ANN-Based Employee Performance Prediction: A Comparative Analysis of Optimization Techniques

Rahaf Mohammed Bajhzer, Yousef Alsenani, Sahar Jambi, Tawfiq Hasanin
Faculty of Computing and Information Technology-Information Systems Department,
King Abdulaziz University, Jeddah 21589, Saudi Arabia

***Abstract*—With the increasing use of artificial intelligence in decision-making systems, predicting employee performance has attracted growing attention in human resource analytics. This study aims to systematically evaluate the impact of data preprocessing and model optimization techniques on artificial neural network (ANN)-based prediction of employee performance in HR analytics. Three publicly available HR datasets were used, and multiple configurations involving feature selection, feature extraction, principal component analysis (PCA), reduced architectures, and regularization were evaluated. The experimental results show that appropriate feature selection and regularization consistently improve predictive performance across datasets, whereas PCA-based dimensionality reduction resulted in lower accuracy in the evaluated datasets, possibly due to the loss of discriminative information. Additionally, simplified ANN architectures yielded modest, but consistent improvements in generalization performance across datasets, highlighting the importance of controlling model complexity. The top-performing configurations across the assessed datasets achieved accuracies ranging from 81% to 96%. These findings offer practical guidance on selecting efficient preprocessing and architectural techniques when applying ANN-based models in human resource analytics.**

Keywords*—*Employee performance prediction; artificial neural networks; data preprocessing; model optimization; HR analytics

I. Introduction

Artificial intelligence (AI) techniques have made a significant impact on information technology (IT), and innovations in AI-based workflows are already known and implemented by organizations, enhancing employee productivity and satisfaction. Information technologies are no longer confined to managing static datasets; they can now analyze massive amounts of data and produce deeper insights. Since employees' skills and talents are sources of growth and strategic advantage for organizations, human resource management (HR) has received greater recognition and attention in recent years [1]. The ability to predict the future performance of an employee has long been a core function of human resource management. HR analytics has significantly improved decision-making processes by analyzing data to inform performance management strategies [2]. Employee performance refers to how an employee completes tasks and fulfills assigned responsibilities. Researchers have designed various methods to assess employee performance in businesses. These include assessments of the employee's performance across personal attributes, psychological factors, and educational levels. However, not all sectors might be examining these aspects [3]. An adequate analysis of performance indicators is also critical to measuring employee productivity. However, distinguishing between high and low performers and offering guidance to workers remains a difficult task in human capital management. Many organizations fail to conduct performance reviews systematically, leading to unreliable evaluations. To enhance the accuracy and effectiveness of workplace performance predictions, machine learning techniques can be used. Employee performance is affected by the use of statistical methods, data visualization, and forecasting techniques in business analytics, which help determine relationships, trends, and patterns. In addition, machine learning algorithms refine the analysis by capturing complex patterns and relationships in the data, enabling more accurate performance predictions. With this comprehensive approach, organizations can improve their analysis, personnel management, performance optimization, and overall success [4].

Most previous research has employed traditional machine learning algorithms or ensemble techniques to predict employee performance, but neural network approaches have received comparatively less attention in this area. Such reliance may limit the ability to discern complex, nonlinear patterns and deep relationships among variables that profoundly affect employee performance. In addition, many studies rely on a single dataset, which can reduce the accuracy and generalizability of predictions, especially in complex organizational environments. Furthermore, there has been a lack of systematic analysis of data preprocessing and model optimization across multiple datasets. This gap underscores the need for further research to assess the effectiveness, robustness, and practical applicability of neural network techniques for predicting employee performance.

Artificial neural networks (ANNs) were selected in this study due to their strong capability to model complex nonlinear relationships between employee attributes and performance outcomes. In contrast to conventional tree-based models, such as Random Forests or Decision Trees, ANNs can capture hierarchical interactions and learn deeper representations from heterogeneous HR data. This capability makes ANN-based models particularly suitable for high-dimensional datasets, where interactions among variables may be complex and difficult to represent using traditional machine learning methods.

The main aim of this research is to logically evaluate the performance of artificial neural networks (ANNs) for predicting employee performance by examining the effects of various data preprocessing methods and model optimization approaches across different publicly available HR datasets. Accordingly, the main research question is: What is the impact of data preprocessing and architectural optimization on the effectiveness of ANN models for predicting employee performance?

This study makes the following contributions:

- It introduces a unified cross-dataset evaluation framework for systematically assessing preprocessing and ANN optimization strategies in HR analytics, providing generalized insights that extend beyond dataset-specific findings.
- It provides empirical evidence on the effectiveness of feature selection, architectural simplification, and regularization strategies, demonstrating their consistent impact on improving model generalization across heterogeneous HR datasets.
- It offers a comparative analysis of dimensionality reduction techniques, showing that PCA may lead to the loss of discriminative information in HR datasets, thereby adversely affecting classification performance.
- It presents practical methodological guidelines for designing efficient ANN-based predictive models in HR analytics, enabling both researchers and HR practitioners to select appropriate preprocessing and optimization strategies.

The study is structured as follows: Section II reviews the literature. Section III outlines the methodology. Section IV describes the experiments. Section V presents the results. Section VI provides the discussion. Finally, Section VII concludes the study and outlines directions for future research.

II. Literature Review

Employee performance prediction has become a growing focus in HR analytics because of its effects on corporate productivity, strategic decision-making, and long-term competitive advantage. The literature review is structured into three interrelated streams. First, previous work on the direct prediction of employee performance using machine learning and artificial intelligence techniques is reviewed. Second, related studies on student performance prediction are considered, which are methodologically close to the present study and implement artificial neural networks and multi-class classification. Finally, research on employee attrition prediction is reviewed as a closely related area of HR analytics that offers similar predictive modeling challenges in workforce management.

A. Employee Performance Prediction

Some early HR analytics research used statistical and regression-based methods to identify factors affecting employee performance. Kumari et al. [2] used regression analysis and a mixed-methods approach to investigate the influence of training, engagement, and job satisfaction on organizational growth, showing that variations in these variables accounted for 29% of the variance in employees' performance. Although such approaches provided useful insights into the determinants of performance, their ability to capture complex nonlinear relationships between employee attributes remained limited, which restricts their effectiveness when dealing with large and multidimensional HR datasets.

Numerous recent studies have employed conventional classifiers, including Decision Trees, Random Forests (RF), Support Vector Machines (SVM), Naive Bayes, and Logistic Regression. Nayem and Uddin [5] reported that Random Forest gave the best performance among the evaluated models. Nasr et al. [6] used three conventional classifiers on a relatively small sample of survey data and found that SVM showed the highest accuracy (86.9%) with years of experience and professional training as key determinants of performance. Although these studies demonstrate the effectiveness of traditional machine learning algorithms, most focus primarily on model comparison or predictive accuracy rather than examining how data preprocessing strategies or dataset characteristics influence model behavior. For example, Asuquo et al. [7] and Jayadi et al. [8] reported strong performance for Random Forest and Naïve Bayes, respectively, yet the role of preprocessing strategies such as feature selection or dimensionality reduction was not systematically investigated. Consequently, it remains unclear how different preprocessing strategies influence model performance across datasets with varying characteristics.

Patel et al. [3] proposed a multi-classifier AI-based framework and demonstrated that ensemble learning approaches outperformed standalone models by 3–4% in accuracy. Similarly, Obiedat and Toubasi [9] showed significant improvements in accuracy when combining AdaBoost and Bagging with conventional classifiers, highlighting the potential of ensemble techniques to enhance classification performance. However, both studies primarily emphasize improvements in accuracy through model aggregation while offering limited discussion of how preprocessing strategies, feature engineering, or dataset complexity affect the overall predictive performance of these models.

Hasan et al. [10] proposed an integrated framework combining business analytics and machine learning to predict employee performance across multiple data sources and analytical stages, including preprocessing, feature selection, and model optimization. The findings suggest that combining machine learning and business analytics can help estimate worker performance more accurately and enhance data-driven decision-making in HR management. However, rather than methodically assessing neural network topologies or preprocessing approaches across several datasets, the study primarily focuses on integrating analytical techniques.

Other studies have focused on hyperparameter adjustment and ensemble learning. Sinha [4] discussed ensemble-based algorithms combined with hyperparameter optimization techniques such as Optuna, Bayesian optimization, and randomized search, with Gradient Boosting achieving the best accuracy (96.2%). Similarly, Tanasescu et al. [11] applied the CRISP-DM framework and found that XGBoost achieved the highest accuracy after systematic hyperparameter tuning using the Optuna framework. While these studies highlight the importance of optimization techniques for improving model performance, they continue to rely heavily on tree-based models, leaving neural network approaches relatively underexplored in HR analytics, particularly in relation to preprocessing strategies and architectural design choices.

B. Student Performance Prediction

Student performance prediction has received considerable attention in educational data mining, and several studies have applied machine learning and neural network models to identify patterns that influence academic outcomes. These studies are methodologically relevant to the present research because they address similar multi-class classification problems and frequently employ neural network architectures.

Yağcı [12] compared several machine learning algorithms for predicting academic performance and reported that the proposed model achieved classification accuracies of 70%–75%. Although these results demonstrate the applicability of machine learning techniques in educational contexts, the relatively moderate accuracy suggests that model performance may depend strongly on feature quality and dataset characteristics.

Recent research has also studied deep learning architectures in conjunction with data-balancing strategies. Aslam et al. [13] achieved good predictive results with their deep learning model, which incorporated SMOTE to address data imbalance, highlighting the importance of preprocessing strategies when dealing with skewed educational datasets.

Other studies have explored neural network–based approaches. Lau et al. [14] developed a hybrid predictive model combining statistical methods with neural networks trained using the Levenberg-Marquardt algorithm. The neural network model consists of 11 input variables, 2 hidden layers, and 1 output layer, achieving an accuracy of 84.8%. Likewise, Abu Naser et al. [15] developed an Artificial Neural Network (ANN) model based on a multilayer perceptron trained via backpropagation, achieving an accuracy of 84.6%. Similarly, Chavez et al. [16] developed an artificial neural network model to predict student academic performance using nonpersonal learning data from the Open University Learning Analytics dataset. Behavioral learning indicators, such as course interactions and evaluation scores, were used as input features, achieving an accuracy of 93.81%.

Several additional studies further support the effectiveness of ANN-based approaches for predicting academic outcomes. Baashar et al. [17] conducted a comprehensive review examining the application of Artificial Neural Networks (ANNs) for predicting students' academic performance. Their findings indicate that ANN-based approaches generally achieve strong predictive performance when combined with data mining and analytics techniques; however, the selection of input variables varies substantially across studies and often depends on data availability and contextual factors. Mengash [18] applied various data mining techniques to predict early academic achievement based on pre-admission standards and reported that the ANN model achieved the highest predictive accuracy of 79%, further demonstrating neural networks' ability to capture complex relationships in educational data. Similarly, Altabrawee et al. [19] examined several machine learning techniques and found that the ANN model achieved the best performance, with an accuracy of 77.04%.

These results demonstrate how neural network models can capture intricate nonlinear correlations in educational data. However, most existing studies evaluate neural network models using a single dataset and focus primarily on predictive accuracy, while providing limited analysis of how architectural configurations or preprocessing strategies influence model performance across different datasets.

C. Employee Attrition Prediction

Several studies have compared multiple machine learning algorithms to identify the most effective predictive models. Setiawan et al. [20] used a logistic regression model in an HR analytics framework to examine employee attrition. Their findings demonstrated the value of statistical methods for identifying key factors influencing employee turnover, with the proposed model achieving an accuracy of almost 75%. Fallucchi et al. [1] evaluated several algorithms to predict employee attrition and concluded that the best predictive performance was obtained by Gaussian Naive Bayes. Similarly, Nandal et al. [21] conducted a comparative analysis of machine learning, ensemble, and deep learning approaches, reporting that the Feedforward Neural Network achieved the highest accuracy of 83.93%. In a similar study, Mansor et al. [22] examined the use of machine learning methods to forecast employee attrition and showed that classification models can successfully identify trends in HR analytics datasets related to employee turnover. Additionally, Jain et al. [23] suggested a machine learning strategy that uses many categorization models to explain and forecast employee attrition. According to their results, Random Forest had the best predictive performance, indicating that ensemble learning methods can offer reliable ways to identify workers who are at risk of leaving the company. Likewise, Pratt et al. [24] evaluated several machine learning methods for predicting employee attrition and found that the Random Forest method performed best, with an accuracy of almost 85%. Punnoose and Ajit [25] used several supervised machine learning algorithms to study employee turnover prediction and found that the Extreme Gradient Boosting (XGBoost) classifier yielded the best predictive performance. Their findings demonstrate how well boosting-based models handle noisy HR information and raise the accuracy of turnover prediction. These studies collectively demonstrate that ensemble-based models, particularly Random Forest and boosting algorithms such as XGBoost, often outperform traditional classifiers in employee attrition prediction tasks.

Marín Díaz et al. [26] applied explainable artificial intelligence (XAI) techniques to analyze employee attrition and improve transparency in predictive models. Their results demonstrate the importance of interpretable machine learning techniques in facilitating better-informed HR decision-making. Alsheref et al. [27] proposed an ensemble-based predictive framework to forecast employee turnover via hyperparameter tuning, highlighting that organizational context influences the model's effectiveness.

Other studies have incorporated additional preprocessing and optimization techniques. Arqawi et al. [28] applied preprocessing and SMOTE-based balancing before training deep learning models and reported improved predictive performance with an accuracy of 94.52%. Likewise, Raza et al. [29] used exploratory data analysis, feature engineering, and SMOTE-based balancing, demonstrating that an optimized Extra Trees classifier could achieve 93% accuracy. Similarly, Al-Darraji et al. [30] developed a deep neural network model by incorporating preprocessing and data balancing techniques; their model achieved accuracies of 91% on the original dataset and 94% on the balanced dataset. These findings demonstrate the importance of preprocessing and class balancing for improving the predictive performance of HR analytics.

These results emphasize the importance of data balancing and preprocessing techniques for enhancing model performance; however, most research focuses on predictive accuracy rather than systematically examining how model architectures or preprocessing techniques affect generalization across datasets.

D. Research Gap

Even though these studies [2], [3], [4], [11], [5], [9], [7], [6], [8], [10] showed strong predictive performance, several limitations remain. The majority of the research is based on a single dataset, limiting the generalizability of the results. Additionally, they concentrate more on hyperparameter tuning and model comparison than on the systematic assessment of preprocessing techniques or neural network architecture design across multiple datasets.

In contrast to studies in other areas on predicting student performance [12], [13], [14], [15], [16], [17], [18], [19], which mainly evaluate neural networks on a single dataset, systematic comparisons of preprocessing strategies and modern optimization techniques are still lacking.

Although employee attrition studies [1], [27], [28], [21], [29], [22], [23], [30], [26], [24], [20], [25] demonstrate how machine learning and deep learning can be applied to HR analytics, their main focus is on model comparison and predictive performance. Comprehensive cross-dataset evaluation of preprocessing methodologies, optimization approaches, and the effects of neural network architectural complexity has received little attention.

To address these gaps, this study utilizes three publicly available HR datasets to systematically evaluate and compare preprocessing methods and model tuning strategies for ANN-based employee performance prediction.

III. Methodology

A. Dataset Description

The first dataset, obtained from Kaggle (HR metrics and analytics), contains information on 311 employees and 36 attributes. Of these 36 attributes, there were no duplicates and no missing values apart from 207 missing values in the 'Date of Termination' attribute and 8 missing values in the 'ManagerID' attribute. The dataset includes both categorical and numerical variables that reveal employees' characteristics, demographics, job details, performance, and identification details. Also included are attributes that capture whether an employee is satisfied with their job, as well as those related to their salary and attendance. The target variable is the performance score (Exceeds, Fully Meets, Needs Improvement, and PIP).

The second dataset from INX Future Inc. consists of 1200 employees and 28 different features. There are no missing values or duplicates in this dataset. The dataset has more numerical features than categorical features. The categorical features include EMP number, gender, education level, marital status, department, position, business travel, overtime, and whether the employee has left the company. The numerical features include age, job satisfaction, years of service, percentage of last salary hike, and promotion. The performance rating is the target and has the values 2, 3, and 4 (2 being the lowest and 4 the highest).

The third dataset (Employees' performance) from Kaggle contains 17,417 employees and 13 features. There are no duplicates or missing variables in the dataset, except for 771 missing values in the education column and 1,363 missing values in the previous year rating column. The dataset contains variables such as department, region, education, gender, and recruitment channel, all of which are categorical. employee ID, number of trainings, age, previous year rating, length of service, KPIs met more than 80, awards won, and average training score are numerical. The target variable (average training score) has been divided into four performance classes to facilitate the determination of employee performance: low, medium-low, medium-high, and high. Table I shows the summary of all three datasets.

The class distribution varies across the evaluated datasets. The HR metrics and analytics dataset exhibits noticeable class imbalance, as some performance categories have few instances. In contrast, the INX dataset shows a relatively balanced distribution across the three classes. The Employees' performance dataset contains a larger number of samples and demonstrates a moderately balanced class distribution.

B. Data Preprocessing

To understand how attributes relate to each other and how they are arranged across datasets, we used Exploratory Data Analysis (EDA) tools, such as histograms, scatterplots, and boxplots, to analyze the data before preprocessing. The data preprocessing stage consisted of several steps applied consistently across all datasets.

1) Categorical encoding: In data science and machine learning, handling categorical data can be challenging. Some target variables are categorical by default, so we applied preprocessing techniques, such as label encoding, to convert non-numeric columns to numeric values to improve the data for machine learning models and extract useful information.

2) Handling missing values: Using scikit-learn's SimpleImputer, we handled missing values (NaNs). For numeric columns, we used the mean as the fill value (imputing the column's mean); for categorical columns, we used the most frequent value (imputing the most frequent category). To avoid data leakage, imputation was applied post train-test split. Imputation was based solely on the properties of the training data set. We carefully handled missing values to ensure the dataset was reliable and complete.

3) Feature selection: Before training the model on the datasets, we identified the most impactful and relevant features that affect an employee's performance. Only training data was used for feature selection to avoid data leakage, and the selected features were then applied to the test set. Initially, irrelevant features, such as unique identifier columns (e.g., ManagerID and EmpID in the HR metrics and analytics dataset, EMP number in the INX dataset, and employee ID in the Employees' performance dataset), were filtered out. These features do not contribute to the predictive model and should not be included. We identified and removed features that could slow down model training or degrade performance, such as columns with weak correlations with the target variable. Based on the correlation analysis below:

TABLE I. SUMMARY OF ALL DATASETS

Datasets	Rows	Columns	Missing Values	Duplications	Categorical Columns	Integer Columns	Target Variable	Target Classes
HR metrics and analytics dataset	311	36	Date of Termination (207 Nulls), ManagerID (8 Nulls)	None	18	18	Performance score	Four (Exceeds, Fully Meets, Needs Improvement, PIP)
INX dataset	1200	28	None	None	9	19	Performance rating	Three (2,3,4)
Employees' performance dataset	17.417	13	education (771 Nulls), previous year rating (1363 Nulls)	None	5	8	Average training score	Four (Medium-High, Low, Medium-Low, High)

In the HR metrics and analytics dataset, the performance score showed a strong positive correlation with DaysLateLast30, EmpID, and Termd; a weak positive correlation with RaceDesc and Manager_Name; a strong negative correlation with PerfScoreID and EngagementSurvey; and a weak negative correlation with Employee_Name and FromDiversityJobFairID.

In the INX dataset, performance rating showed a strong positive correlation with EMP environment satisfaction, EMP last salary hike percent, and EMP work-life balance; a weak positive correlation with overtime, marital status, num companies worked, EMP education level, education background, and EMP job satisfaction; and a weak negative correlation with other columns.

In the employees' performance dataset, the average training score showed a weak positive correlation with the number of awards won and with KPIs met more than 80, as well as a weak negative correlation with length of service and age.

4) Feature extraction: By merging the two features most closely related to the target, we created new features to improve the dataset. Potential interaction effects between the most predictive features were captured using simple arithmetic combinations (addition and multiplication). To avoid data leakage, feature extraction was performed independently for each dataset using only the training data, and the constructed features were subsequently applied to the test set. These combination features increase the dataset's expressive power and may improve model performance. The process of feature extraction seeks to capture essential information by retaining the most critical patterns or structures in the data, capturing interactions or combined effects that are not represented by individual features, eliminating irrelevant or redundant information, improving interpretability through the creation of features that are easier to understand or visualize, and enhancing generalization performance, particularly when the original features are strongly correlated with one another and with the target variable.

5) Principal Component Analysis (PCA): We employed PCA across all datasets, setting n_components = 0.90 to retain 90% of the total variance. PCA was fitted only on the training data and then applied to the test data to prevent data leakage. PCA projects the original feature space into a lower-dimensional space by constructing a set of principal components that capture directions of maximum variance in the data. This method reduces dimensionality and computational complexity and can help mitigate noise. PCA relies on linear algebra concepts, particularly eigenvectors and eigenvalues.

The covariance matrix measures how different variables relate to each other. If two variables, X3 and X4, are strongly correlated, PCA can combine them into a single component. Eigenvectors represent the orientations (axes) of the new coordinate system formed by the principal components, and each principal component is associated with its own eigenvector. Eigenvalues indicate how much of the variance in the dataset is captured by each principal component, where larger eigenvalues correspond to components that explain more variance. Finally, the data is projected into a lower-dimensional space defined by the selected principal components.

6) Feature scaling: Feature scaling is necessary because it enables the creation of accurate and efficient machine learning models. We used StandardScaler on all datasets to standardize the features, setting them to have a mean of zero and a standard deviation of one. This step is critical because it improves the convergence of gradient-based optimization algorithms used in neural networks, ensures that features with different scales contribute equally to the model, and mitigates the impact of outliers by normalizing the data distribution. We applied StandardScaler after splitting the data to prevent data leakage and ensure the test set remained unseen during training. Applying scaling implemented before the split may have integrated test-set data into the training set, leading to an underestimate of model performance.

C. Data Splitting

We divided each dataset into a training and test set (70% for training and 30% for testing). While training the model, we set aside an additional 20% of the training set as a validation set to assess the model's fit on novel data. This splitting strategy ensures sufficient training data, an independent test set for unbiased evaluation, and a validation dataset for hyperparameter tuning and overfitting detection.

D. Proposed ANN Model Implementation

After the objectives were set, we carefully studied all the datasets in the research. Next, we built a sequential neural network model using the Keras library. We developed 12 versions of the model using two primary work strategies: manipulating the input dataset and tuning parameters. The manipulation of input datasets included removing the least-correlated features, creating features by combining the two features with the highest correlation with the target, and performing principal component analysis. The corresponding changes to the model architecture included reducing the number of layers and using regularization and dropout.

E. Model Training and Evaluation

We trained the models for 100 epochs with a batch size of 32 for each dataset. Hyperparameters were chosen based on preliminary experimental trials and commonly used configurations in the neural network literature. To ensure adequate model convergence, the number of epochs was set to 100, while a batch size of 32 was used to balance computational efficiency and training stability. To further mitigate overfitting and improve generalization, L2 regularization with a coefficient of 0.01 and a dropout rate of 0.5 were applied. We set aside 20% of the training data for validation to monitor loss and accuracy. During training, accuracy and categorical cross-entropy loss were measured on both the training and validation sets. Learning curves were plotted across epochs to analyze convergence behavior and detect potential underfitting or overfitting, thereby supporting informed decisions regarding model architecture, regularization strategies, and training duration. For each model, a confusion matrix was used to evaluate class-level performance by analyzing true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). Accuracy, precision, recall, and F1-score were calculated to provide a comprehensive performance assessment. The confusion matrix presented actual classes as rows and predicted classes as columns, enabling identification of misclassification patterns.

- True Positives (TP): Instances where the model correctly predicted the performance rating.
- True Negatives (TN): Instances correctly identified as not belonging to a given class.
- False Positives (FP): Instances where the model incorrectly classified a sample as belonging to a given.
- False Negatives (FN): Instances where the model failed to predict the correct class.

As shown in the flowchart of Fig. 1, we developed a methodological framework to predict employee performance. The flowchart presented the methodology at each step, beginning with the description and analysis of employee performance datasets. Subsequent steps included data preprocessing and dividing the dataset into training, validation, and test sets. Following this, an artificial neural network (ANN) model was implemented. The last steps involved training the model and evaluating its performance to determine its predictive accuracy.

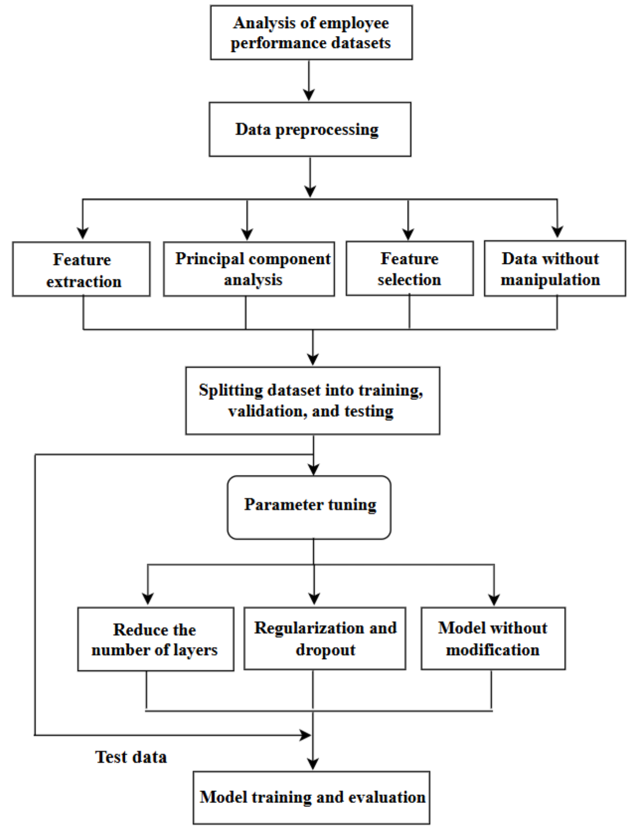

Fig. 1. Methodology framework.

IV. Experiments

The initial model (without data manipulation or architectural modifications) consisted of three layers: 128 neurons, 64 neurons, and an output layer with a softmax activation function for multi-class classification. The categorical cross-entropy loss function was used as it is suitable for multi-class problems. The Adam optimizer was employed for efficient gradient-based optimization. ReLU activation was applied in the hidden layers to introduce non-linearity. The initial results served as a baseline for subsequent experiments.

A. Data Preprocessing Techniques

Three strategies were used to preprocess the data to increase the ANN's predictive capability.

1) Feature selection: To reduce noise and dimensionality, we targeted and excluded features that were minimally correlated with the prediction variable. For instance, in the HR metrics and analytics dataset (RaceDesc, MaritalDesc, and Manager_Name), the INX dataset (education background, EMP job satisfaction, gender, and training times last year), and the Employees' performance dataset (recruitment channel and education). The ANN was trained on these reduced datasets to enhance computational efficiency while improving model accuracy.

2) Feature extraction: Based on the two features with the highest correlation with the prediction variable, we performed arithmetic operations (multiplication and addition) to create new features. For example, in the HR metrics and analytics dataset (DaysLateLast30 and Salary), the INX dataset (EMP environment satisfaction and EMP last salary hike percent), and the Employees' performance dataset (KPI met more than 80 and age). This approach was taken to better capture potential nonlinear interactions between the features and the target, thereby improving the model's overall performance.

3) Principal Component Analysis (PCA): PCA was used to reduce the dataset's dimensionality while retaining the components that explain most of the variance. We set n_components = 0.90 to retain 90% of the total variance. For the HR metrics and analytics dataset, the original set comprised 311 instances and 36 features. After applying PCA, the dimensionality was reduced to 20 components. In the INX dataset, which initially consisted of 1200 instances and 28 features, PCA reduced it to 19 components. For the Employees' performance dataset, which initially had 17417 instances and 13 features, PCA reduced it to 9 components. We trained the ANN on these transformed datasets.

B. Parameter Tuning

We implemented two model control strategies to improve the ANN's performance.

1) Reducing the number of layers: To continue the investigation into model optimization and complexity reduction, we introduced a new architecture with fewer layers and neurons. The model consisted of two hidden layers with 64 and 32 neurons, and an output layer with a softmax activation function for multi-class classification. The categorical cross-entropy loss function was used since it is suitable for multi-class problems. The Adam optimizer was used for gradient-based optimization. ReLU activation was applied in the hidden layers to introduce nonlinearity and enable the model to learn complex patterns. This reduced architecture aimed to mitigate overfitting and decrease computational complexity while maintaining effective multi-class classification performance.

2) Regularization and dropout: To mitigate overfitting, while improving generalization, dropout and L2 regularization were applied. L2 regularization added a penalty term (0.01) to the loss function to constrain the magnitude of the weights, thereby reducing model complexity and overfitting. Dropout randomly deactivated 50% of neurons during training to prevent over-reliance on specific neurons and enhance generalization.

V. RESULTS

A. Results of HR Metrics and Analytics Dataset

Results for the HR metrics and analytics dataset showed that the initial ANN model achieved the highest accuracy of 93.62% compared to the other datasets. This percentage reflected the model's ability to accurately classify cases and to identify significant patterns in the data. The initial ANN model served as a baseline for comparison.

Removing features with lower correlations from the initial model and combining feature selection with a reduced-layer model yielded the highest accuracy of 96.81%, suggesting that feature selection was successful for this dataset. The accuracy of feature selection with regularization and dropout was 95.74%, which indicated an improvement over the first model but not an optimal configuration.

The model's performance increased to 95.74% when feature extraction techniques were applied to the initial model, demonstrating a noticeable improvement in classification. In addition, employing feature extraction in a model with fewer layers, as well as combining feature extraction with regularization and dropout, resulted in similar accuracy levels. This enhancement implied that the ANN was better able to identify underlying patterns and relationships by highlighting the key characteristics.

The accuracy dropped to 88.30% when Principal Component Analysis (PCA) was used on the initial ANN model. When PCA was applied to a model with fewer layers and with the regularized model incorporating dropout, the accuracy further decreased to 86.17%. These outcomes suggest that PCA was less effective for this dataset, as class-discriminative information may have been lost during dimensionality reduction.

The ANN model's accuracy increased significantly to 96.81% when the number of layers was reduced, as well as when regularization and dropout were applied to the initial model. These outcomes suggest that limiting the model's capacity helped mitigate overfitting, thereby improving its ability to generalize to new data.

The confusion matrices shown in Fig. 2 give insight into the best-performing model across the four classes (Exceeds, Fully Meets, Needs Improvement, and PIP). The improved ANN resulted in improved overall accuracy (96.81%) compared to the baseline model (93.62%). A significant improvement was observed in Class Exceeds, where the F1-score rose from 0.80 to 1.00, indicating better class discrimination. The remaining classes showed no serious change and retained their comparatively small F1-scores. The main reasons for this performance gap include the small sample sizes and extreme class imbalance, which make it harder for the model to generalize to underrepresented categories.

Class Exceeds: Precision = 1.00, Recall = 1.00, F1 = 1.00

Class Fully Meets: Precision = 0.99, Recall = 1.00, F1 = 0.99 (unchanged)

Class Needs Improvement: Precision = 0.67, Recall = 0.50, F1 = 0.57 (unchanged)

Class PIP: Precision = 0.50, Recall = 0.50, F1 = 0.50 (unchanged)

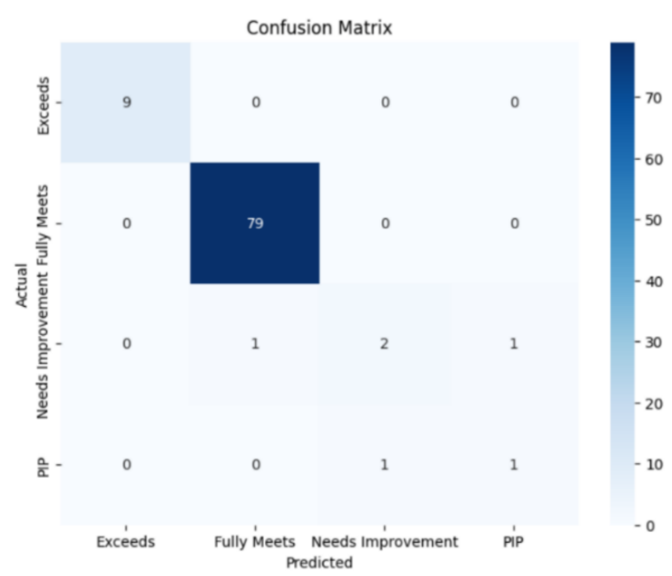

Fig. 2. Confusion matrices of HR metrics and Analytics dataset.

Key insights from the HR metrics and Analytics dataset:

- The observed improvement with feature extraction suggests that the extracted features provided relevant information and improved the model's predictive capability.

- In this dataset, PCA did not help construct effective predictive models. This analysis indicates that PCA may have led to the loss of important information and an oversimplification of the feature space.

- The combination of regularization and dropout proved effective in mitigating overfitting. In addition, applying feature selection improved model performance by achieving a better balance between model robustness and data quality.

- This dataset demonstrates stable behavior, as all evaluated models achieved consistently high accuracy, suggesting a relatively low noise level and reliable data patterns.

B. Results of the INX Dataset

The initial ANN model achieved an accuracy of 81.11%, which we set as the baseline model for comparison. The model showed strong predictive performance but was outperformed by some modified versions. This discrepancy suggested that the first model was not the best in terms of both feature relevance and overfitting control.

In another experiment, features with weak correlation to the target variable were removed, resulting in an accuracy improvement to 83.89%. Applying feature selection to a reduced-layer model yielded the same accuracy. This improvement probably reflected a decrease in feature-space noise and an increase in model generalization. However, the combination of feature selection, regularization, and dropout achieved an accuracy of 81.94%. This implied that adding more regularization constraints to this configuration reduced the advantages of feature removal.

When the new features were extracted with arithmetic operations (addition and multiplication) and added to the original features in the initial model, the ANN achieved an accuracy of 81.94%. Feature extraction was then applied to a reduced-layer model. Similarly, using feature extraction on a regularized model with dropout yielded the same accuracy. These outcomes showed that the extracted features were informative and improved the model's performance.

In contrast, using principal component analysis (PCA) reduced model accuracy to 79.72% compared to the initial model, suggesting that the dimensionality-reduction step may have removed informative features. We applied PCA to the model with fewer layers and to the model with regularization and dropout, both resulting in an accuracy of 79.17%. These outcomes demonstrated that PCA's dimensionality reduction was less effective for this dataset.

The ANN model's accuracy increased to 82.78% when the number of layers was reduced, implying that fewer layers improved generalization by reducing overfitting.

The combination of regularization and dropout resulted in the most noticeable improvement. This configuration achieved an accuracy of 85.28%, clearly higher than that of the initial model. These results showed that regularization helped mitigate overfitting, making the combination the most balanced in terms of complexity and generalization.

The confusion matrices in Fig. 3 provided more detail on the classification performance of the best model across three classes (Class 2, Class 3, and Class 4). The results showed that performance was relatively balanced, with Class 3 achieving the highest prediction accuracy. The total accuracy of the enhanced model increased from 81.11% to 85.28%, with the largest gains in Class 4 (F1 $\approx$ 0.75) and modest gains in Class 2 (F1 $\approx$ 0.67), while Class 3 remained at the highest level (F1 $\approx$ 0.91). The overall outcomes indicated that the suggested strategy improved classification performance across all categories and yielded balanced precision–recall trade-offs.

Class 2: Precision = 0.67, Recall = 0.67, F1 = 0.67

Class 3: Precision = 0.90, Recall = 0.91, F1 = 0.91

Class 4: Precision = 0.78, Recall = 0.72, F1 = 0.75

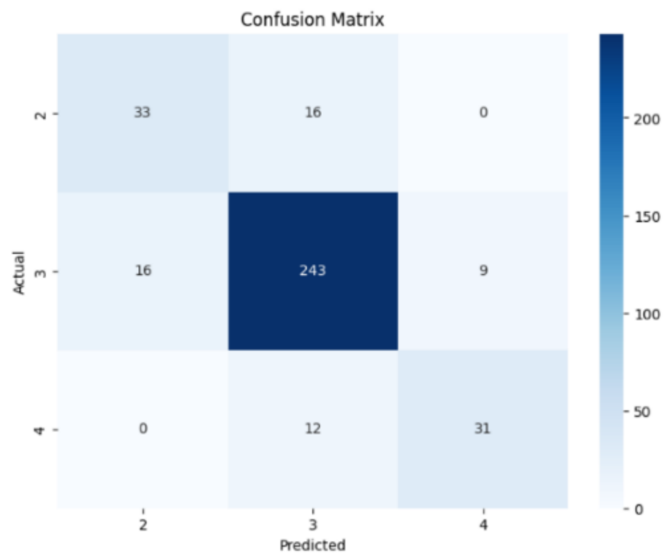

Fig. 3. Confusion matrices of INX dataset.

Key insights from the INX dataset:

- Moderate performance improvement was achieved after implementing feature selection. This suggested that moderately increasing the model's accuracy may be possible through focused optimization of the feature space, with removing low-correlation features being more effective than introducing new ones in this dataset.

- There appeared to be a consistent performance deficiency across all PCA-based models. This suggested that PCA might have removed too much potentially useful information and, therefore, cannot be the right approach for this dataset.

- The findings showed that the best methods for enhancing model performance are regularization and dropout. These techniques addressed overfitting in neural networks, a prevalent problem, by promoting stronger generalization to unseen data.

C. Results of the Employees' Performance Dataset

The results of this dataset indicate that the ANN model performs quite consistently using various preprocessing and parameter-tuning techniques. The initial accuracy of the model (86.57%) suggests the existence of natural discriminative patterns in the dataset. However, the other modified models were more accurate.

The accuracy increased to 87.85% after applying low-correlation-based feature selection on the initial model. Also, when applying feature selection to the model with lower layers, the same level of accuracy was obtained, suggesting that removing irrelevant features was enough to improve performance without adding model complexity. However, when feature selection was combined with regularization and dropout, the accuracy dropped to 86.39%, indicating that other regularization constraints might have diminished the marginal benefits of feature removal.

When the feature extraction technique was performed, the performance decreased slightly, with the accuracy going down to 86.39%. It is interesting to note that this result can be observed in a wide range of models, including ones that use feature extraction with fewer layers or include regularization and dropout. This means that the extracted features may have introduced redundancy and did not provide any additional discriminative information.

Likewise, applying Principal Component Analysis (PCA) to the initial model resulted in a more pronounced decline in accuracy to 83.72%. Furthermore, the accuracy decreased slightly to 86.05% after applying PCA to a model with fewer layers or to a model incorporating regularization and dropout, suggesting that dimensionality reduction in this dataset may have led to the loss of crucial information related to class separation.

The accuracy increased to 87.29% over the initial model when layers were reduced, which means that simplification of the architecture is somewhat beneficial.

Combining regularization and dropout gave the best overall result on this dataset with an accuracy of 88.16%. This result implies that these strategies are effective in reducing overfitting and improving the generalization of the model.

The confusion matrices in Fig. 4 provided comprehensive details for the best-performing model across 4 classes (Class Low, Class Medium-Low, Class Medium-High, and Class High). The model's predictive ability is good overall, especially for the Low and Medium-Low classes, which have high true-positive rates. However, there are discernible misclassifications between the Medium-High and nearby classes, indicating that this group has lower recall and class overlap.

Class Low: Precision = 0.88, Recall = 0.95, F1 = 0.91

Class Medium-Low: Precision = 0.83, Recall = 1.00, F1 = 0.91

Class Medium-High: Precision = 0.89, Recall = 0.75, F1 = 0.81

Class High: Precision = 0.95, Recall = 0.79, F1 = 0.87

Key insights from the employees' performance dataset:

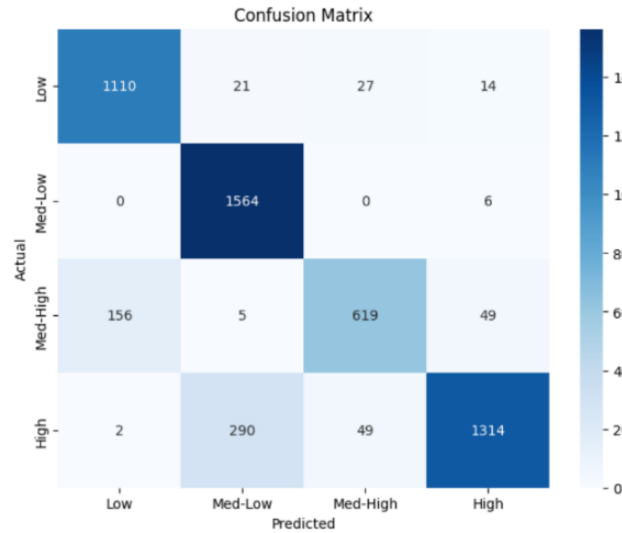

Fig. 4. Confusion matrices of the Employees' performance dataset.

- Principal component analysis (PCA)-based models in this dataset performed worse than the initial model since important features may have been lost during dimensionality reduction.
- Reducing the number of layers resulted in small improvements over the initial model, but the effect is limited compared to regularization and dropout.
- Overall, the results indicate that feature selection and model regularization are more useful for this dataset compared to dimensionality reduction or feature augmentation techniques.

D. Comparative Summary

We compared the experiment's outcomes across three datasets using the same methods. This comparative analysis helped reveal dataset-specific behaviors and find trends in technique performance. Table II shows a summary of results.

The comparison focuses on neural network-based techniques, serving as a methodological reference point for evaluating the usefulness of the proposed ANN model despite variances in application areas and network topologies. Table III compares our results with those of previous studies.

VI. Discussion

The baseline ANN results yield the following observations:

The HR metrics and Analytics dataset has the highest baseline accuracy and indicates strong predictive signals and a well-calibrated model. In contrast, the INX dataset is the most challenging dataset, as indicated by the lower baseline accuracy. The Employees' performance dataset is of moderate complexity and moderate baseline accuracy. The baseline is the measure of improvement. The HR metrics and Analytics dataset begins close to its ideal point, with little room for improvement. On the other hand, the INX dataset and the Employees' performance dataset provide a bigger scope for improvement.

TABLE II. EXPERIMENT CONFIGURATIONS AND ACCURACY RESULTS

Experiment Configuration		**HR metrics and analytics dataset Accuracy**	**INX dataset Accuracy**	**Employees' performance dataset Accuracy**
Initial ANN		93.62%	81.11%	86.57%
Data Preprocessing				
Feature Selection (Low Correlation)		**96.81%**	83.89%	87.85%
Feature Extraction (Addition/Multiplication)		95.74%	81.94%	86.39%
PCA		88.30%	79.72%	83.72%
Parameter Tuning				
Reduced Layers		94.68%	82.78%	87.29%
Regularization and Dropout		**96.81%**	**85.28%**	**88.16%**
Data preprocessing and parameter tuning				
Feature selection	reduced layers	**96.81%**	83.89%	87.85%
	regularization and dropout	95.74%	81.94%	86.39%
Feature extraction	reduced layers	95.74%	81.94%	86.39%
	regularization and dropout	95.74%	81.94%	86.39%
PCA	reduced layers	86.17%	79.17%	86.05%
	regularization and dropout	86.17%	79.17%	86.05%

TABLE III. RELATED WORK COMPARISON

Related work comparison	**Method**	**Results**
Our proposed work	We develop several neural network models that involve preprocessing steps for the data and fine-tuning their hyperparameters	Our accuracy is 90.08%, the average of the best-performing configurations obtained for each dataset
Study [14] Predict students' academic performance	They developed a neural network model with 11 input variables, 2 hidden layers, and 1 output layer	84.8%
Study [15] Forecast sophomores' performance	They created an artificial neural network model based on the multilayer perceptron topology	84.6%
Study [16] predict students' academic performance	They developed an artificial neural network model	93.81%
Study [18] predict students' academic performance	They applied several data mining classification methods	The ANN model achieved the best accuracy of 79%
Study [19] predict students' academic performance	They examined multiple machine learning techniques	The ANN model achieved the best accuracy of 77.04%
Study [21] proposed a thorough method for forecasting employee attrition	They compared and implemented multiple machine learning, ensemble, and deep learning approaches	The Feedforward Neural Network (FNN) achieved an accuracy of 83.93%

The feature selection applied to both the initial model and the model with a reduced number of layers leads to the following observations:

In both configurations, the model outperforms the baseline across all datasets (+3.19%, +2.78% and +1.28%, respectively). Moreover, it achieves the highest performance on the HR metrics and Analytics dataset, suggesting that removing low-correlation features helps the model to focus on the most informative features, especially in this dataset.

The feature selection applied to the model incorporating regularization and dropout results in the following observations:

The model performs better than the baseline on the HR metrics and analytics dataset and the INX dataset (+2.11% and +0.83%, respectively). However, in the Employees' performance dataset, a decrease in performance of -0.18% is observed. This slight decline raises the possibility of an over-regularization impact, which could restrict the model's applicability to this dataset.

Observations from results of feature extraction (addition/multiplication) applied to the initial model, the reduced-layer model, and the model that includes regularization and dropout:

The feature extraction shows consistent performance across the different model configurations and achieves results comparable to feature selection combined with the regularization and dropout, despite the difference in the methodology.

The principal component analysis (PCA) applied to the initial model provides the following observations:

PCA results in lower accuracy compared to the baseline across all datasets by (-5.32%, -1.39%, and -2.85%) respectively). The HR metrics and analytics dataset shows the largest performance loss. Thus, PCA ranks as the overall weakest technique for the evaluated datasets. One possible explanation for PCA's reduced performance is that it prioritizes variance preservation over class discrimination. As a result, principal components may retain high-variance information that is not necessarily relevant for distinguishing between performance categories. Consequently, some discriminative features may be suppressed during dimensionality reduction.

The principal component analysis (PCA) applied to the model with a reduced number of layers and to the model incorporating regularization and dropout yielded the following observations:

Compared to the baseline model, accuracy decreased for all of the datasets analyzed. The HR Metrics and analytics dataset showed the greatest decrease (-7.45%), followed by the INX dataset (-1.94%), while the Employees' Performance dataset showed only a slight decrease (-0.52%). These results suggest that PCA may have eliminated some relevant discriminative information despite the use of architectural simplification and reduction of overfitting, and therefore failed to improve the classification performance.

The reduced number of layers yielded the following observations:

Results show improvement compared to the baseline across all the datasets (+1.06%, +1.67%, and +0.72%, respectively). This outcome suggests that a simpler architecture may facilitate better generalization, possibly by mitigating overfitting.

The regularization and dropout yielded the following observations:

The approach outperformed the baseline across all datasets (+3.19%, +4.17%, and +1.59%, respectively). The most significant improvement was attained in the INX dataset, followed by the HR Metrics and Analytics dataset. These results suggest that regularization and dropout helped reduce overfitting and improve generalization as well as overall classification capabilities of the model.

The overall findings across all datasets reveal several consistent patterns. PCA-based dimensionality reduction resulted in lower classification accuracy, suggesting that some discriminative information was lost. In contrast, all datasets demonstrated consistent performance improvements when regularization and dropout were applied, highlighting their effectiveness in mitigating overfitting and promoting stronger generalization. Furthermore, feature selection consistently outperformed feature extraction and remained stable across multiple parameter settings, demonstrating its robustness and suitability for the classification tasks under investigation. Overall, the optimal configuration appears to depend on the inherent characteristics of each dataset rather than on a universally superior technique, as factors such as dimensionality, data complexity, feature interactions, and noise levels influence the effectiveness of a given approach.

Several previous studies have demonstrated the strong performance of tree-based models, such as Random Forest and Gradient Boosting, in HR analytics. However, the present study focuses on investigating how preprocessing strategies and architectural optimization influence ANN-based models.

A bar chart in Fig. 5 illustrates accuracy across all datasets and highlights overall performance trends.

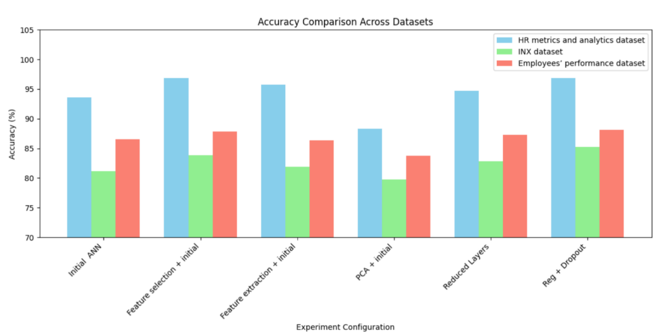

Fig. 5. Comparing accuracies across datasets.

VII. Conclusion and Future Work

This study demonstrated the effectiveness of artificial neural networks in predicting employee performance by systematically evaluating preprocessing techniques and model optimization strategies across three independent HR datasets. The results indicate that dropout and regularization consistently improved generalization performance, especially for datasets prone to overfitting. Feature selection also yielded stable performance gains, highlighting the importance of eliminating low-correlation features. On the other hand, predictive accuracy consistently decreased with PCA-based dimensionality reduction, suggesting a potential loss of discriminative information in the transformed feature space. Moderate improvements were achieved by simplifying the architecture through layer reduction, underscoring the importance of balancing model complexity with dataset characteristics. Across the assessed datasets, the top-performing configurations achieved accuracies ranging from 81% to 96%. The use of diverse datasets enhances the robustness of the findings and provides practical guidance for ANN-based predictive modeling in HR analytics. The results of this study can help data analysts and HR specialists select effective neural network configurations and preprocessing techniques for employee performance prediction systems. Specifically, organizations can prioritize feature selection, architectural simplification, and regularization strategies to improve model generalization and ensure robust predictive performance across different HR datasets. These insights enable the development of more reliable data-driven decision-support systems for workforce management and performance evaluation.

Although the results are encouraging, this study has some limitations. A single train-test split was used, and hyperparameter tuning was performed within predefined limits. Moreover, methods for mitigating class imbalance were not thoroughly investigated. Addressing these limitations could further improve model generalizability and robustness.

A. *Future Work*

Future work may extend this study in several directions. First, a more comprehensive hyperparameter search that accounts for differences in network architectures, regularization strengths, dropout rates, batch sizes, and learning rates may further improve prediction performance across datasets. Second, other deep learning architectures could be explored to compare their performance in HR analytics applications. Third, k-fold cross-validation would provide more reliable performance estimates across different data splits. Furthermore, advanced imbalance mitigation strategies, such as SMOTE and weighted loss functions, could be explored to address potential class imbalance issues. Lastly, deeper insights into the most significant variables in each dataset could be obtained through feature importance analysis using techniques such as SHAP or permutation-based methods.

Acknowledgment

The research was funded by the KAU Endowment (WAQF) at the King Abdulaziz University, Jeddah, Saudi Arabia. The authors would like to thank the WAQF and the Deanship of Scientific Research (DSR) for technical and financial support.

References

[1] F. Fallucchi, M. Coladangelo, R. Giuliano, and E. W. D. Luca, "Predicting employee attrition using machine learning techniques," *Computers*, vol. 9, no. 4, p. 86, Nov. 2020, doi: https://doi.org/10.3390/computers9040086.

[2] B. K. Kumari, V. M. Sundari, C. Praseeda, P. Nagpal, J. EP, and S. Awasthi, "Analytics-based performance influential factors prediction for sustainable growth of organization, employee psychological engagement, work satisfaction, training and development," *Journal for ReAttach Therapy and Developmental Diversities*, vol. 6, no. 8s, pp. 76–82, 2023.

[3] K. Patel, K. Sheth, D. Mehta, S. Tanwar, B. C. Florea, D. D. Taralunga, and R. Sharma, "Ranker: An ai-based employee-performance classification scheme to rank and identify low performers," *Mathematics*, vol. 10, no. 19, p. 3714, 2022.

[4] H. Sinha, "Predicting employee performance in business environments using effective machine learning models," *International Journal of Novel Research and Development*, vol. 9, no. 9, pp. 875–881, 2024, doi:https://doi.org/10.5281/zenodo.13771036.

[5] Z. Nayem and M. A. Uddin, "Unbiased employee performance evaluation using machine learning," *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 10, no. 1, p. 100243, 2024, doi:https://doi.org/10.1016/j.joitmc.2024.100243.

[6] M. Nasr, E. Shaaban, A. Samir, and M. Nasr, "A proposed model for predicting employees' performance using data mining techniques: Egyptian case study," *International Journal of Computer Science and Information Security (IJCSIS)*, vol. 17, no. 1, pp. 31–40, 2019.

[7] D. E. Asuquo, U. A. Umoh, F. B. Osang, and E. W. Okokon, "Performance evaluation of c4.5, random forest and naïve bayes classifiers in employee performance and promotion prediction," *African Journal of Management Information System*, vol. 2, no. 4, pp. 41–55, 2020.

[8] R. Jayadi, H. M. Firmantyo, M. T. Dzaka, M. F. Suaidy, and A. M. Putra, "Employee performance prediction using naïve bayes," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 8, no. 6, pp. 3031–3035, 2019, doi:https://doi.org/10.30534/ijatcse/2019/59862019.

[9] R. Obiedat and S. A. Toubasi, "A combined approach for predicting employees' productivity based on ensemble machine learning methods," *Informatica*, vol. 46, no. 5, 2022, doi:https://doi.org/10.31449/inf.v46i5.3839.

[10] M. R. Hasan, R. K. Ray, and F. R. Chowdhury, "Employee performance prediction: An integrated approach of business analytics and machine learning," *Journal of Business and Management Studies*, vol. 6, no. 1, pp. 215–219, 2024, doi: https://doi.org/10.32996/jbms.2024.6.1.14.

[11] L. G. Tanasescu, A. Vines, A. R. Bologa, and O. Vîrgolici, "Data analytics for optimizing and predicting employee performance," *Applied Sciences*, vol. 14, no. 8, p. 3254, 2024, doi:https://doi.org/10.3390/app14083254.

[12] M. Yağcı, "Educational data mining: Prediction of students' academic performance using machine learning algorithms," *Smart Learning Environments*, vol. 9, no. 1, p. 11, 2022.

[13] N. Aslam, I. Khan, L. Alamri, and R. Almuslim, "An improved early student's academic performance prediction using deep learning," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 16, no. 12, pp. 108–122, 2021.

[14] E. T. Lau, L. Sun, and Q. Yang, "Modelling, prediction and classification of student academic performance using artificial neural networks," *SN Applied Sciences*, vol. 1, no. 9, p. 982, 2019.

[15] S. A. Naser, I. Zaqout, M. A. Ghosh, R. Atallah, and E. Alajrami, "Predicting student performance using artificial neural network: In the faculty of engineering and information technology," *International Journal of Hybrid Information Technology*, vol. 8, no. 2, pp. 221–228, 2015.

[16] H. Chavez, B. Chavez-Arias, S. Contreras-Rosas, J. M. Alvarez-Rodriguez, and C. Raymundo, "Artificial neural network model to predict student performance using nonpersonal information," *Frontiers in Education*, vol. 8, p. 1106679, 2023, doi: https://doi.org/10.3389/feduc.2023.1106679.

[17] Y. Baashar, G. Alkawsi, A. Mustafa, A. A. Alkahtani, Y. A. Alsariera, A. Q. Ali, W. Hashim, and S. K. Tiong, "Toward predicting student's academic performance using artificial neural networks (anns)," *Applied Sciences*, vol. 12, no. 3, p. 1289, 2022, doi: https://doi.org/10.3390/app12031289.

[18] H. A. Mengash, "Using data mining techniques to predict student performance to support decision making in university admission systems," *IEEE Access*, vol. 8, pp. 55 462–55 470, 2020, doi: https://doi.org/10.1109/ACCESS.2020.2981905.

[19] H. Altabrawee, O. A. J. Ali, and S. Q. Ajmi, "Predicting students' performance using machine learning techniques," *Journal of University of Babylon for Pure and Applied Sciences*, vol. 27, no. 1, pp. 194–205, 2019.

[20] I. Setiawan, S. Suprihanto, A. C. Nugraha, and J. Hutahaean, "Hr analytics: Employee attrition analysis using logistic regression," *IOP Conference Series: Materials Science and Engineering*, vol. 830, p. 032001, 2020, doi: https://doi.org/10.1088/1757-899X/830/3/032001.

[21] M. Nandal, V. Grover, D. Sahu, and M. Dogra, "Employee attrition: Analysis of data driven models," *EAI Endorsed Transactions on Internet of Things*, vol. 10, 2024.

[22] N. Mansor, N. S. Sani, and M. Aliff, "Machine learning for predicting employee attrition," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 11, 2021, doi: https://doi.org/10.14569/IJACSA.2021.0121149.

[23] P. K. Jain, M. Jain, and R. Pamula, "Explaining and predicting employees' attrition: A machine learning approach," *SN Applied Sciences*, vol. 2, no. 4, p. 757, 2020, doi: https://doi.org/10.1007/s42452-020-2519-4.

[24] M. Pratt, M. Boudhane, and S. Cakula, "Employee attrition estimation using random forest algorithm," *Baltic Journal of Modern Computing*, vol. 9, no. 1, pp. 49–66, 2021, doi: https://doi.org/10.22364/bjmc.2021.9.1.04.

[25] R. Punnoose and P. Ajit, "Prediction of employee turnover in organizations using machine learning algorithms: A case for extreme gradient boosting," *International Journal of Advanced Research in Artificial Intelligence (IJARAI)*, vol. 5, no. 9, pp. 22–26, 2016.

[26] G. M. Díaz, J. J. G. Hernández, and J. L. G. Salvador, "Analyzing employee attrition using explainable ai for strategic hr decision-making," *Mathematics*, vol. 11, p. 4677, 2023, doi: https://doi.org/10.3390/math11224677.

[27] F. K. Alsheref, I. E. Fattoh, and W. M. Ead, "Automated prediction of employee attrition using ensemble model based on machine learning algorithms," *Computational Intelligence and Neuroscience*, vol. 2022, p. 7728668, 2022.

[28] S. Arqawi, M. Rumman, E. Zitawi, B. Abunasser, and S. A. Abu-Naser, "Predicting employee attrition and performance using deep learning," *Journal of Theoretical and Applied Information Technology*, vol. 100, no. 21, pp. 6526–6536, 2022.

[29] A. Raza, K. Munir, M. Almutairi, F. Younas, and M. M. S. Fareed, "Predicting employee attrition using machine learning approaches," *Applied Sciences*, vol. 12, no. 13, p. 6424, 2022.

[30] S. Al-Darraji, D. G. Honi, F. Fallucchi, A. I. Abdulsada, R. Giuliano, and H. A. Abdulmalik, "Employee attrition prediction using deep neural networks," *Computers*, vol. 10, no. 11, p. 141, 2021, doi: https://doi.org/10.3390/computers10110141.