Journal of Global Research in Electronics and Communication
Volume 1, No. 11, November 2025
Available Online at: www.jgrec.info
RESEARCH PAPER

Human Resource Analytics for Enhancing Employee Performance through Machine Learning Techniques

Dr. Amit Jain
Professor
Department of Computer Science and Engineering
OP Jindal University
Raigarh (C.G)
amit.jain@opju.ac.in

***Abstract*—Human Resources (HR) onboarding is a critical step in integrating new employees into an organization. Effective onboarding can improve employee satisfaction, reduce turnover, and accelerate productivity. However, traditional methods for onboarding and predicting employee success are often limited to individual-level attributes, neglecting the complex network of relationships and team dynamics that influence an employee's experience. This study focuses on predicting employee performance using machine learning and deep learning techniques, with a particular emphasis on effective feature selection and preprocessing. The Employee Dataset, consisting of diverse demographic and organizational attributes, was preprocessed through normalization, outlier removal, and dimensionality reduction using PCA to ensure improved data quality and reduced redundancy. Deep Neural Network (DNN) model, was developed and evaluated. The results demonstrate that the proposed DNN model achieved a high accuracy of 96.25%, along with a precision of 98%, a recall of 97%, and an F1-score of 97%. These findings highlight the superior capability of the DNN in capturing complex patterns within employee data, making it a highly effective approach for enhancing prediction accuracy and supporting data-driven decision-making in employee performance management**.

Keywords—HR Analytics, Artificial Intelligence, Human Resource Management, Predictive Analytics, Machine Learning, Deep Learning.

I. Introduction

Human resources (HR) is the organization of a company that is in charge of analysis, recruiting, monitoring, and training job candidates, and managing employee benefit programs. It refers to both the people who are working for a enterprise or an organization and the division in charge of dealing with all employee-related issues [1][2]. Employees are among the most important resources in any organization or company [3]. HR is critical in enabling businesses to engage with a business in a socially responsible business atmosphere and a higher improvement in the quality of workers in the twenty-first decade [4][5].

The essential role of Performance Management within Human Resource Management (HRM) is to evaluate, guide, and improve employee contributions toward Organizational goals through assessments [6]. Traditional performance management systems encounter multiple difficulties when staff members conduct appraisals due to subjective practices that yield inconsistent results and biased judgments [7]. The issues with performance management systems impact both employee trust and commitment, and limit correct skill development and Organizational decisions. Numerous Organizations experience increasing pressure to create performance evaluation approaches that maintain fairness and improve both efficiency and scalability [8].

The growing issues have necessitated the use of Artificial Intelligence (AI) and Machine Learning (ML) in Performance Management, which has greatly impacted the HR technological landscape [9]. The rapid advancements in artificial intelligence (AI) and machine learning (ML) have significantly transformed various business functions, including human resource management (HRM) [10][11][12]. Traditional HR analytics primarily focused on data-driven insights to improve decision-making. However, with the integration of AI-driven tools, HRM has evolved into a more proactive, efficient, and strategic function that enhances workforce productivity and engagement [13]. machine learning algorithms that improve Human Resource performance management through unbiased data-driven decision-making [14]. ML models for human resource management, along with deployment implementation elements, while offering flexible implementation guidelines for real-world applications [15].

A. Motivation and Contribution

The motivation behind this study lies in the growing need for organizations to accurately evaluate and enhance employee performance in order to remain competitive in dynamic business environments. Traditional performance assessment methods often rely on subjective judgments, which can lead to inconsistencies and bias, whereas data-driven approaches offer more objective, reliable, and scalable solutions. By leveraging machine learning and deep learning models, particularly the Deep Neural Network (DNN), organizations can uncover hidden patterns in employee data, predict performance outcomes more effectively, and make informed decisions regarding workforce management. This not only improves productivity and efficiency but also supports fair and transparent evaluation processes that contribute to overall organizational growth. This research makes several key contributions to the field of Employee Performance:

- Used an Employee Dataset containing diverse demographic and organizational attributes of both active and resigned employees.
- Implemented data cleaning, handling of missing values, outlier removal, encoding, normalization, and duplicate elimination to ensure high-quality data.
- Applied Principal Component Analysis (PCA) to reduce dimensionality, remove redundancy, and retain the most significant features for better model performance.

- Provided a reliable framework for organizations to predict and analyze employee performance, supporting data-driven decision-making for workforce optimization.
- Demonstrated that the DNN model achieved the highest accuracy, outperforming baseline models.
- Evaluated model accuracy using multiple metrics (accuracy, precision, recall, ROC, F1-score) for thorough performance assessment.

B. *Justification and Novelty*

The justification for this study stems from the limitations of conventional employee performance evaluation methods, which often fail to capture the complex and nonlinear relationships among multiple employee attributes. Unlike traditional approaches, the proposed framework integrates advanced preprocessing techniques such as normalization, outlier removal, and Principal Component Analysis (PCA) for feature selection, ensuring cleaner and more representative input data. The novelty of this work lies in the application of a Deep Neural Network (DNN), which significantly outperforms classical models by achieving higher accuracy and robust predictive capability. By combining feature selection with deep learning, this study provides a unique and efficient methodology that enhances prediction reliability, reduces redundancy, and delivers actionable insights for improving workforce performance management.

II. Literature Review

Several significant research studies on enhancing employee performance have been reviewed and analyzed to provide a strong foundation and guidance for the development of this work.

Goyal et al. (2025) Real-world data evaluation confirmed that the Hybrid SSL + RL model delivers superior performance to SVM, along with SSL-only and RL-only models. Acceptance for the Hybrid model reached 92.7%, while the SSL model achieved 84.3% accuracy, and the RL model scored 90.2% accuracy. Within this study, the Hybrid model showed the maximum precision levels (91.8%) and recall (93.0%), thus indicating its ability to make precise productive predictions. During 40 iterations, the RL agent substantially enhanced its performance until it reached an average reward rate of 0.95, indicating its enhanced ability to improve productivity strategies over time. The Hybrid model displayed adaptability across different tasks at 94.2%, whereas the baseline stood at 85.0%, SSL at 90.5%, and RL at 92.1% [16].

Tanmayi et al. (2025) aim of this research is to examine how various factors affect employee attrition, to understand why workers leave the company and predict future departures. In the proposed work, machine learning, ensemble models and deep learning are used to predict employee attrition accurately so that it helps the company to improve their strategies to satisfy their employees. A dataset from Kaggle is used to train and evaluate these models. Many models like CatBoost, AdaBoost, Random Forest, SVM, DNN, Ensemble model of CatBoost and AdaBoost, and Ensemble model of hard Voting Classifier are used. Random forest gave the best results out of all the models tested with an F-score of 93.1 percent and an accuracy of 93.5 percent [17].

Kosuri et al. (2025) proposes an AI-based sentiment analysis model using the GRU-Attention mechanism to enhance employee satisfaction. Data collected from Kaggle includes employee feedback from various organizations, which undergoes preprocessing using techniques such as tokenization, stopword removal, and lemmatization. Text features are extracted using the TF-IDF method, which converts textual data into numeric form. The GRU-Attention model leverages the strengths of gated recurrent units (GRUs) and attention mechanisms to focus on key phrases in employee feedback for sentiment classification into positive, negative, or neutral categories. Results demonstrate that the proposed model achieves higher accuracy (91.8%) compared to baseline methods like Random Forest (82.1%) and GRU without Attention (88.5%) [18].

Prakash et al. (2024) came up with an AI model that incorporates machine learning algorithms with corporate culture measures to estimate employees' motivation levels. The model relies on the IBM HR Analytics dataset that focuses on employee behavior and feedback data, for performance records. Approaches used herein are Principal Component Analysis for feature extraction and XGBoost and Logistic Regression for modeling. The proposed model obtained an accuracy of 90.5%, 89.2% precision, and an AUC of 0.92, surpassing typical models. Also, the system analyzed different employees and labeled them as high-risk employees with an accuracy level of 87.5%. The findings also show how AI can boost employee motivation by offering precise information on their requirements [19].

Subha et al. (2024) proposes a novel cascaded methodology combining Multi-Layer Perceptron (MLP) and Support Vector Machine (SVM) models, enhanced by deep gradient optimization techniques, to predict employee attrition. The dataset, preprocessed through feature selection and synthetic oversampling (SMOTE), included key employee features such as monthly income, job satisfaction, and overtime. The proposed methodology was tested against other models, including Random Forest (RF) and Logistic Regression (LR). The cascaded MLP+SVM model achieved a training accuracy of 94.5% and a test accuracy of 92.1%, outperforming standalone MLP (89.7%) and SVM (86.5%). The model also showed minimal overfitting, with only a 2.4 % difference between training and test accuracy. Key performance metrics, including an F1 score of 0.83 and an AUC score of 0.91, confirmed the model's ability to classify employees likely to leave accurately [20].

Priyana et al. (2024) involved detailed analysis of employee profiles and application of SMOTE techniques to harmonize data imbalance, along with fine tuning of hyperparameters in the XGBoost model to improve model performance. The results of the first experiment using the XGBoost model showed that the model produced an accuracy rate of 0.81, but after going through the hyperparameter tuning process, the accuracy increased significantly to 0.85. The results of this study provide a deeper understanding of the factors that contribute to employee retention in different types of companies. The implementation of SMOTE and hyperparameter tuning in the XGBoost model proved its effectiveness in improving prediction accuracy. This research has the potential to provide strategic guidance for human resource management in improving employee retention in the future [21].

Mascarenhas, Savant and Aswale (2023) presents a study that leverages machine learning techniques to determine employee performance using available data in organizations. The objective is to detect the key causes of attrition and minimize them through data analysis. The methodology involves data acquisition, conditioning, visualization, and classification using classification models like Random Forest, Support Vector Machine, K-Nearest Neighbor, Decision Tree, and Naive Bayes algorithms in Python. The results are evaluated using an accuracy score and confusion matrix, revealing that the Decision Tree XGBoost algorithm achieved the highest accuracy at 82.7%. However, Naive Bayes and Support Vector Machine models performed better in classifying true positives. This research provides valuable insights for organizations seeking to utilize data to understand and address attrition issues [22].

Table I presents an overview of recent studies on enhancing employee performance, highlighting the models applied, datasets used, key outcomes, and the challenges faced.

TABLE I. OVERVIEW OF RECENT STUDIES ON PREDICTIVE MODELING OF ENHANCING EMPLOYEE PERFORMANCE USING MACHINE LEARNING

Author	Proposed Work	Dataset	Key Findings	Challenges/Recommendation
Goyal et al. (2025)	Hybrid Self-Supervised Learning (SSL) + Reinforcement Learning (RL) model using typing & cursor movements with software usage patterns.	Real-world workplace data	Hybrid model outperformed SSL-only and RL-only, achieving highest accuracy, precision, recall, adaptability, and reward rate.	Further testing on diverse work environments to ensure scalability.
Tanmayi et al. (2025)	ML, ensemble models, and deep learning for predicting employee attrition.	Kaggle Employee Attrition Dataset	Random Forest gave best results with high accuracy and F-score; ensemble models also effective.	Explore larger datasets and more hybrid models for improved generalization.
Kosuri et al. (2025)	GRU-Attention model for AI-based sentiment analysis of employee feedback.	Kaggle Employee Feedback Dataset	Achieved higher accuracy (91.8%) compared to RF and GRU; outperformed LSTM-Attention.	Focus on more diverse feedback sources and multilingual datasets.
Prakash et al. (2024)	AI model combining PCA, XGBoost, and Logistic Regression for employee motivation analysis.	IBM HR Analytics Dataset	High accuracy and precision; identified high-risk employees effectively.	Apply model to larger, more complex datasets for broader validation.
Subha et al. (2024)	Cascaded MLP + SVM model with deep gradient optimization for employee attrition prediction.	Preprocessed dataset with SMOTE	Outperformed standalone MLP and SVM; achieved high accuracy with minimal overfitting.	Extend study with real-time employee monitoring for proactive retention.
Priyana et al. (2024)	XGBoost with SMOTE and hyperparameter tuning for employee retention prediction.	Employee Dataset (SMOTE applied)	Accuracy improved from 0.81 to 0.85 after tuning; identified retention factors.	Need for generalization across different industries.
Mascarenhas, Savant & Aswale (2023)	ML models (RF, SVM, KNN, DT, Naive Bayes) to analyze employee performance and attrition causes.	Organizational employee data	Decision Tree XGBoost achieved highest accuracy (82.7%); SVM & NB better at true positives.	Improve balance between precision and recall; validate across larger datasets.

III. RESEARCH METHODOLOGY

The proposed methodology begins with the collection of the Employee Dataset, comprising 858 records with 18 relevant features of both active and resigned employees. Data preprocessing was carried out through multiple steps, including handling missing values, removing duplicate entries, eliminating outliers, encoding categorical variables, and normalizing features using the StandardScaler () to ensure uniformity in scale. To reduce dimensionality and select the most significant attributes, Principal Component Analysis (PCA) was applied, thereby improving computational efficiency and minimizing redundancy. The dataset was then split into training (70%) and testing (30%) sets to ensure unbiased evaluation. Finally, a Deep Neural Network (DNN) model was developed and trained on the processed data, leveraging its capability to capture complex, nonlinear relationships among features to accurately predict employee performance and provide actionable insights for organizational decision-making. The model's performance was evaluated using standard metrics such as accuracy, precision, recall, and F1-score to ensure reliable prediction and classification of Employee Performance, with the complete process illustrated in Figure 1.

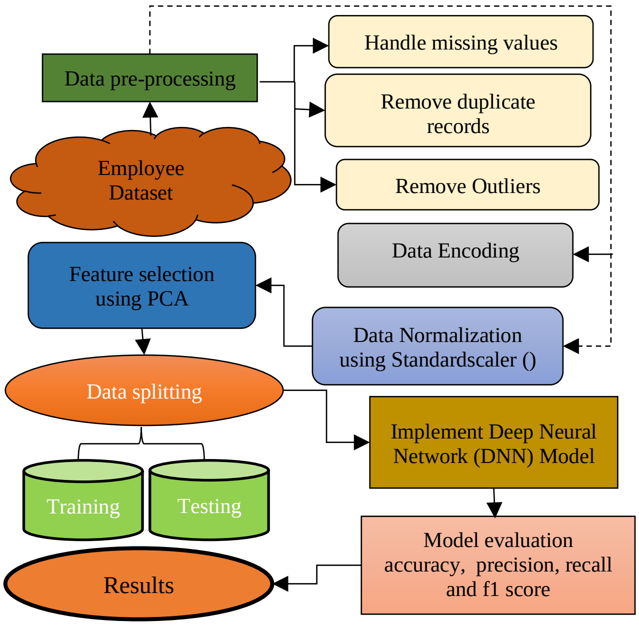

Fig. 1. Proposed flowchart for Enhancing Employee Performance

A detailed explanation of each step in the proposed flowchart for predictive modeling of enhancing employee performance is provided below.

A. Data Collection

This study utilized an Employee Dataset from INX Future Inc. that contains a complete list of 1200 employees along with their organizational and personal details. Although the dataset is composed of 28 fields capturing various demographic, job-related, and performance attributes, only selected features were included for model development. These features are EmpDepartment, EmpJobRole, DistanceFromHome, EmpLastSalaryHikePercent, EmpWorkLifeBalance, EmpHourlyRate, ExperienceYearsAtThisCompany, Age, EmpEnvironment Satisfaction, and Performance Rating. These attributes were used to analyze employee characteristics and support the development of the proposed predictive model. Data visualizations such as bar plots and heatmaps were used to examine Performance distribution, feature correlations etc., are given below:

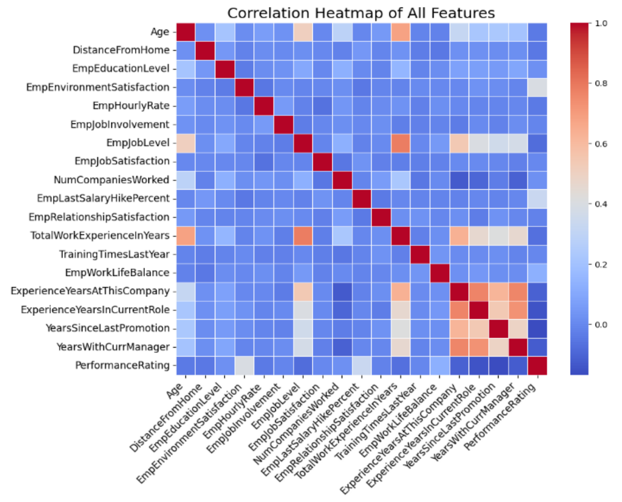

Fig. 2. Correlation Matrix

Figure 2 represents a correlation heatmap of the dataset's features, where each cell indicates the correlation coefficient between two variables. The diagonal line of dark red squares shows perfect self-correlation (value = 1), while the off-diagonal cells display varying degrees of positive (red/orange) and negative (blue) correlations. Most feature relationships appear to be weak to moderate, as reflected by the dominance of blue shades, suggesting low correlation among many attributes. However, a few pairs exhibit stronger positive or negative associations, highlighted in brighter red or light blue tones, which may indicate potential redundancy or interaction effects between those variables. This visualization helps in identifying patterns, dependencies, and multicollinearity within the dataset.

Figure 3 illustrates scatter plots showing the relationship between employee performance ratings and four different attributes: hourly rate, experience, distance from home, and age. In subplot (a), performance ratings appear across different hourly pay levels, with no strong linear relationship but visible clustering within certain pay ranges. Subplot (b) indicates that employees with higher performance ratings generally have a wider distribution of experience, though many clusters around lower experience levels. Subplot (c) shows performance ratings relative to distance from home, where ratings are spread across all distances without a distinct trend, suggesting weak correlation. Finally, subplot (d) displays performance ratings against age, showing that employees of varying ages can achieve similar ratings, though younger employees are more concentrated in mid-level ratings. Overall, the plots suggest that while these features influence performance ratings to some extent, none show a very strong direct relationship.

Fig. 3. Comparison of performance rating with different columns of the dataset

B. Data Pre-Processing

Data preparation began with collecting the Employee Dataset, followed by concatenation and data cleaning to ensure consistency. Relevant features were then extracted, and preprocessing steps were applied to handle missing values, remove duplicates and remove outliers. Further, the dataset underwent transformation and normalization to enhance model performance. The detailed preprocessing steps are outlined as follows:

- **Handle missing value:** To handle missing values, the rows or columns can be eliminated, particularly if the dataset is large and the amount of missing data is small, or if an entire column has a very high percentage of missing values.
- **Remove duplicate records:** To remove duplicate records, use the built-in "Remove Duplicates" feature in Excel (found under the Data tab) to delete rows with identical values in specified columns.
- **Remove Outliers:** Removing outliers during data preprocessing is a common practice to improve the performance and accuracy of statistical analyses and machine learning models.

C. Data Encoding

Data encoding is the process of converting raw or complex data into a structured, standardized format, often numerical, that is suitable for storage, transmission, and processing by machines and algorithms.

D. Standardscaler() For Data Normalization

Given the different scales of each descriptor, the dataset was standardized using the StandardScaler() method to transform the data so that the mean of the resulting distribution is zero and the standard deviation is one. This transformation is achieved by subtracting the mean value of each observation and dividing by the standard deviation, as shown in Equation (1):

$$z = \frac{x-\mu}{\sigma} \tag{1}$$

Where z is the transformed value of the feature, x is the original value of each descriptor, μ is the mean, and σ is the standard deviation of the feature in the dataset.

E. PCA for Feature Selection

Feature selection is a data preprocessing technique that identifies and selects a subset of relevant input variables (features) from a larger dataset to build machine learning models, improving model performance, reducing noise, preventing overfitting, decreasing computational cost, and enhancing interpretability [23]. PCA can be used for feature selection by identifying original features with high loadings on the most significant Principal Components (PCs), which represent the directions of maximum variance in the data.

F. Data Splitting

To evaluate efficiency, the dataset was divided into training and testing subsets, with 70% of the data used for model development and parameter estimation, and the remaining 30% reserved for testing and performance assessment.

G. Proposed Deep Neural Network (DNN)

A deep neural network (DNN) is a deep learning algorithm widely recognized by scholars. The network structure of DNN includes the input layer, hidden layer and output layer and each layer is fully connected [24][25]. Each neuron has no connection with the neurons between the layers and is connected with all the neurons in the next layer. After each layer of the network, there is an activation function acting on the output, which strengthens the effect of network learning. Therefore, DNN can also be understood as a large perceptron composed of multiple perceptrons. Take the ith layer forward propagation calculation as an example, the formula is as follows Equation (2):

$$x_{i+1} = \sigma(\sum w_i x_i + b) \tag{2}$$

Where x represents the input value, w represents the weight coefficient matrices and b represents the bias vector[26]. In a multi-class network, ReLU is usually used as an activation function, the formula is as follows Equation (3):

$$\sigma(x) = max(0, x) \tag{3}$$

The loss function measures the output loss of training samples and calculates the backpropagation of the network through the loss function to optimize the network structure. In the classification task, usually cross-entropy is chosen as the loss function, the formula is as follows (4):

$$C = -\frac{1}{N}\sum_x \sum_{i=1}^{M}(y_i log p_i) \tag{4}$$

Where N represents the number of the input data set, M represents the number of categories, yi represents whether the classification i corresponds to the real category and pi represents the probability of predicting into category i. The Deep Neural Network (DNN) was configured with ReLU activation, Adam optimizer (learning rate 0.001), batch size 32, and 4,000 epochs. Cross-entropy loss handled binary classification, with dropout (0.2) preventing overfitting. The initialization improved stability, and early stopping based on validation loss ensured optimal training.

H. Evaluation Metrics

The efficacy of the suggested design was investigated through a number of performance parameters. The predicted outcomes of trained models were compared with the real values. It comprises of four parameters, such as true positive (TP), false positive (FP), true negative (TN), and false negative (FN) [27]. The following matrix, including accuracy, precision, recall, and F1-score, is explained below:

Accuracy: The ratio of the number of instances correctly predicted by the trained model to the total number of instances in the dataset (input samples). It is given as Equation (5)-

$$Accuracy = \frac{\text{TP+TN}}{\text{TP+Fp+TN+FN}} \tag{5}$$

Precision: Precision is the proportion of positive instances successfully predicted to all positive instances predicted by the model. Precision indicates. How good the classifier is in predicting the positive classes is expressed as Equation (6):

$$Precision = \frac{\text{TP}}{\text{TP+FP}} \tag{6}$$

Recall: This metric, the ratio of events that were accurately predicted as positive to all instances that should have proved positive. In mathematical form it is given as Equation (7):

$$Recall = \frac{\text{TP}}{TP+FN} \tag{7}$$

F1 score: It is a combination of the harmonic mean of precision and recall, that is, it helps to balance recall and precision. Its range is [0, 1]. Mathematically, it is given as Equation (8):

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{8}$$

IV. Results and Discussion

This section outlines the experimental setup and presents the performance of the proposed model during both training and testing phases. The proposed model runs 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. The proposed model was trained on the Employee Dataset and assessed using key performance metrics such as accuracy, precision, recall, and F1-score, as summarized in Table II. The experimental results of the proposed Deep Neural Network (DNN) model for enhancing employee performance on the Employee Dataset indicate highly efficient and reliable performance. The model achieved an impressive accuracy of 96.25%, supported by a precision of 98%, a recall of 97%, and an F1-score of 97%, reflecting its strong ability to correctly classify employee performance with minimal errors. These balanced metrics demonstrate the robustness of the DNN model, highlighting its effectiveness in handling complex patterns within the dataset and making it superior for predictive analysis compared to traditional approaches.

TABLE II. Experimental Results of Proposed Model for Enhancing Employee Performance on Employee Dataset

Performance Matrix	**Deep Neural Networks (DNNs) Model**
Accuracy	96.25
Precision	98
Recall	97
F1-score	97

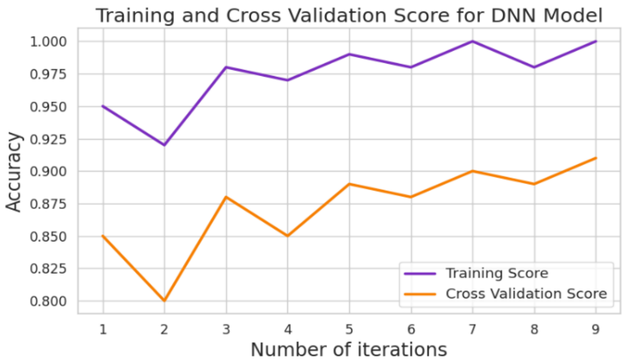

Fig. 4. Accuracy curves for the DNN Model

Figure 4 presents the accuracy curves of a model across nine iterations, showing the comparison between training score (blue) and cross-validation score (green). The training accuracy begins around 0.95, slightly dips, and then steadily improves, approaching nearly 1.0 in later iterations, indicating that the model learns effectively on the training data. The cross-validation accuracy starts lower at about 0.85, drops slightly, and then gradually improves to around 0.91, demonstrating that the model generalizes reasonably well. Although the training score remains consistently higher than the validation score, the relatively parallel trends suggest that the model maintains stability with minimal overfitting, achieving strong predictive performance.

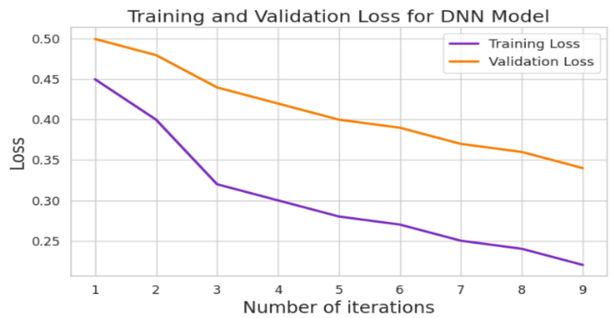

Fig. 5. Loss curves for the DNN Model

Figure 5 shows the loss curves for the DNN model across 9 iterations, depicting training loss (purple) and validation loss (orange). Both curves demonstrate a steady decline, indicating effective learning and convergence of the model. The training loss decreases consistently from about 0.45 to nearly 0.22, while the validation loss reduces from around 0.50 to approximately 0.34, closely following the training curve. The relatively small gap between the two curves suggests that the model generalizes well to unseen data with minimal overfitting, and the final low loss values confirm the robustness and stability of the DNN model in achieving reliable predictive performance.

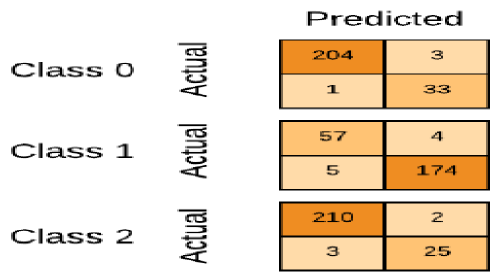

Fig. 6. Confusion matrix for DNN

Figure 6 represents a confusion matrix that illustrates the classification performance of a model across three classes. For Class 0, the model correctly predicted 204 instances while misclassifying a few as Class 1 (3 cases), and similarly achieved 33 correct predictions with only 1 misclassification. In Class 1, the model predicted 174 samples correctly but confused 57 as Class 0 and 4 as Class 2, with 5 instances misclassified in the opposite direction. For Class 2, it accurately classified 210 instances, while misclassifying 2 as Class 1, along with 25 correct predictions, but 3 cases were misclassified. Overall, the diagonal dominance in the matrix indicates strong predictive performance, though some overlap between classes highlights areas for potential improvement.

A. *Comparative Analysis*

To demonstrate the effectiveness of the proposed DNN model, a comparative accuracy analysis with other existing models is presented in Table III. In this evaluation, the table presents the accuracy comparison of different predictive models for enhancing employee performance using the Employee Dataset. Among the models, the Decision Tree (DT) achieved an accuracy of 79.6% with reasonably good precision (87.1%), recall (89%), and F1-score (88.1%), reflecting balanced performance but limited overall accuracy. The Logistic Regression (LR) model showed an accuracy of 86.73%, though its precision (43.37%) and recall (50%) were comparatively low, indicating weaker predictive reliability. The Random Forest (RF) model performed better, with 88.4% accuracy and strong precision (87.23%), recall (88.5%), and F1-score (87.8%), suggesting more stable predictions. However, the Deep Neural Network (DNN) significantly outperformed all other models, achieving the highest accuracy of 96.25% along with superior precision (98%), recall (97%), and F1-score (97%), highlighting its effectiveness and robustness in predicting employee performance.

TABLE III. ACCURACY COMPARISON OF DIFFERENT PREDICTIVE MODELS OF ENHANCING EMPLOYEE PERFORMANCE USING THE EMPLOYEE DATASET

Model	Accuracy	Precision	Recall	F1-Score
DT[28]	79.6	87.1	89	88.1
LR[29]	86.73	43.37	50	-
RF[30]	88.4	87.23	88.5	87.8
DNN	96.25	98	97	97

The proposed Deep Neural Network (DNN) model, with an accuracy of 96.25%, offers a significant advantage by effectively capturing complex, non-linear relationships within the Employee Dataset that traditional models often fail to address. Its high precision, recall, and F1-score further emphasize its reliability in minimizing both false positives and false negatives, ensuring more accurate and consistent predictions of employee performance. This superior performance demonstrates the robustness and adaptability of the DNN model, making it a powerful tool for enhancing decision-making in human resource management and employee performance evaluation.

V. CONCLUSION AND FUTURE STUDY

Human resource management (HRM) plays a crucial role in the effective functioning of modern businesses. However, as the volume of data continues to increase, HR professionals are facing growing challenges in objectively gathering, measuring, and interpreting human resources data. The experimental results reveal that among the evaluated models, the Deep Neural Network (DNN) achieved the highest accuracy of 96.25%, outperforming Decision Tree (79.6%), Logistic Regression (86.73%), and Random Forest (88.4%). This superior performance can be attributed not only to the powerful learning capability of the DNN but also to the effective feature selection process, which involved dimensionality reduction using PCA to eliminate irrelevant and redundant attributes. By retaining only, the most significant features, the model's efficiency and generalization ability were enhanced, leading to more accurate predictions. Hence, the integration of feature selection with DNN provides a robust and reliable framework for employee performance analysis. In the future, a decentralized mechanism can be incorporated with the proposed model to secure the AI-based classification to rank and identify low performers efficiently and reliably.

References

[1] J. Y. Yong, M. Y. Yusliza, T. Ramayah, C. J. Chiappetta Jabbour, S. Sehnem, and V. Mani, "Pathways towards sustainability in manufacturing organizations: Empirical evidence on the role of green human resource management," *Bus. Strateg. Environ.*, 2020, doi: 10.1002/bse.2359.

[2] P. R. Marapatla, "Next-Gen Enterprise BI : A Strategic Guide to AI-Infused Reporting Solutions," *TPM – Testing, Psychom. Methodol. Appl. Psychol.*, vol. 32, no. S7, pp. 248–262, 2025.

[3] H. Kali, "The Future of HR Cybersecurity: AI-Enabled Anomaly Detection in Workday Security," *Int. J. Recent Technol. Sci. Manag.*, vol. 8, no. 6, pp. 80–88, 2023.

[4] M. M. Abdeldayem and S. H. Aldulaimi, "Trends and opportunities of artificial intelligence in human resource management: Aspirations for public sector in Bahrain," *Int. J. Sci. Technol. Res.*, 2020.

[5] A. K. Mishra, "Assessment of Human Resource Capacity of Construction Companies in Nepal," *J. Adv. Res. HR Organ. Manag.*, vol. 05, no. 04, pp. 14–25, Nov. 2018, doi: 10.24321/2454.3268.201804.

[6] A. Parupalli, "Business-Oriented Employee Performance Assessment via Machine Learning in ERP Systems," *Tijer – Int. Res. J.*, vol. 11, no. 11, 2024.

[7] L. Parisi, N. RaviChandran, and M. L. Manaog, "Decision support system to improve postoperative discharge: A novel multi-class classification approach," *Knowledge-Based Syst.*, vol. 152, pp. 1–10, Jul. 2018, doi: 10.1016/j.knosys.2018.03.033.

[8] A. Adeniran and O. Obembe, "The significance of strategic management accounting on the performance of transport businesses in Nigeria," *Insights into Reg. Dev.*, vol. 2, 2020, doi: 10.9770/IRD.2020.2.3(5).

[9] J. R. Vummadi, H. Volikatla, and S. Dodda, "Smart HR for Smart Enterprises: A Machine Learning-Based Approach to Payroll Automation and Time Optimization," *Int. J. Artif. Intell. Data Sci. Mach. Learn.*, vol. 6, no. 3, 2025, doi: 10.63282/3050-9262.IJAIDSML-V6I3P113.

[10] P. Budhwar, A. Malik, M. T. T. De Silva, and P. Thevisuthan, "Artificial intelligence – challenges and opportunities for international HRM: a review and research agenda," *Int. J. Hum. Resour. Manag.*, vol. 33, no. 6, pp. 1065–1097, Mar. 2022, doi: 10.1080/09585192.2022.2035161.

[11] R. Q. Majumder, "Machine Learning for Predictive Analytics: Trends and Future Directions," *Int. J. Innov. Sci. Res. Technol.*, vol. 10, no. 4, pp. 3557–3564, May 2025, doi: 10.38124/ijisrt/25apr1899.

[12] S. Thangavel, S. Srinivasan, S. B. V. Naga, and K. Narukulla, "Distributed Machine Learning for Big Data Analytics: Challenges, Architectures, and Optimizations," *Int. J. Artif. Intell. Data Sci. Mach. Learn.*, vol. 4, no. 3, pp. 18–30, Oct. 2023, doi: 10.63282/3050-9262.IJAIDSML-V4I3P103.

[13] Selvamohana K, "From HR Analytics to AI-Driven HRM: Enhancing Workforce Productivity and Engagement," *J. Inf. Syst. Eng. Manag.*, vol. 10, no. 21s, pp. 578–585, Mar. 2025, doi: 10.52783/jisem.v10i21s.3395.

[14] N. Prajapati, "The Role of Machine Learning in Big Data Analytics: Tools, Techniques, and Applications," *ESP J. Eng. Technol. Adv.*, vol. 5, no. 2, pp. 16–22, 2025, doi: 10.56472/25832646/JETA-V5I2P103.

[15] R. Shettigar, A. Ravisankar, V. Pawar, and G. N. Mandala, "Optimizing Remote Work Performance Management: Ethical Implications of Ai and Ai Advancements," *SSRN Electron. J.*, 2025, doi: 10.2139/ssrn.5076173.

[16] R. Goyal, G. Gupta, S. Gupta, J. K. Kakar, P. Singla, and S. Kaur, "Enhancing Employee Productivity Prediction through Self-Supervised Learning and Reinforcement Learning," in *2025 International Conference on Computing Technologies (ICOCT)*, 2025, pp. 1–6. doi: 10.1109/ICOCT64433.2025.11118832.

[17] B. Tanmayi, H. Rishab, N. N. Yadav, A. A. S, and M. Venugopalan, "Comprehensive Analysis of Employee Attrition Using Machine Learning, Deep Learning and Ensemble Methods," in *2025 International Conference on Knowledge Engineering and Communication Systems (ICKECS)*, 2025, pp. 1–6. doi: 10.1109/ICKECS65700.2025.11034779.

[18] S. Kosuri, N. K. Bhasin, G. Dhanalakshmi, R. Murugan, V. Selvi, and S. Muthuperumal, "AI-based Sentiment Analysis for Enhancing Employee Satisfaction using GRU-Attention Mechanism," in *2025 International Conference on Next Generation Information System Engineering (NGISE)*, 2025, pp. 1–6. doi: 10.1109/NGISE64126.2025.11085256.

[19] S. Prakash, P. S, G. S, and K. Raja.K, "Enhancing Employee Motivation Through AI: The Role of Corporate Culture in the Workplace," in *2024 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, 2024, pp. 1–7. doi: 10.1109/ICSES63760.2024.10910789.

[20] T. D. Subha, S. Vijayalakshmi, M. N. Pooja, R. K. V, J. Pavithira, and R. V, "Experimental Evaluation of a Novel Methodology to Predict Employee Attrition Using Elevated Deep Gradient Optimization Logic," in *2024 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, 2024, pp. 1–6. doi: 10.1109/ICSES63760.2024.10910261.

[21] I. Priyana, N. Alamsyah, Budiman, A. P. Sarifiyono, and E. Rusnendar, "Predictive Boosting for Employee Retention with SMOTE and XGBoost Hyperparameter Tuning," in *2024 International Conference on Smart Computing, IoT and Machine Learning (SIML)*, IEEE, Jun. 2024, pp. 92–97. doi: 10.1109/SIML61815.2024.10578116.

[22] J. Mascarenhas, S. Savant, and S. Aswale, "Employee Appraisal Prediction Using Classification Models," in *2023 6th International Conference on Contemporary Computing and Informatics (IC3I)*, 2023, pp. 2170–2175. doi: 10.1109/IC3I59117.2023.10398114.

[23] U. A. Korat and A. Alimohammad, "A Reconfigurable Hardware Architecture for Principal Component Analysis," *Circuits, Syst. Signal Process.*, vol. 38, no. 5, pp. 2097–2113, 2019, doi: 10.1007/s00034-018-0953-y.

[24] S. Nokhwal, P. Chilakalapudi, P. Donekal, S. Nokhwal, S. Pahune, and A. Chaudhary, "Accelerating Neural Network Training: A Brief Review," in *2024 8th International Conference on Intelligent Systems Metaheuristics & Swarm Intelligence (ISMSI)*, New York, NY, USA: ACM, Apr. 2024, pp. 31–35. doi: 10.1145/3665065.3665071.

[25] R. Tarafdar and Y. Han, "Finding Majority for Integer Elements," *J. Comput. Sci. Coll.*, vol. 33, no. 5, pp. 187–191, 2018.

[26] S. Nokhwal, S. Nokhwal, S. Pahune, and A. Chaudhary, "Quantum Generative Adversarial Networks: Bridging Classical and Quantum Realms," in *2024 8th International Conference on Intelligent Systems, Metaheuristics & Swarm Intelligence (ISMSI)*, New York, NY, USA, NY, USA: ACM, Apr. 2024, pp. 105–109. doi: 10.1145/3665065.3665082.

[27] G. James, D. Witten, T. Hastie, and R. Tibishirani, *An Introduction to Statistical Learning: With Applications in R*, 1st ed. Springer, 2013.

[28] M. G. T. Li, M. Lazo, A. K. Balan, and J. de Goma, "Employee Performance Prediction using Different Supervised Classifiers," in *Proceedings of the International Conference on Industrial Engineering and Operations Management*, Michigan, USA: IEOM Society International, Mar. 2021. doi:

10.46254/AN11.20211188.

[29] S. Sonavane, S. Pawar, T. Gupta, and S. Bhoir, “HR Analytics to Track Employee Performance using Python,” *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 13, no. 3, pp. 1698–1704, 2025, doi: 10.22214/ijraset.2025.67642.

[30] G. Thakre, P. Wankhede, S. Patle, S. Joshi, and A. Chauhan, “Implementation of Machine Learning Model for Employee Retention Prediction,” *Int. J. Sci. Res. Sci. Technol.*, vol. 8, no. 3, pp. 503–508, 2021.