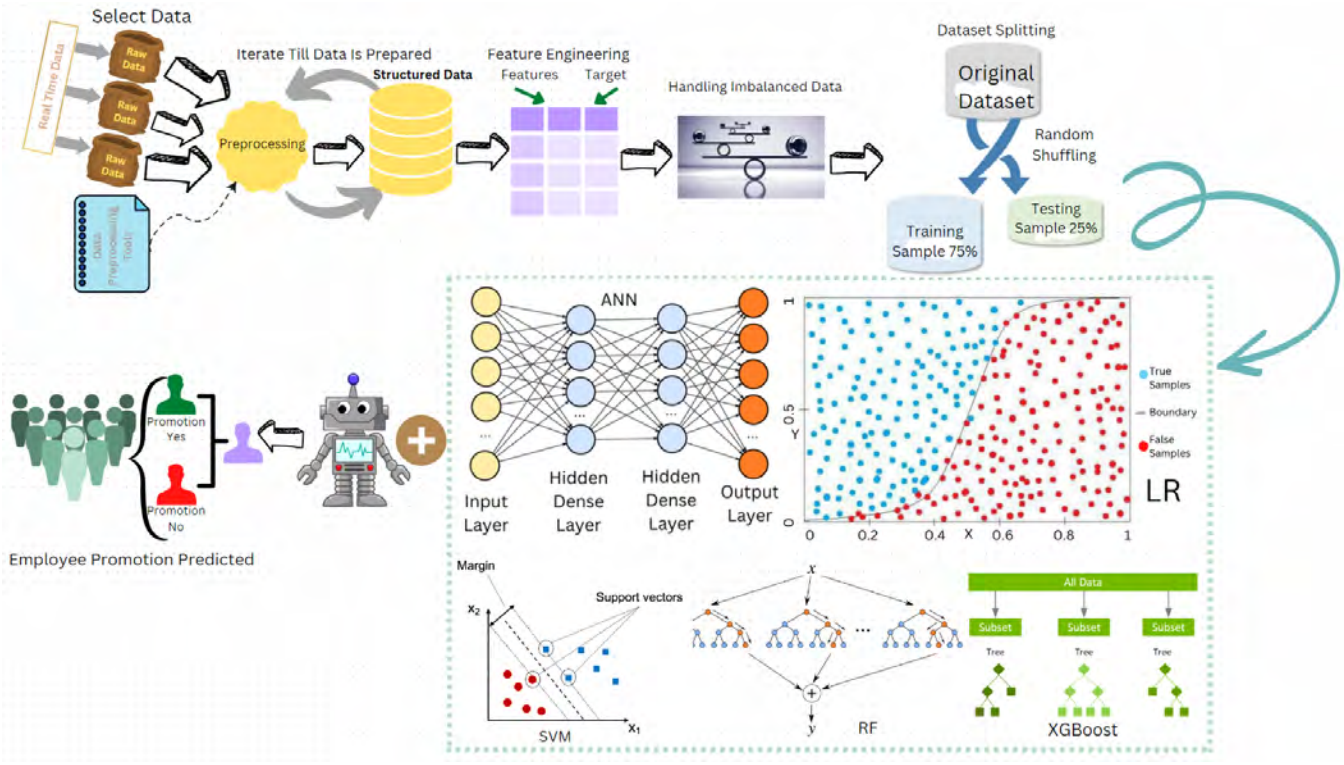

Fig. 3. Graphical view of the proposed model

C. Preprocessing

Among all attributes of the dataset, only one target variable exited, i.e., the is_promoted attribute. This attribute indicates whether one employee is given a promotion or not. In this algorithm, after considering all of the attributes the specific attribute shows the targeted output. Our proposed algorithm is trained to give accurate predictions based on other variables of a particular observation. The value of the target variable can have either 1 or 0. The value 1 indicates that the employee is promoted and 0 for not promoted. For better visualization and understanding of the data of the target variable 1 and 0 are represented as YES and NO respectively.

D. Balancing dataset

In the raw dataset, all records were not unlike and have some sort of similarity. So, to make the dataset more useful we used a data re-examining technique [37] called Synthetic Minority Oversampling Technique (SMOTE). This approach made the dataset balanced from the unbalancing dataset. After developing a model, it requires training and testing with real-life data. This process can eliminate the complexity and increases the reliability of any proposed method. For this purpose, we divided our dataset into two parts with a ratio of 3:1. For training and testing purposes we assigned 75% and 25% respectively of the whole dataset. In addition, another reason for dataset splitting was to prepare our model more generalize.

E. Feature Engineering

We needed a tool called FAT (Feature Analysis Tool), and when it was applied to the employee dataset, we obtained useful insights. The FAT was used to carefully scrutinize the features and factors which were the main things of employee promotion. We scrutinized the features through different types of comparative study analysis in Fig. 4. The FAT illustrated the pattern of data and also it proved that it was helpful for the analysis of data features in the case of employee promotion.

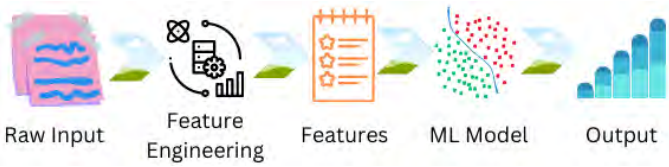

Fig. 4. Feature Engineering

F. Extracting Major and Minor Features

Extracting features from the dataset was a vital point of our learning method. Exactly finding major and minor features was a real challenge. We never get a good result, until we stay free from an amalgam of major and minor features. So, applying appropriate feature-extracting techniques is really a crucial part to find the best-fit result of features that will be applied for the learning method. To handle features of the above-mentioned company dataset, major and minor feature-extracting techniques have been applied in our research. For analyzing the correlation of the features, we used the features like Innovation, Experience, Punctuation, Managerial Ability,

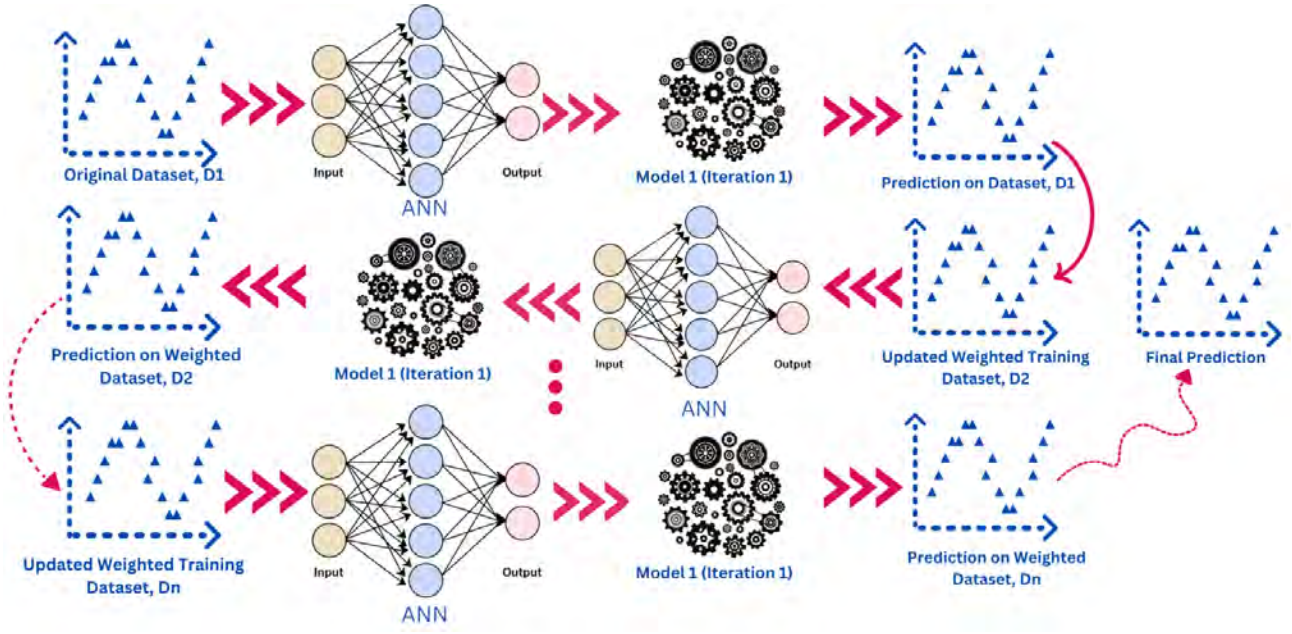

Fig. 5. Prediction using modified AdaBoost with ANN classifier

Workplace Assistance, Dependency, Loyalty, Punctuation, and so on. These features are more functional than other minor features. These identified features were encoded to obtain high accuracy and effective result using an encoding technique. Finally, we found these feature-extracting techniques were productive for our research study.

G. Proposed Method:

At first, we used some baseline classifiers like Support Vector Machine (SVM), Logistic Regression (LR), Artificial Neural Network (ANN), Random Forest (RF), and XGBoost (XGB) to automatically predict employee promotion. Then we proposed a model combining AdaBoost and ANN classifiers to obtain high accuracy. AdaBoost, also known as Adaptive Boosting, is a boosting algorithm that is widely used to refine imbalanced data and used as an Ensemble Method in Machine Learning. It builds a model and gives equal weights to all the data points. It then assigns higher weights to points that were incorrectly classified. Now all the points with higher weights are given more priority in the next model. These training models will keep continuing until and unless a lower error is accepted as true or valid. In this paper, AdaBoost Fig. 5 illustrates the iterations of using the AdaBoost classifier for prediction. At the very beginning, a raw dataset D1 was taken which was used as input in the first iteration of the model. The model returned a prediction on dataset D1. After updating weights to the dataset we found training dataset D2. It was then used as input in the second iteration of the model and got a second prediction on weighted dataset D2. Then its weights were updated and used as input in the next iteration of the model. These steps would be continued until we got the desired result. This result will be the final prediction of the model. Normally, AdaBoost uses Decision Tree as the default classification model in each iteration, but in this paper, we modified the default classification with ANN and performed better accuracy than the default AdaBoost algorithm.

IV. Result and Discussion:

A. Experimental setup

In this section, we need to setup an environment where our proposed method will be run and tested. All the experiments have run on a machine with specifications of Intel (R) Xeon (R) CPU, 13 GB RAM, 2249.998 MHz CPU, 512KB cache size, and the CPU model name is AMD EPYC 7B12. For our proposed model we used base_estimator=SVM, n_estimators=50, learning_rate=1, and random_state=0.

B. Evaluation metrics

Employee promotion prediction were carefully scrutinized. In our machine learning-based research study, the evaluation metrics holds the training and testing accuracy, precision score, recall score, and f1 score. This metrics is acquired using confusion metrics. The followings are the key components of evaluation metrics:

True Positive (TP): when positive is the value for both of the forecasted values and real values;

True Negative (TN): when negative is the value for both of the forecasted values and real values;

False Positive (FP): when the approach predicts a positive value, but the real values are negative;

False Negative (FN): when the approach predicts a negative value, but the real values are positive.

The accuracy scores were measured when the training and testing process is performed. We found the 90% of accuracy on real-time data, then the model was generalized. The formula of calculating the accuracy score is given below:

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN}$$

The following equation depicts the precision calculation. It is the measurement of the model that correctly identifies positive values among all values.

$$Precision = \frac{TP}{TP + FP}$$

The measurement of accurately identifying true positive values is called recall, that is calculated using the equation below:

$$Recall = \frac{TP}{TP + FN}$$

The F1 score illustrates the efficiency and effectiveness of this model which is shown in Fig. 6. It combines the recall and precision score values.

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall}$$

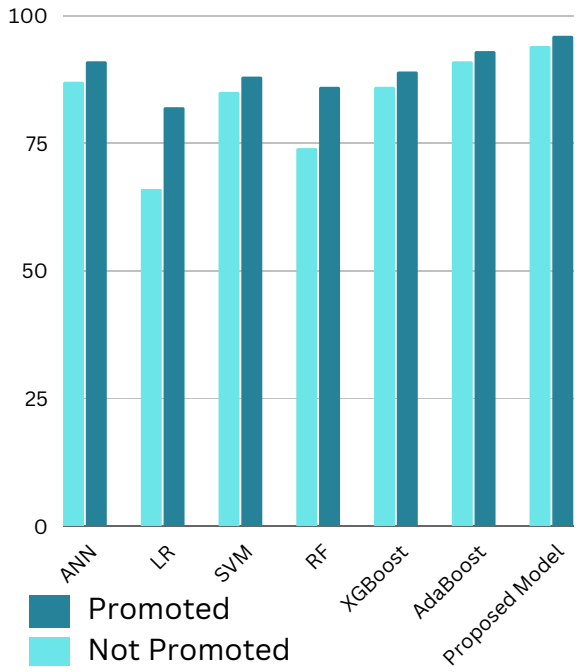

Fig. 6. F1-score bar graph of different machine learning models

C. Result Analysis

Table-I depicts a comparison of Accuracy, Precision, Recall, and F1-score among different traditional machine-learning models. Our proposed method is also included in comparison with these traditional models. From the table, it is shown that AdaBoost model has the highest accuracy in comparison to the rest of the models but less than our Proposed_model which has 95.20% accuracy. The Logistic Regression is the lowest accuracy 76.65%. The accuracy rate is 89.37% and 92.16% respectively for ANN and AdaBoost. In addition, the precision rate of ANN and SVM is almost the same. But, like the accuracy rate, LR is also the lowest in the precision value among all models. Moreover, the Random Forest model

TABLE I
PERFOMANCE ANALYSIS OF DIFFERENT ML MODELS

Model Name	Acc(%)	Pre(%)	Rec(%)	F1(%)
ANN	89.37	84.08	98.93	90.90
LR	76.65	70.59	96.79	81.64
SVM	87.08	84.40	93.13	88.55
RF	81.15	74.00	99.00	85.06
XGBoost	87.44	83.66	95.16	89.04
AdaBoost	92.16	87.96	98.93	93.12
Proposed_Model	**95.30**	**92.79**	**98.93**	**95.76**

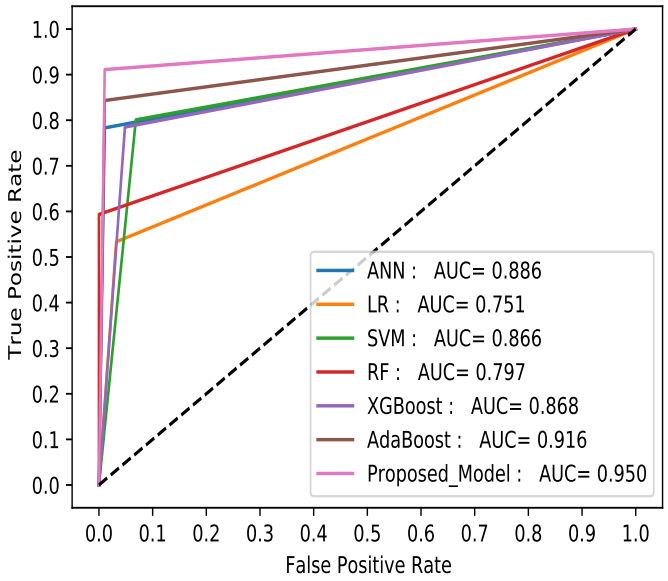

Fig. 7. ROC curve of different machine learning models

shows the value of recall, which is 99.00%. The XGBoost model is the second highest in the F1 score, and the value is 89.04% whereas 95.76% is our Proposed_model as shown in Fig. 6. To surmise Table-I, Proposed_model shows the highest value in all categories except the recall value. Instead of the highest values, the LR model is the lowest in all categories except the recall value. A ROC curve stands for the Receiver Operating Characteristic curve. It is a graph that shows the performance of a classification model at all classification thresholds. This curve plots two parameters: True Positive Rate and False Positive Rate. Fig. 7 shows the ROC curve of different traditional classification models like ANN, LR, SVM, RF, XGBoost, AdaBoost, and Proposed_model. In addition, there are some existing works available where they work on the same dataset that we used. Like, the author [3] applied only three traditional methods on the same dataset, and there is no modified method in their work. But, we applied six methods alongside our modified proposed method. Luckily, our proposed method demonstrates the best result over the above-mentioned paper [3], the result is shown in Table I.

V. CONCLUSION

During this research work, it has been identified that the prediction of employee promotion is a very important aspect of the human-capital building process. The employee promotion prediction by using the five popular advanced machine

learning techniques ANN, LR, SVM, RF, and XGBoost were applied in comparison in this study. Training datasets of employees were created using real data of the company and further processed. A modified A daBoost c lassifier wa s used here for prediction and later popular supervised machine learning algorithms were studied and tested on performance using numerical metrics. The applied machine learning techniques achieved accuracy scores of 89.37 by ANN, 87.08 by LR, 76.61 by SVM, 81.15 by RF, and 87.44 by XGBoost. In addition, F1 score of the ANN is also greater than all of the other models. Though the Precision and Recall of the ANN model are not the highest percentages among all of the models, it is almost the same as that model which has the highest ratings. From these scores, it is clear that the Artificial Neural Network (ANN) provides better results on this dataset. Organizations will be beneficiary f rom o ur r esearch findings to overcome employee attrition by giving promotions to the employees.

REFERENCES

[1] P. Ajit, "Prediction of employee turnover in organizations using machine learning algorithms," *algorithms*, vol. 4, no. 5, p. C5, 2016.

[2] S. Hosen, J. F. Eva, A. Hasib, A. K. Saha, M. Mridha, and A. H. Wadud, "Hqa-data: A historical question answer generation dataset from previous multi perspective conversation," *Data in Brief*, vol. 48, p. 109245, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352340923003645

[3] K. Şahinbaş, "Employee promotion prediction by using machine learning algorithms for imbalanced dataset," *International Conference on Computing and Machine Intelligence (ICMI), IEEE Xplore, DOI: 10.1109/ICMI55296.2022.9873744*, 2022.

[4] J. Liu, T. Wang, J. Li, J. Huang, F. Yao, and R. He, "A data-driven analysis of employee promotion: the role of the position of organization," in *2019 IEEE international conference on systems, man and cybernetics (SMC)*. IEEE, 2019, pp. 4056–4062.

[5] V. V. A. E. B. M. N. De Pater, I. E. and KLEHE, "Employees' challenging job experiences and supervisors' evaluations of promotability," *Personnel Psychology. 62, 2 (Summer 2009), 297-325.DOI= https://doi.org/10.1111/j.1744-6570.2009.01139.x*, 2009.

[6] R. B. E. F. Herbert Gintis, Samuel Bowles, "Explaining altruistic behavior in humans," *Evolution and Human Behavior, 24(3), 153–172. doi:10.1016/s1090-5138(02)00157-5*, 2002.

[7] Y. Long, J. Liu, M. Fang, T. Wang, and W. Jiang, "Prediction of employee promotion based on personal basic features and post features," in *Proceedings of the International Conference on Data Processing and Applications*, 2018, pp. 5–10.

[8] J.-W. H. Chung-Jen Chen, "Strategic human resource practices and innovation performance — the mediating role of knowledge management capacity," *Journal of Business Research, 62(1), 104–114. doi:10.1016/j.jbusres.2007.11.016*, 2007.

[9] S. P. Jeffrey H. Greenhaus and W. M. Wormley, "Effects of race on organizational experiences, job performance evaluations, and career outcomes," *Academy of Management Journal, 33(1), 64–86. doi:10.5465/256352*, 1990.

[10] M. A. H. Wadud, M. Mridha, and M. M. Rahman, "Word embedding methods for word representation in deep learning for natural language processing," *Iraqi Journal of Science*, pp. 1349–1361, 2022.

[11] Kaggle, "Employees evaluation for promotion," *Available online: https://www.kaggle.com/datasets/muhammadimran112233/employees-evaluation-for-promotion*, accessed on 01 March 2023.

[12] F. D. Blau and DeVaro, "New evidence on gender differences in promotion rates: An empirical analysis of a sample of new hires," *Industrial Relations: A Journal of Economy and Society. 46, 3 (July. 2007), 511-550. DOI= https://doi.org/10.1111 j.1468-232x.2007.00479.x.*, July, 2007.

[13] Song, "Does gender make a difference?-career mobility in urban china," *China Economic Quarterly. 6, 2 (Jan. 2007), 629-654. DOI= https://doi.org/10.13821/j.cnki.ceq.2007.02.007*, 2007.

[14] P. K. L. Roth, P. L. and Bobko, "A metaanalysis of gender group differences for measures of job performance in field studies," *Journal of Management. 38, 2 (Mar. 2012), 719-739. DOI= https://doi.org/10.1177/0149206310374774*, 2012.

[15] C. S. Machado and Portela, "Age and opportunities for promotion," *IZA Discussion Paper No.7784*, 2013.

[16] S. J. Adams, "Passed over for promotion because of age: an empirical analysis of the consequences," *Journal of Labor Research. 23, 3 (Sep. 2002), 447-461. DOI= https://doi.org/10.1007/s12122-002-1046-y*, 2002.

[17] S. Spilerman and Lunde, "Features of educational attainment and job promotion prospects," *American Journal of Sociology. 97, 3 (Nov. 1991), 689-720. DOI= https://doi.org/10.1086/229817*, 1991.

[18] M. L. Bognanno and Melero, "Promotion signals, experience, and education," *Journal of Economics Management Strategy. 25, 1 (Spring 2016), 111-132. DOI= https://doi.org/10.1111/jems.12132*, 2016.

[19] J. B. Kulkarni, P. M. and Kumar, "Emotional intelligence and employee performance as an indicator for promotion, a study of automobile industry in the city of belgaum, karnataka, india," *International Journal of Business and Management. 4, 4 (Apr. 2009), DOI= https://doi.org/10.5539/ijbm.v4n4p161*, 2009.

[20] C. C. F. P. A. H. N. Woolley, A. W. and Malone, "Evidence for a collective intelligence factor in the performance of human groups," *Science. 330, 6004 (Oct. 2010), 686-688. DOI= https://doi.org/10.1126/science.1193147*, 2010.

[21] Pentland, "The new science of building great teams," *Harvard Business Review. 90, 4 (Apr. 2012), 60-69*, 2012.

[22] H. G. Ridder and Hoon, "Introduction to the special issue: qualitative methods in research on human resource management," *German Journa of Human Resource Management: Zeitschrift für Personalforschung. 23, 2 (May.2009), 93-106. DOI= https://doi.org/10.117/239700220902300201*, 2009.

[23] C. J. A. Sanders, K. and Bainbridge, "Research methods for human resource management," *Routledge*, 2014.

[24] M. Karthika and T. Meyyappan, "Employee performance prediction using epp framework," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, vol. 6, no. 6, pp. 297–302, 2020.

[25] T. Raut, P. Kale, R. Sonkusare, and A. K. Gaikwad, "Employee performance prediction system using data mining," *International Research J. of Engineering and Technology (IRJET)*, vol. 7, 2020.

[26] D. J. M. Reddy, S. Regella, and S. R. Seelam, "Recruitment prediction using machine learning," in *2020 5th International Conference on Computing, Communication and Security (ICCCS)*. IEEE, 2020, pp. 1–4.

[27] C. E. A. Pah and D. N. Utama, "Decision support model for employee selection based on data mining and fuzzy logic," 2020.

[28] M. F. Mridha, M. A. H. Wadud, M. A. Hamid, M. M. Monowar, M. Abdullah-Al-Wadud, and A. Alamri, "L-boost: Identifying offensive texts from social media post in bengali," *IEEE Access*, vol. 9, pp. 164 681–164 699, 2021.

[29] M. A. H. Wadud, M. M. Kabir, M. Mridha, M. A. Ali, M. A. Hamid, and M. M. Monowar, "How can we manage offensive text in social media-a text classification approach using lstm-boost," *International Journal of Information Management Data Insights*, vol. 2, no. 2, p. 100095, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2667096822000386

[30] S. S. Alduayj and K. Rajpoot, "Predicting employee attrition using machine learning," in *2018 international conference on innovations in information technology (iit)*. IEEE, 2018, pp. 93–98.

[31] A. J. Keya, M. A. H. Wadud, M. Mridha, M. Alatiyyah, and M. A. Hamid, "Augfake-bert: Handling imbalance through augmentation of fake news using bert to enhance the performance of fake news classification," *Applied Sciences*, vol. 12, no. 17, p. 8398, 2022. [Online]. Available: https://www.mdpi.com/2076-3417/12/17/8398

[32] W. Dai and Z. Zhu, "Employee resignation prediction model based on machine learning," in *International Conference on Applications and Techniques in Cyber Security and Intelligence*. Springer, 2020, pp. 367–374.

[33] M. A. H. Wadud, M. Alatiyyah, and M. Mridha, "Non-autoregressive end-to-end neural modeling for automatic pronunciation error detection," *Applied Sciences*, vol. 13, no. 1, p. 109, 2022. [Online]. Available: https://www.mdpi.com/2076-3417/13/1/109

[34] P. Usha and N. Balaji, "A comparative study on machine learning algorithms for employee attrition prediction," in *IOP Conference Series: Materials Science and Engineering*, vol. 1085, no. 1. IOP Publishing, 2021, p. 012029.

[35] M. A. H. Wadud and M. R. H. Rakib, "Text coherence analysis based on misspelling oblivious word embeddings and deep neural network," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 1, 2021. [Online]. Available: http://dx.doi.org/10.14569/IJACSA.2021.0120124

[36] X. Zhang, P. Zheng, T. Peng, Q. He, C. K. Lee, and R. Tang, "Promoting employee health in smart office: A survey," *Advanced Engineering Informatics*, vol. 51, p. 101518, 2022.

[37] S. Krishna and S. Sidharth, "Hr analytics: Employee attrition analysis using random forest." *International Journal of Performability Engineering*, vol. 18, no. 4, 2022.

MD. ANWAR HUSSEN WADUD is an Assistant Professor in the Department of Computer Science and Engineering, Bangladesh University of Business and Technology, Dhaka, Bangladesh. He received his B.Sc. and M.Sc. Engineering degree in CSE from Mawlana Bhashani Science and Technology University, Tangail, Bangladesh. He participated in several ACM ICPC programming contests during his university life. He worked on several programming platforms such as Java Spring & Hibernate, Android app developments, Python NumPy, Keras, etc. for big data and deep learning analysis in several software companies. His area of interest is Big Data Analysis, Deep Learning, Natural Language Processing, Internet of Things, and Machine Learning.

KAMRUDDIN NUR (Senior Member, IEEE) is currently serving as a professor in the Department of Computer Science at American International University-Bangladesh (AIUB). He also served as the chairman in the Department of Computer Science and Engineering at Stamford University Bangladesh (SUB) and Bangladesh University of Business and Technology (BUBT). Dr. Nur completed his PhD from UPF, Barcelona, Spain, Masters from UIU, and Bachelor from Victoria University of Wellington (VUW), New Zealand. Dr. Nur authored many prestigious journals and conferences in IEEE and ACM, served as TPC members, and reviewed articles in IEEE, ACM, Springer journals, and conferences. His research area includes Ubiquitous Computing, Computer Vision, Machine Learning, and Robotic Automation.

Md Abu Jafor received his B.Sc. Engineering degree in Computer Science and Engineering from Mawlana Bhashani Science and Technology University, Tangail, Bangladesh. He has good programming knowledge and skills. He participated in several ACM ICPC programming contests during his undergraduate program. His research interest includes Machine Learning, Network Security, Big Data Manipulation, and Data Mining. He also published a paper on the PageRank Algorithm which was related to Big Data and Data Mining. Alongside his academic study, he also developed different Android apps and completed two projects on Ludu Game and Lan Communication using Java.

MOHAMMAD MOTIUR RAHMAN is a professor in the department of Computer Science and Engineering, Mawlana Bhashani Science and Technology University, Tangail, Bangladesh. He was the department head for two times. He Completed his Ph.D degree from Jahangirnagar University, Bangladesh. His research interests include digital image processing, medical image processing, computer vision and digital electronics. He has many international journal and conference publications.