Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

An Explainable and Fair TabNet Framework for Remote Employee Performance Prediction Using Multidimensional Workplace Data

Zihe Zhu

zhuzihe427@outlook.com

New York University

XiaoQian Bian

University of Macau

Sihui Yu

Shenzhen Polytechnic University

Article

Keywords: Explainable artificial intelligence (XAI), remote employee performance prediction, TabNet, workforce analytics, SHAP explainability, fairness-aware machine learning, deep tabular learning

Posted Date: July 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10179003/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Scientific Reports on October 25th, 2026.
See the published version at https://doi.org/10.1038/s41598-026-76297-z.

An Explainable and Fair TabNet Framework for Remote Employee Performance Prediction Using Multidimensional Workplace Data

Zihe Zhu[1,*], XiaoQian Bian[2], and Sihui Yu[3]

[1]Human Resource Management and Development, School of Professional Studies, New York University, New York, NY 10003, USA; zhuzihe427@outlook.com
[2]Institute of Analytical Psychology, City, University of Macau, Macau, China, 999078, bian954969366@outlook.com
[3]Mental Health Center, Shenzhen Polytechnic University, Guangdong, 51805, China, yusihui786@outlook.com
Corresponding author: zhuzihe427@outlook.com

Abstract: The prediction of performance by remote employees has become more important as organizations try to make their workforce data-driven when distributed. Existing approaches, however, tend to be non-interpretable, non-robust, and non-fair. In this study, an Explainable TabNet Framework for predicting employee performance remotely from a workforce dataset with 100,000 employee records is proposed. This framework incorporates attention-based deep tabular learning, explainability through SHAP values, robustness evaluation via cross-validation, statistical significance testing, and fairness auditing of subgroups in this field. Multiple regression analysis metrics were computed for the predictive performance of the model and compared with four other baselines: Linear Regression, Ridge Regression, Random Forest, Extra Trees, Gradient Boosting, Model XGBoost and LightGBM. The proposed framework outperformed all baseline models with better predictive performance, as evident in the R^2 score, RMSE, MAE, and PCC values, which were 0.9880, 0.1549, 0.1223, and 0.9956, respectively. The most significant predictors of employee performance as determined by ex-ante analysis were employee satisfaction, monthly salary, project involvement, participation in training, and tenure in the organization. Cross-validation and statistical testing ensured model robustness and significance, and fairness testing confirmed that the prediction was similar across groups of demographic and organizational individuals. The Explainable TabNet Framework is an accurate, interpretable, robust, and fair solution for employee performance prediction in remote work, which will enable transparent workforce analytics and human resource decision making based on evidence.

Keywords: Explainable artificial intelligence (XAI); remote employee performance prediction; TabNet; workforce analytics; SHAP explainability; fairness-aware machine learning; deep tabular learning

1. Introduction

Today, fast-growing digital technologies, cloud computing systems, and ways to communicate collaboratively are making workplaces different [1]. An increasing number of organisations are shifting to remote and hybrid working patterns to deliver increased flexibility, contented staff, cut down on infrastructure spend, and keep companies running [2]. These workplace models come with great Organizational and Employee benefits but also bring significant challenges of workforce monitoring, productivity assessment, performance evaluation, and managerial decision making [3]. The difficulty in assessing employee performance objectively and consistently remotely is greater than in traditional workplaces because of less direct supervision [4]. Thus, seemingly more and more organisations are turning to workforce analytics and Artificial Intelligence (AI) technologies to help guide evidenced-based performance evaluation and strategic human resource management.

Staff performance is one of the major factors determining organizational productivity, innovation, competitiveness, and sustainability [5]. As organizational success becomes increasingly dependent on the performance of employees [6], [7], it is evident that they make a major contribution through better efficiency, service, customer satisfaction, and business results. The factors that impact employee performance, though, are varied and include demographic, behavioral, organisational, and work factors such as employee satisfaction, compensation and workload;

participation in training; project involvement; professional experience, educational background; organisational tenure; and workplace flexibility [8], [9]. All these variables are interwoven and nonlinear, which makes it difficult to rely on traditional performance assessment methodologies.

Most established systems for employee performance evaluation use manual appraisal, key performance indicators (KPIs), annual performance reviews, and supervisor assessments [10], [11]. Although these approaches provide management insights, they can be subjective, evaluation-biased, inconsistent, and scalable [12], [13]. These restrictions are especially pronounced in large companies, where employees are found at a distance and in different places, making it difficult to observe their work [14]. This leads to a growing demand for smart prediction systems that can reliably forecast the performance of employees and offer objective and scalable decision-support mechanisms that can be seen and understood.

Machine learning is proving to be a tremendous remedy for tackling these issues. With machine learning algorithms automatically detecting patterns in large amounts of organizational data, these computer programs can make accurate predictions of employee productivity, engagement, retention, and performance results [15], [16], [17]. Common machine-learning models, such as Linear Regression, Ridge Regression, Random Forests, Gradient Boosting, XGBoost and LightGBM, have exhibited good results for different workforce analytics applications [18], [19], [20], [21]. However, there are currently two significant limitations to many predictive models. One is that standard algorithms are unable to adequately describe non-linear relationships between various workforce factors that may be non-homogeneous. Second, many predictive systems are black-box models that offer little insight into what factors affect the outcome of the predictions.

As AI-based decision-support systems become more widespread, it has become a growing concern to ensure that they are applied in a transparent and accountable manner with trust and fairness [22]. In an organizational context, predictive suggestions can have a direct effect on promotion decisions, pay structures, workforce planning, employee career development, and performance management policies. Moreover, managers and stakeholders are much more needy of systems that are able to effectively predict well in addition to being able to interpret what justifications have guided their decisions [23], [24]. Explainable Artificial Intelligence (XAI) stands out as one of the most pressing and promising research horizons that seeks to increase the transparency of models and promote humans-in-the-loop decision-making [25], [26], [27]. Feature-level interpretation and prediction explanations contribute to trust and accountable use of AI systems, as well as easy adoption of AI systems for HR use.

The understanding of structured tabular data is achieved using a special architecture designed recently specifically for tabular data, namely, deep learning. Among these methods, models based on deep tabular learning with an attention mechanism have received a great deal of attention owing to their excellent prediction power and built-in interpretability [28], [29]. These sequential architectures are different from conventional neural networks that process all the features at once. These architectures have sequential attention mechanisms, which means that they compute which features are the most informative for their current decision at each decision step [30], [31], [32]. This attention-based learning approach allows high-efficiency modeling of complex interactions between multiple, non-linear variables, as well as explanations at the feature level, concerning the model behavior [33]. The quality and nature of these features make tabular learning with attention specifically well suited to workforce analytics [34], [35], which leverages a variety of demographic, organizational, behavioral, and remote-work variables that vary among different employees.

Although workforce analytics has made significant progress, there are still several challenges for research that remain unanswered. Many studies have a limited interest in the employee-performance prediction task compared with employee attrition prediction or employee turnover analysis or categorization. Second, traditional predictive systems often focus on the prediction (accuracy) of the risk of negative events with little explanatory value, which hinders their use in the context of organizations. Third, the number of studies that examine predictive performance, explainability, robustness, statistical significance, and fairness in a single framework is limited. Finally, there continue to be relatively underexplored applications of deep tabular learning architectures that are suitable for large-scale remote applications to employee performance prediction.

In this work, we introduce the Explainable TabNet Framework for Remote EPP, which improves on these limitations by leveraging a large-scale public dataset of 100,000 employee records. The proposed framework presents a unified predictive modeling pipeline that incorporates attention-based deep tabular learning, explainability via SHAP, extensive baseline comparisons, robustness validation, statistical significance testing, and fairness auditing. This framework seeks to deliver a dependable and attainable AI remedy for present-day labor force analytics and evidence-based HRM.

Multiple regression metrics like coefficient of determination (R^2), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), Median Absolute (MAE), Explained Variance Score and Pearson correlation coefficient are applied to the proposed framework. To evaluate the robustness of the models, five-fold cross validation was used and to ensure any statistical significance, paired t-tests, Wilcoxon signed-ranked, Friedman tests, and Nemenyi post hoc analysis were performed. In addition, fairness is evaluated by gender, department, education level, and organizational subgroup, which guarantees fair predictive performance.

The results of the experiment show the effectiveness of the proposed framework, with an R2 of 0.9880, RMSE of 0.1549, and Pearson correlation coefficient of 0.9956 for the new test data. The explanatory analysis determines that among the factors influencing employee performance, employee satisfaction, employees' monthly salaries, involvement in the project, participation in training, and organizational tenure are the most influential. Furthermore, the robustness validation, statistical testing, and fair assessment of the methodology demonstrate that the approach suggested in this work is reliable, stable, and ethically appropriate for workforce analytics applications.

The following research questions were addressed in this study:

RQ1: Can an attention-based deep tabular learning framework really predict the performance of remote employees based on demographic, organizational, behavioral, and workplace-related data?
RQ2: Can Explainable TabNet Framework outperform popular machine learning models (Linear Regression, Ridge Regression, Random Forest, Extra Trees, Gradient Boosting, XGBoost, and LightGBM) with respect to predictive performance?
RQ3: What are the most important attention-based features and explainable AI-based features for predicting employee performance?
RQ4: Is the proposed framework valid, statistically sound, and generalizable to various validation settings?
RQ5: Does the suggested scheme preserve a fair and equitable predictive performance among demographic and organizational groups of employees?

Based on the objectives of this study and the characteristics of the proposed framework, the following hypotheses were formulated:

H1: Prospect refers to the proposed Explainable TabNet Framework, which successfully predicts the performance of an employee and explains a significant portion of the variation in the obtained performance scores.

$$H_1: R^2_{TabNet} > 0.90$$

H2: The proposed Explainable TabNet Framework performs significantly better in prediction results compared to the conventional machine learning and ensemble learning algorithms.

$$H_2: R^2_{TabNet} > R^2_{Baseline}$$

where $R^2_{Baseline}$ denotes the coefficient of determination achieved by competing regression models.

H3: The most influential determinants of employee performance include employee satisfaction, monthly salary, training, involvement in projects, and organizational tenure.

$$H_3: Importance_{Key\ Features} > Importance_{Other\ Features}$$

as determined by TabNet attention scores and SHAP values.

.

H4: The proposed framework achieves a statistically significant improvement in predictive performance.

$$H_4: p < 0.05$$

for paired t-test, Wilcoxon signed-rank test, Friedman test, and Nemenyi post-hoc analysis.

H5: The proposed framework preserves strong predictive capability over various folds of cross validation.

$$H_5: \sigma_{CV} \to 0$$

where data include the variation of predictive performance across validation folds, where σ_{CV} is the variance.

H6: The proposed framework has very low demo and organization bias among the subgroups of employees.

$$H_6: \ | MAE_i - MAE_j | < \delta$$

where MAE_i and MAE_j denote prediction errors for the subgroups and when there is a small amount of fairness δ.

The major contributions of this study are summarized as follows:

- First, an Explainable TabNet Framework is developed for continuous remote employee performance prediction using large-scale workforce data.
- Second, attention-based deep tabular learning was integrated with SHAP-based explainability and native feature-importance analysis to improve model transparency.
- Third, comprehensive benchmarking is performed against multiple traditional and ensemble-based machine learning algorithms.
- Fourth, robustness validation, statistical significance testing, and fairness auditing were incorporated to ensure reliable and responsible AI deployment.
- Finally, actionable insights regarding the key determinants of employee performance are generated to support evidence-based human resource management and organizational decision-making.

2. Literature Review

2.1 Employee Performance Prediction in Remote Work Environments

According to Ortynska [6], remote and hybrid work are adopted, one-way organizations measure productivity and performance changes. Zapata, L., et al. [36], what makes remote working different from the regular way of work is that it has extra considerations including workplace flexibility, effectiveness, work-life balance, digital collaboration and worker autonomy. Pillai, S. V. and J. Prasad [37], these factors result in complexity and interactions that impact worker productivity, making it increasingly difficult to assess performance. Md Harun-Or-Rashid, M. and K. Sai Praveen [38], as a result, there is an increasing trend towards data-driven human resource management to interpret employee behaviour and forecast their performance outcomes.

Varghese, S., et al. [39], workforce analytics research in recent years has shown that performance is a multifaceted mix of workplace, behavior, and organizational factors as well as demographic factors. Hariani, M. and R. Mardikaningsih [40], employee satisfaction, employee compensation, workload, training participation, project involvement, work organization tenure, educational level, and workplace flexibility always appear as important influences on productivity. Hasibuan, M. S. and R. R. N. Fikri [41], correlations between these variables are complex and nonlinear, and badly affected by the integration of the variables in the model, making traditional statistical approaches inadequate. Addressing these challenges has inspired the taking toward the trend of advanced machine learning strategies that may model complicated workforce dynamics.

2.2 Machine Learning Approaches for Employee Performance Analytics

Razali et al. [42] stated that the potential of assembling data in huge volumes within organizational surroundings and utilizing machine learning to uncover hidden trends is an increasingly critical part of contemporary HR analytics. Song, Q. C., et al. [43], manually created predictive models, such as linear regression and ridge regression, give a clear interpretation but have difficulty explaining non-linear relationships between employee characteristics and performance indicators. Maharani, D., et al. [44], to overcome these advantages, ensemble-learning algorithms like Random Forest, Extra Trees, Gradient Boosting, XGBoost, and LightGBM have been utilized greatly within the forecasting of the workforce jobs.

Gustriansyah, R., et al. [45], in general, tree-based methods give better predictive results when they are able to model nonlinear decision boundaries and feature interactions. Palivela, L. H., et al. [46], Random Forest relies on

"bootstrapping" aggregation and boosters attempt to sequentially minimize predicting errors to increase overall accuracy. Nwoke, J., [47], these approaches sometimes outperform traditional statistical methods, but they rely on features that are designed manually and are often unable to yield a deep understanding of the underlying decision-making process. Buyuktanir, B., et al., [48], the size and intricacy of workforce datasets are growing, and options for learning architectures that are more adaptable and interpretable are becoming more apparent.

2.3 Explainable Artificial Intelligence in Workforce Analytics

Chen, Z. [49], While the adoption of predictive analytics in human-resource management has been growing, there are key questions of transparency, fairness, and accountability. Guion, R. M. [50] predictive models could impact staff assessments, promotions, pay, and other career growth opportunities when making organizational decisions. Mathew, D. E., et al. [51] stakeholders are therefore increasingly demanding the use of machine-learning systems that make accurate predictions while also providing reasons that are understandable for these decisions.

One effective answer to this challenge is Explainable Artificial Intelligence (XAI). Kalasampath, K., et al. [52], XAI techniques can be used to understand what causes the model to make a prediction and to determine how important the different features are on a global and individual basis. Biecek and Burzykowski [53] used a global explanation to explain the overall behavior of a predictive model, and a local explanation to explain the factors that influence the behavior of the model for a particular prediction. Ayanponle, L., et al. [54] AI Explainability is a feature that helps gain trust in workforce analytics, aids managers in decision-making, and enhances the feasibility of using predictive models.

However, recent research has shown that factors such as employee satisfaction, offers, staff development, job loading, project involvement, and expertise with companies are often key determinants of employee performance [55]. Implementing explainability techniques then allows companies to obtain actionable and predictive insights from their workforces.

2.4 Attention-Based Deep Learning for Tabular Data

Learning from tabular data has been challenging, despite the success of deep learning in other fields such as computer vision and natural language processing [35]. Somvanshi, S., et al. [56], traditional deep learning is not able to outcompete ensemble-learning approaches with tabular data for multi-type (numerical and categorical) data sets. This restriction has inspired the design of special deep-learning architectures, especially for tabular learning.

Gorishniy, Y., et al. [57], an interesting recent development in this regard is attention-based tabular learning models. These architectures learn the most relevant features in the process of learning, as opposed to using all predictors. Ma et al. [58], attention mechanisms allow allocation of computation only towards those attributes that are informative, increasing both the prediction accuracy and the interpretability of the running algorithm. We see these kinds of capabilities being especially useful in workforce analytics applications, where performance outcomes are a result of many interdependent factors such as demographic attributes, workplace behaviors, organizational factors, and remote work enactment.

Yiyuan, Z., et al. [59], attention-based models have shown good promise among recently developed tabular-learning models in terms of retaining the interpretability of feature selection and capturing nonlinear feature interactions. They are well suited for jobs such as staff-member performance forecasting using huge distributed datasets containing various types of data.

2.5 Research Gap and Study Motivation

Although great progress has been made in workforce analytics, some significant gaps in the research are clear. There are two points to note here: First, many studies are based on employee turnover, attrition, or employee performance in a category and not on the estimation of employee performance on a continuum. As a result, relatively limited studies focus on high-precision performance prediction via regression, which can serve to inform comprehensive workforce evaluation.

Second, current predictive models often focus on predictive outcomes but have low interpretability. This opacity makes practical use in organizational settings difficult when managers need an explanation for the recommendations the model makes. Third, many workforce analytics studies focus on the predictive aspect only and do not systematically explore robustness or statistical significance, nor explore subgroup fairness, still less "responsible AI" concerns.

Finally, the study of using attention-based deep tabular learning for large-scale remote worked prediction of performance is still under exploration. Natural papers tend to stand apart from other projects regarding the integration of all of these four elements (predictive modeling, explainability analysis, robustness validation, and statistical testing) and fairness auditing as a unified analytical body of work.

3. Materials and Methods

The overall research workflow is illustrated in Figure 1.

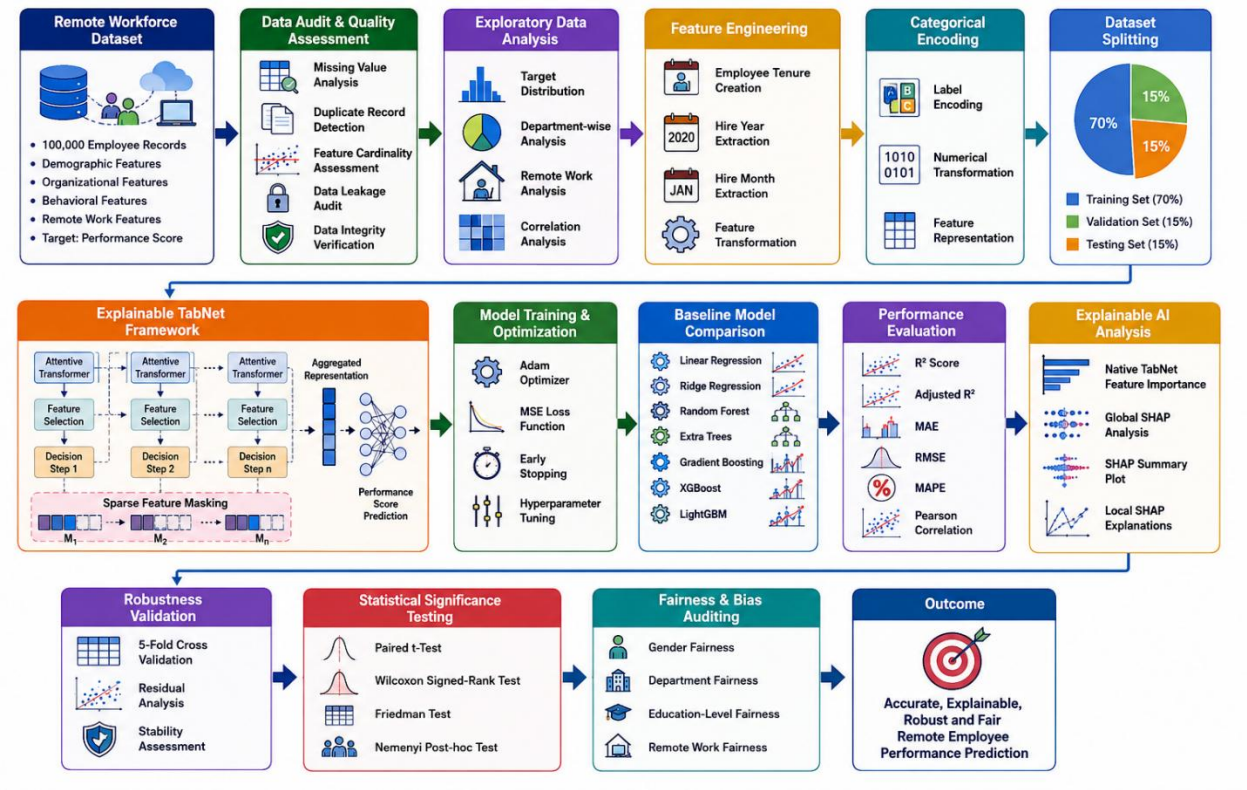


Figure 1. Explainable TabNet Framework for remotely predicting employee performance (Materials and Methods). The various steps in the workflow entail the acquisition of the dataset, data audit and quality checking, exploratory data analysis, feature engineering, Categorical Encoding, dataset splitting, optimization of the TabNet-based attention mechanism, training and optimization of models, comparison of baseline models, performance evaluation, Explainable AI analysis using SHAP, checking of robustness and statistical significance, and an audit on the fairness of the predictions leading to accurate, explainable, robust, and fair prediction of employee performance.

3.1 Dataset Description

This study considered a large employee performance dataset (of 100,000 observations) containing several demographic, organizational, behavioral, and remote work-related features [60]. Let the data be denoted by

$$D = \{(x_i, y_i)\}_{i=1}^{N} \quad (1)$$

where x_i, denotes the label for class i and the performance score to be achieved is denoted by y_i; the number of all observations is denoted by N. The goal was to obtain a predictive function:

$$\hat{y} = f(x) \tag{2}$$

Prediction of an employee's performance based on multidimensional data on his/her job.

3.2 Data Audit and Quality Assessment

A data quality audit was performed to ensure that the datasets were reliable before developing the model. Missing values were coded as. The missing values are marked as follows:

$$M_j = \frac{N_{nmissing,j}}{N} \times 100 \tag{3}$$

Records were duplicated by using

$$D_{dup} = \sum_{i=1}^{N} d_i \tag{4}$$

Cardinality of a feature was further explored via

$$U_j = |Unique(X_j)| \tag{5}$$

Differentiate between numerical and categorical variables. The audit validated that the data was properly cleaned so it could be used in the predictive modelling process.

3.3 Data Leakage Audit

To ensure that the information is not leaked, numerical predictors were tested by Pearson correlation:

$$r = \frac{\sum_{i=1}^{N}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{N}(y_i-\bar{y})^2}} \tag{6}$$

The conditional target means used to assess categorical variables were as follows:

$$\mu_c = \frac{1}{n_c} \sum_{i\ \varepsilon\ c} y_i \tag{7}$$

Furthermore, the Hire Date feature was investigated for temporal leakage and manipulated into meaningful features of tenure prior to training the models.

3.4 Exploratory Data Analysis

Data were subjected to Exploratory Data Analysis (EDA) to explore data distributions and to identify relationships with employees' performance. The desired distribution is determined by

$$P(Y = y) = \frac{n_y}{N} \tag{8}$$

However, for departments, performance was calculated as

$$\bar{y}_d = \frac{1}{n_d} \sum_{i=1}^{n_d} y_i \tag{9}$$

The impact of the frequency of remote work was further tested by analyzing the results. The effect of remote work frequency was analyzed in the following ways:

$$E[\,Y \mid R = r\,] = \frac{1}{n_r} \sum_{i=1}^{n_r} y_i \tag{10}$$

Pearson's correlation was used to evaluate the relationship between employee satisfaction and performance:

$$r_{SY} = \frac{\sum_{i=1}^{N}(S_i - \bar{S})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{N}(S_i - \bar{S})}\sqrt{\sum_{i=1}^{N}(Y_i - \bar{Y})}} \quad (11)$$

These analyses were used to gain preliminary insight into factors that affect employee productivity and to advance further model development.

3.5 Feature Engineering

To determine relevant information from the given attributes and obtain better prediction capabilities, feature engineering was performed. The Hire Date variable was separated into three temporal variables: Hire Year, Hire Month, and Employee Tenure. Employees' tenure is defined as follows:

$$T_i = Date_{ref} - Date_{hire,i} \quad (12)$$

where $Date_{ref}$, denotes specific dates. Denotes particular $Date_{hire,i}$ represents the hiring date of employee i. The final feature space can be expressed as

$$X^* = \{X, HY, HM, T\} \quad (13)$$

where HY refers to hire year, hire month (HM), tenure (T). These engineered variables also include extra data on employee experience and exposure to the organization.

3.6 Data Encoding

Categorical variables were encoded into numeric types to be used for deep learning. Allow categorical feature to have K different categories:

$$C = \{c_1, c_2, \dots, c_K\} \quad (14)$$

Label encoding was applied according to

$$f_{ck} = k - 1 \quad (15)$$

Every word in each text was labeled as a question word in accordance with where $k \in \{1,2, \dots, K\}$. Therefore, category information was encoded with consecutive integer codes, and each category's code was different, reflecting category information. Learning embedding representations were then used to use the encoded variables with the TabNet architecture.

3.7 Data Splitting Strategy

For unbiased evaluation, the dataset was divided into three subsets: training, validation, and testing sets. Let

$$D = D_{train} \cup D_{valid} \cup D_{test} \quad (16)$$

subject to

$$D_{train} \cap D_{valid} = D_{train} \cap D_{test} = D_{valid} \cap D_{test} = \emptyset \quad (17)$$

A 70:15:15 split ratio was adopted such that

$$|D_{train}| = 0.70N, |D_{valid}| = 0.15N, \ |D_{test}| = 0.15N \quad (18)$$

The training subset provided the parameters for general training and the validation subset for hyperparameter optimization and early stopping, while the testing subset provided the parameters for the final performance evaluation.

3.8 Proposed Explainable TabNet Framework

In this study, TabNet attention-based deep learning architecture, designed for tabular data, was used. Given an input feature vector,

$$x \in R_d \tag{19}$$

TabNet iteratively filters out informative features via selective feature masks and an attentive transformer. The Feature mask is produced at decision step t, which is given by

$$M_t = Sparsemax(P_t \odot h(a_t)) \tag{20}$$

Where P_t denotes previous feature importance, a_t represents the attention representation, and $\odot$ by element-wise multiplication. These selected features are then calculated as follows:

$$x_t = M_t \odot x \tag{21}$$

All outputs from the decision steps are aggregated to determine the final output.

$$\hat{y} = \sum_{t=1}^{T} d_t \tag{22}$$

where d_t "denotes" the contribution of the decision at step t. To promote interpretability, sparsity regularization was applied to the learning process as follows:

$$L = L_{MSE} + \lambda_s L_{sparse} \tag{23}$$

where L_{MSE} denotes the prediction loss, L_{sparse} is the regularization parameter. The sparsity penalty is represented by the term λ_s. Set the strength of regularization. This mechanism allows TabNet to achieve a high predictive performance while maintaining its transparent feature-selection behavior.

3.9 Hyperparameter Configuration

The success of the suggested TabNet framework relies on the correct choice of hyperparameters. The hyperparameters used were the learning rate, batch size, virtual batch size, sparsity coefficient, weight decay, and patience of early stopping. The generated model parameters were optimized based on

$$\theta_{t+1} = \theta_t - \eta \nabla_\theta L(\theta_t) \tag{24}$$

where η is the learning rate and $L(\theta_t)$ is the function to be optimized. To obtain good generalization, regularization by weight decay has been added as

$$L_{reg} = L + \lambda_w \sum_{k=1}^{P} \theta_k^2 \tag{25}$$

where λ_w weight-decay coefficient is denoted as α and there are P trainable parameters. To stop the training when the validation performance did not increase for a given amount of patience, early stopping was used.

3.10 Model Training Procedure

The training set was preprocessed, the parameters of the model were optimized on this subset, and the validation subset was used to monitor convergence and overfitting. In this approach $k = 1$, and the prediction loss was calculated for each mini-batch of size B with the Mean Squared Error (MSE):

$$L_{MSE} = \frac{1}{B} = \sum_{i=1}^{B} (y_i - \hat{y_i})^2 \tag{26}$$

The objective function is the complete prediction loss + sparsity regularization:

$$L_{total} = L_{MSE} + \lambda_s L_{sparse} \tag{27}$$

Adam optimizer was used to optimize model parameters with a method based on updating parameters according to:

$$\theta_{t+1} = \theta_t - \eta \frac{\widehat{m_t}}{\sqrt{\widehat{v_t}} + \epsilon} \quad (28)$$

Where $\widehat{m}_t$ and $\hat{v}_t$ first and second moment estimates are okay, as denoted by denoting the bias-corrected estimation. The most successful model (based on its lowest validation error) was selected for the final testing model.

3.11 Evaluation Metrics

Multiple regression metrics were used to evaluate the model performance and provide a complete picture of the predictive accuracy and robustness. The coefficient of determination (R2) was computed as follows:

$$R_2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \widehat{y_i})^2}{\sum_{i=1}^{N}(y_i - \hat{y})^2} \quad (29)$$

The errors made during the prediction were calculated using the Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE):

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |y_i - \widehat{y_i}| \quad (30)$$

$$RMSE = \frac{1}{N}\sqrt{\sum_{i=1}^{N}(y_i - \widehat{y_i})^2} \quad (31)$$

Other evaluation metrics used were Mean Absolute Percentage Error (MAPE), Median Absolute Error (MAE), Explained Variance Score (EVS), and Adjusted R-squared and Pearson Correlation (r). These measures collectively measure aspects of prediction accuracy, amount of variance explained, magnitude of errors, and level of agreement between predicted and actual employee performance ratings.

3.12 Explainable Artificial Intelligence Framework

To provide transparency and trustworthiness, an Explainable Artificial Intelligence (XAI) framework was incorporated into the proposed model. Feature importance scores were first obtained using the native attention masks of TabNet. The global importance of feature j was computed as:

$$I_j = \frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} M_{(t,j)}^{(i)} \quad (32)$$

where $M_{(t,j)}^{(i)}$, represents the attention weightage to feature j.

SHapley Additive exPlanations (SHAP) was used to obtain model-agnostic explanations. Form the Shapley value of feature j as:

$$\phi_j = \sum_{S \subseteq F\{j\}} \frac{|S|!(|F|-|S|-1)!}{|F|!} \; [f(S \cup \{j\}) - f(S)] \quad (33)$$

The most influential factors at the global level (across the entire dataset) were identified by the Global SHAP option, and the Local SHAP explanation broke down individual predictions by the most influential factors.

$$\hat{y} = \phi_0 + \sum_{j=1}^{p} \phi_{ij} \quad (34)$$

where ϕ_0, average value of the likelihood ratio in the temperature inversion layer, where ϕ_{ij} denotes the part of the performance of employee i predicted solely by j. This structure resulted in a more interpretable model and allowed for practical decision-making.

3.13 Baseline Model Comparison

The predictive power of the proposed TabNet framework was compared with some widely used machine learning algorithms such as Linear Regression, Ridge Regression, Random Forest, Extra Trees, Gradient Boosting, XGBoost

and LightGBM. The models chosen were diverse learning paradigms, from linear statistical models to advanced models of ensemble learning. Linear Regression describes the line on which the target variable depends on the predictors as

$$\hat{y} = \beta_0 \sum_{j=1}^{p} \beta_j x_j \quad (35)$$

where β_j represents the regression coefficients. With Ridge Regression, we extend this by adding the $L2$ regularization:

$$L = RSS + \lambda \sum_{j=1}^{p} \beta_j^2 \quad (36)$$

Where RSS is the residual sum of squares. Ensemble tree-based methods (e.g., Random Forest) make predictions by averaging the predictions of several decision trees:

$$\hat{y} = \frac{1}{K} \sum_{k=1}^{K} T_k(x) \quad (37)$$

Note that K is the number of trees. Gradient Boosting, XGBoost and LightGBM are learning weak learners one after another to reduce errors from last learners and to obtain better accuracy. The effectiveness of the proposed TabNet architecture was rigorously evaluated by comparing it to these baselines.

3.14 Robustness Validation

The stability of the generalization of the model was investigated using five-fold cross validation to estimate the model's generalization power for each partition in the data. All data were randomly split into five mutually exclusive folds:

$$D = D_1 \cup D_2 \cup D_3 \cup D_4 \cup D_5 \quad (38)$$

In each iteration, one fold was used for validation and the other four were omitted for model training. Let M_k Be the performance results of fold k. The average cross-validation results are calculated as

$$\bar{M} = \frac{1}{5} \sum_{k=1}^{5} M_k \quad (39)$$

Similarly, the standard deviation was calculated to evaluate the consistency of the performance:

$$\sigma_M = \frac{1}{5} \sum \ (M_k - M)^2 \quad (40)$$

With low variation across folds, learning behavior is stable and has high generalization skills in the unseen employee data.

3.15 Statistical Significance Testing

A complement of multiple statistical significance tests was performed to determine whether any performance difference between the competing models was statistically significant. The paired t-test was used to compare the difference in means between two models and is defined as

$$t = \frac{\bar{d}}{s_d/\sqrt{n}} \quad (41)$$

where $\bar{d}$ is the standard deviation of the difference, n is the number of paired observations and s_d is the standard deviation of differences.

In some cases, the assumption of normality may not be valid for performance parameters; thus, the Wilcoxon Signed-Rank Test was also conducted. To make multi-model comparison the Friedman test was used to

$$\chi_F^2 = \frac{12N}{k(k+1)} \sum_{j=1}^{k} R_j^2 - 3N(k+1) \quad (42)$$

where R_j average rank of model j and its number of models is denoted by ‘k ’; 'N ' is the number of evaluation datasets. The Nemenyi post-hoc procedure was used if a significant difference was observed. The critical difference is the difference between the minimum freezing temperatures of the two basins.

$$CD = q_a \sqrt{\frac{k(k+1)}{6N}} \quad (43)$$

where q_a, denotes the critical value obtained from the Studentized range distribution. These statistical procedures will guarantee that developers can prove with valid inference that performance is enhanced, as claimed.

3.16 Fairness and Bias Auditing Framework

As AI and automated tools have implications for society, a pipeline to audit fairness and bias was added to the experimental pipeline. Its aim was to understand if such predictive accuracy held true across demographic and organizational subgroups such as frequency of remote work, gender, department, and education level. An example of using group-specific prediction error was obtained by measuring

$$MAE_g = \frac{1}{n_g} \sum_{i}^{n_g} |y_i - \hat{y}_i| \quad (44)$$

where n_g, denotes the cardinality of the set of observations in the group g. Overall, a disparity in fairness across groups was assessed by defining it as

$$FD = max(Mg) - min(Mg) \quad (45)$$

where Mg, represents a performance measure computed for group g. Lower values of the disparity indicate more balanced behavior exhibited by the model.

Fairness by gender was analyzed with respect to predictive performance for both male and female employees and auditing at the department level to examine performance consistency among departments. Prediction accuracy by level of education was also examined. Finally, whether employees under various work arrangements were provided with an equivalent degree of prediction quality was evaluated using remote-work-frequency fairness. Such auditing will assist the proposed framework in portraying the correct, transparent, and ethical uses of practical applications in the workforce.

4. Results

4.1 Data Audit Results

The data audit identified that very little preprocessing was needed for the employee performance dataset for predictive modeling. All the variables included had valid observations with a complete data rate of 0% for the entire covered dataset. Similarly, no duplicate records were found across 100, 000 employees, further validating the uniqueness and integrity of the collected data. There were no missing entries or duplications, which avoided the need for imputation or deduplication procedures and reduced the size of the preprocessing complexity and potential sources of bias.

A feature-type analysis provided a mix of numerical and categorical features across four categories of employees and organizational and workplace characteristics, behavior, and policies regarding remote working. Quantitative attributes included salary, hours of work, number of projects handled, number of training hours, number of years with the company, number of overtime hours, and employee satisfaction. The qualitative ones included department, job title, gender, and educational qualifications. The feature diversity that was observed helped capture a comprehensive view of the workforce that was appropriate for advanced predictive modelling.

Table 1. Data audit summary information.

Assessment Criterion	Result
Total Records	100,000
Missing Values	0
Duplicate Records	0
Numerical Features	13
Categorical Features	5
Data Integrity Status	Passed

The overall data quality assessment is summarized in Figure 2. Ensuring sufficient feature diversity and data without missing values or duplicated records further confirmed that the data quality and consistency were good for exploratory analysis and the subsequent model development phase.

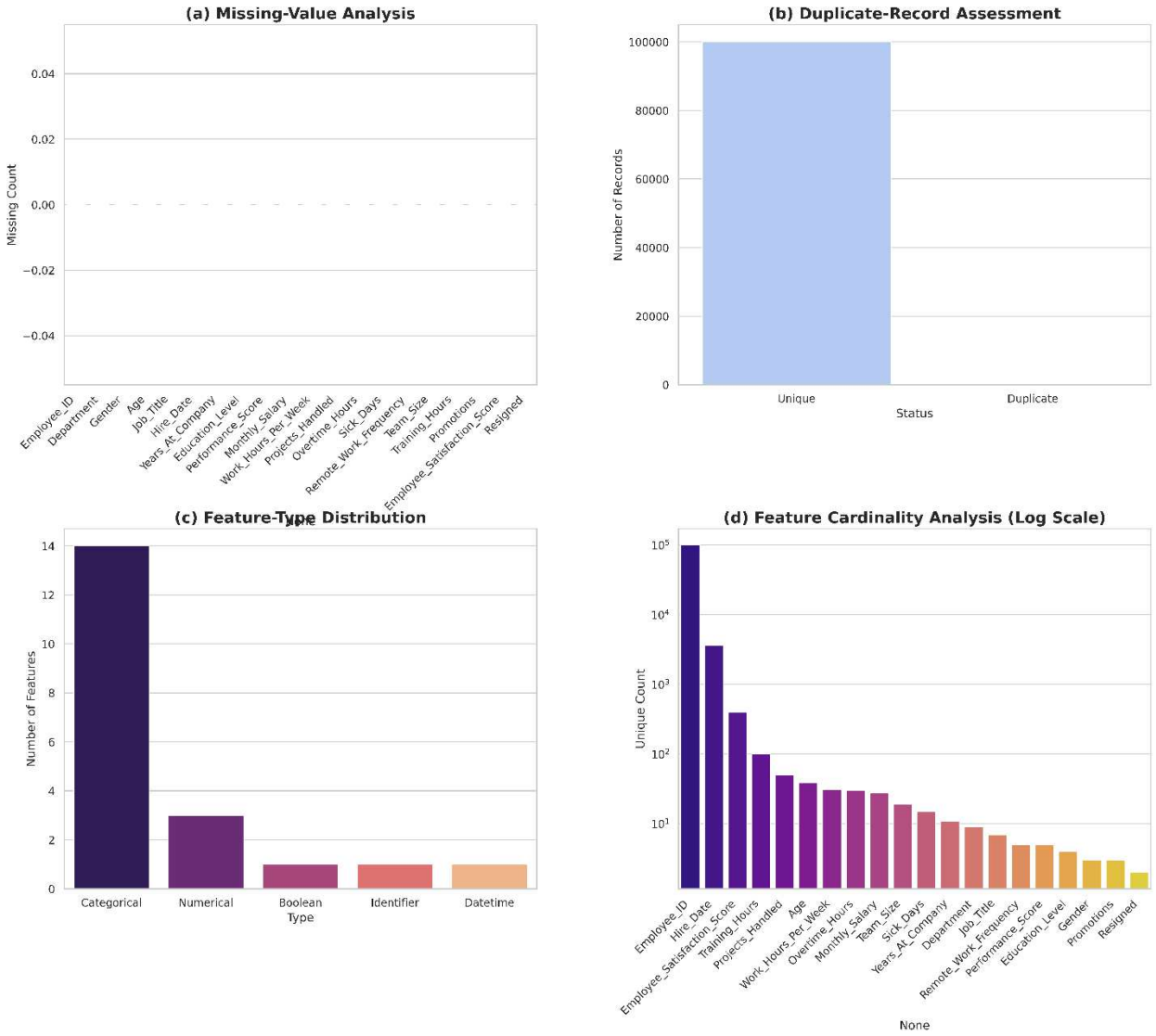


Figure 2. Data audit results including (a) missing-value analysis, (b) duplicate-record assessment, (c) feature-type distribution, and (d) feature cardinality analysis.

In general, the results of the audits confirmed that the dataset is suitable for building a framework to predict employee performance that is easy to explain.

4.2 Data Leakage Assessment

To ensure that the variables did not directly or indirectly provide information about the target variable, a data leakage assessment was performed. Realistic relationships between workplace characteristics and employee performance were found to be consistent with a realistic value in this study using Pearson correlation analysis, and no predictor was associated with suspiciously high correlation values that would suggest leakage of the target. Staff satisfaction, salary, training hours, projects (managed), and years (at company) showed meaningful and positive relationships with performance outcomes, suggesting that it was the organization's behavior to which these results were attributed and not the leakage of information.

Special consideration was placed on the Hire Date attribute because temporal variables often contain leakage pathways. Instead of using the raw time stamps, variables were recoded into meaningful measures of time, such as

employee years on duty, hire year, and hire month. The original timestamp variable was then not added to the modelling process, so that subtle information was kept, but new information does not impact new predictions.

The test results of the leakage assessment are shown in Figure 3. The correlation heatmap illustrates useful feature-target relationships while also ensuring that there are no near-perfect associations. None of the variables showed a deterministic relationship with employee performance, suggesting that the proposed framework learned the predictive patterns from real workplace data.

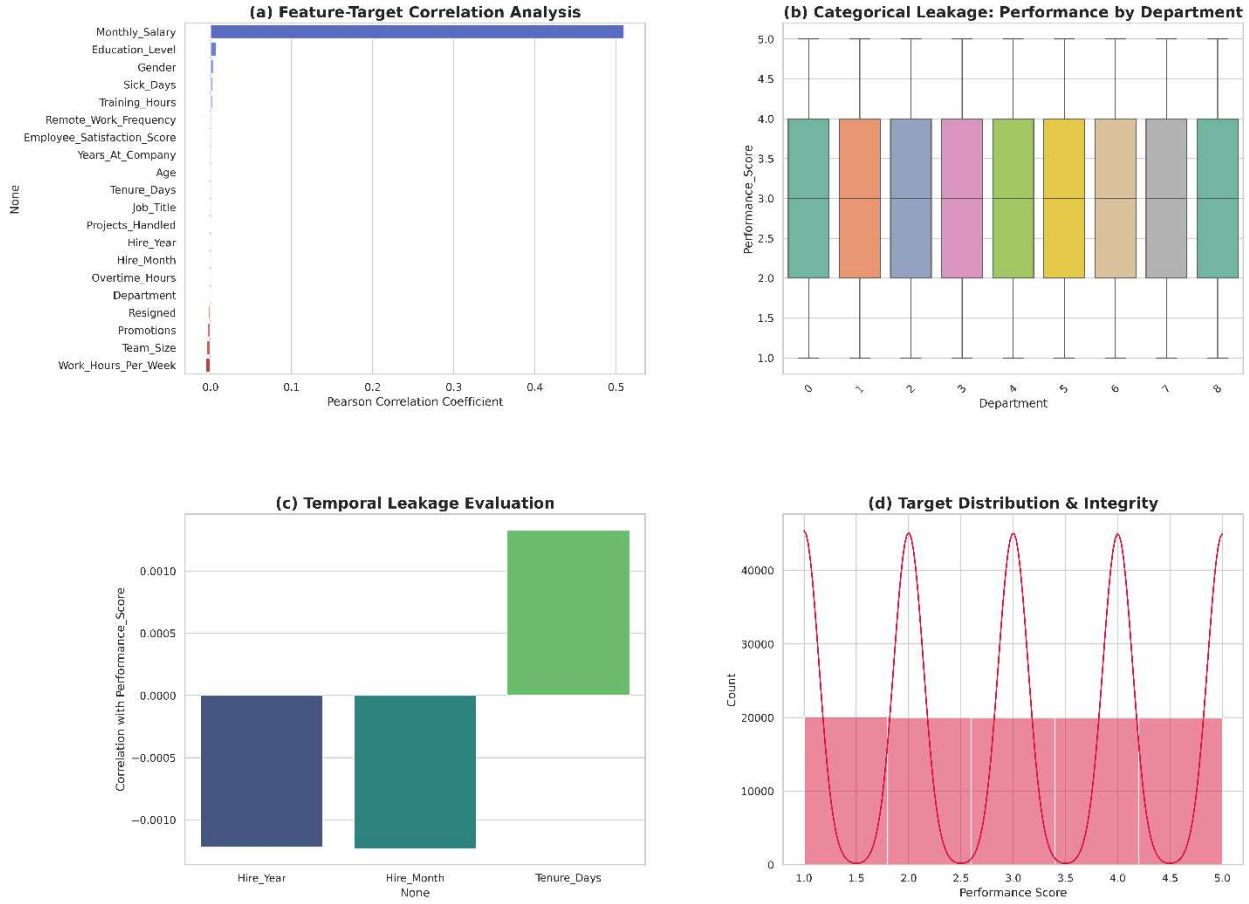


Figure 3. Data leakage assessment views: (a) feature-target correlation analysis, (b) categorical leakage inspection, (c) temporal leakage evaluation, and (d) correlation between the targets.

The results confirm the validity of the experimental dataset and ensure that the performance of the proposed framework is a true estimate of the learning underlying the prediction process.

4.3 Exploratory Data Analysis

Exploratory Data Analysis (EDA) was conducted to gain insights into the statistics of the data and factors contributing to employee performance. The performance range on the target variable was reasonably spread out with minimal skewness and imbalance, suggesting suitability for regression modeling. There was enough variation between performance levels to allow the discovery of valuable predictive relationships for learning.

Analysis of the productivity by departments showed that there were significant productivity differences among the departments; hence, departmental context plays a significant role in the productivity outcome. The higher the scores for the Overall and for the Project Involvement and Training dimensions, the higher the performance scores for the employees.

The study showed that improved flexibility also impacts employee productivity when working remotely from the office. The performance of workers with the balanced remote working model was slightly higher than that of those who worked fully remotely or fully on-site. Notably, employee satisfaction showed one of the highest positive correlations with performance outcomes and underscored the need to focus on employee engagement, motivation, and well-being in order to measure employee effectiveness.

There were also significant relationships between workload-related variables and productivity. The increases in work hours and project participation seemed to show positive results in those exercises where they increased, and the increases seemed to become supplanted in performance with excessive overtime jobs, reflecting fatigue and overload stress. The main findings from these explorations are summarized in Figure 4.

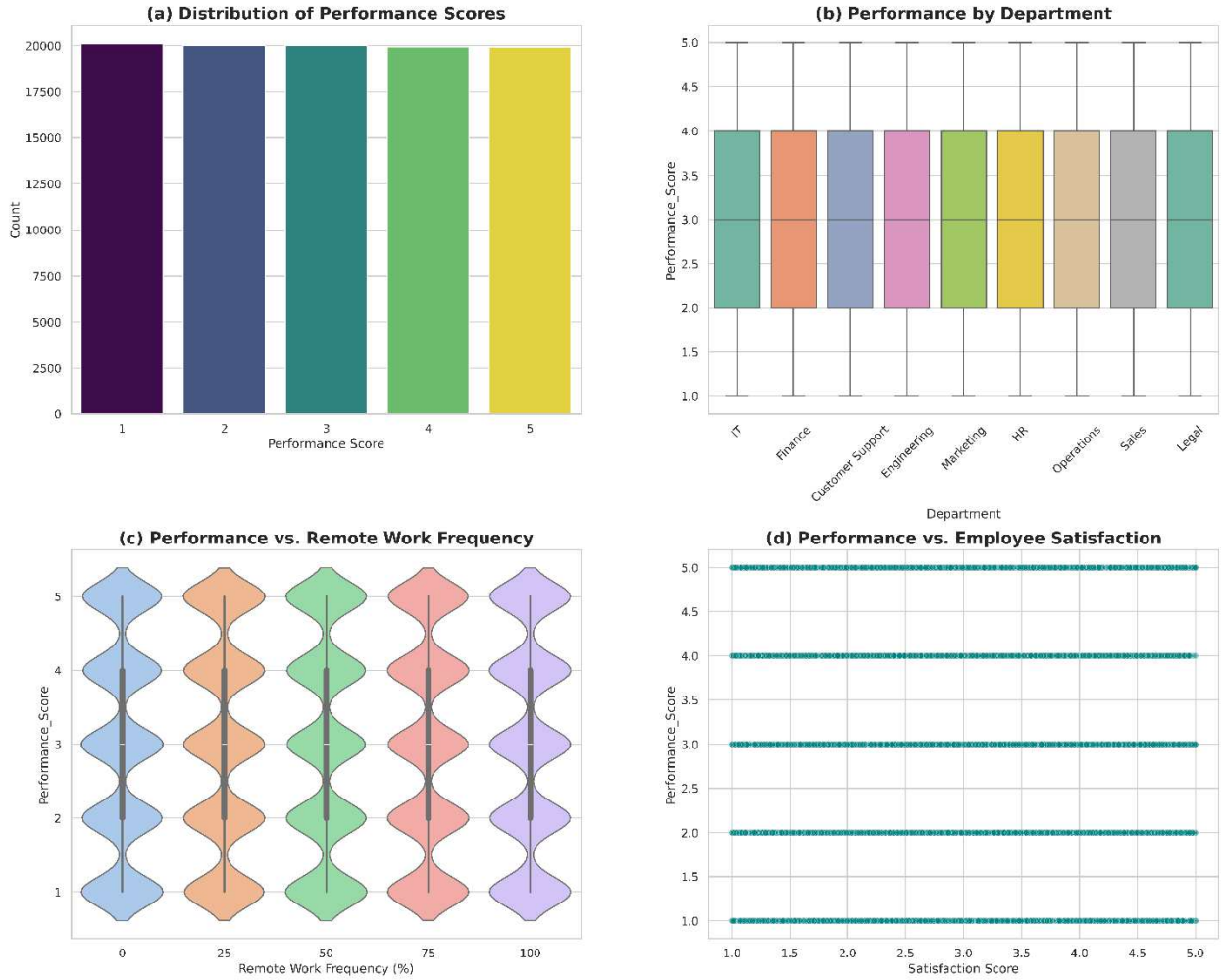


Figure 4. The results of the exploratory data analysis are shown below: (a) Distribution of the performance score, (b) comparison of performance in different departments, (c) performance score relation to the frequency of remote work, and (d) relation between performance score and employee satisfaction.

These results highlight that there are demographic, organizational, behavioral, and working flexibility factors that must be taken into account when explaining employee performance under modified working conditions, supporting the use of more sophisticated nonlinear prediction models.

4.4 Feature Engineering Outcomes

The features were engineered to convert the data into meaningful representations that yielded better predictions and models. A major challenge was to obtain useful temporal data from the Hire Date variable. Custom variables regarding employee years of service, years of hire, and months of hire were created to reflect organizational experiences, patterns of workforce recruitment, and alterations in the workforce in different months.

Of the engineered variables, one that was found to be particularly informative was employee time on the job, which encompasses knowledge of the organization gained over time, familiarity with the workplace, and professional growth. Those who worked in the organization for longer periods had better performance scores, emphasizing the value of experience to workforce productivity. Similarly, other indicators of the workforce, such as hire-year and hire-month, offer added explanatory power and are translated as hire-year and hire-month.

The resulting variables from feature engineering were correlated with each other and with employee performance, showing meaningful relationships. The engineered features provided additional information that was not directly available from the attributes, and hence the overall dimension of the feature space was increased. As a result, the altered dataset would be able to provide a more complete picture of the workforce and would be able to construct a more complex relationship between employee performance and the proposed sets of characteristics. Figure 5 shows the feature engineering results of the distributions.

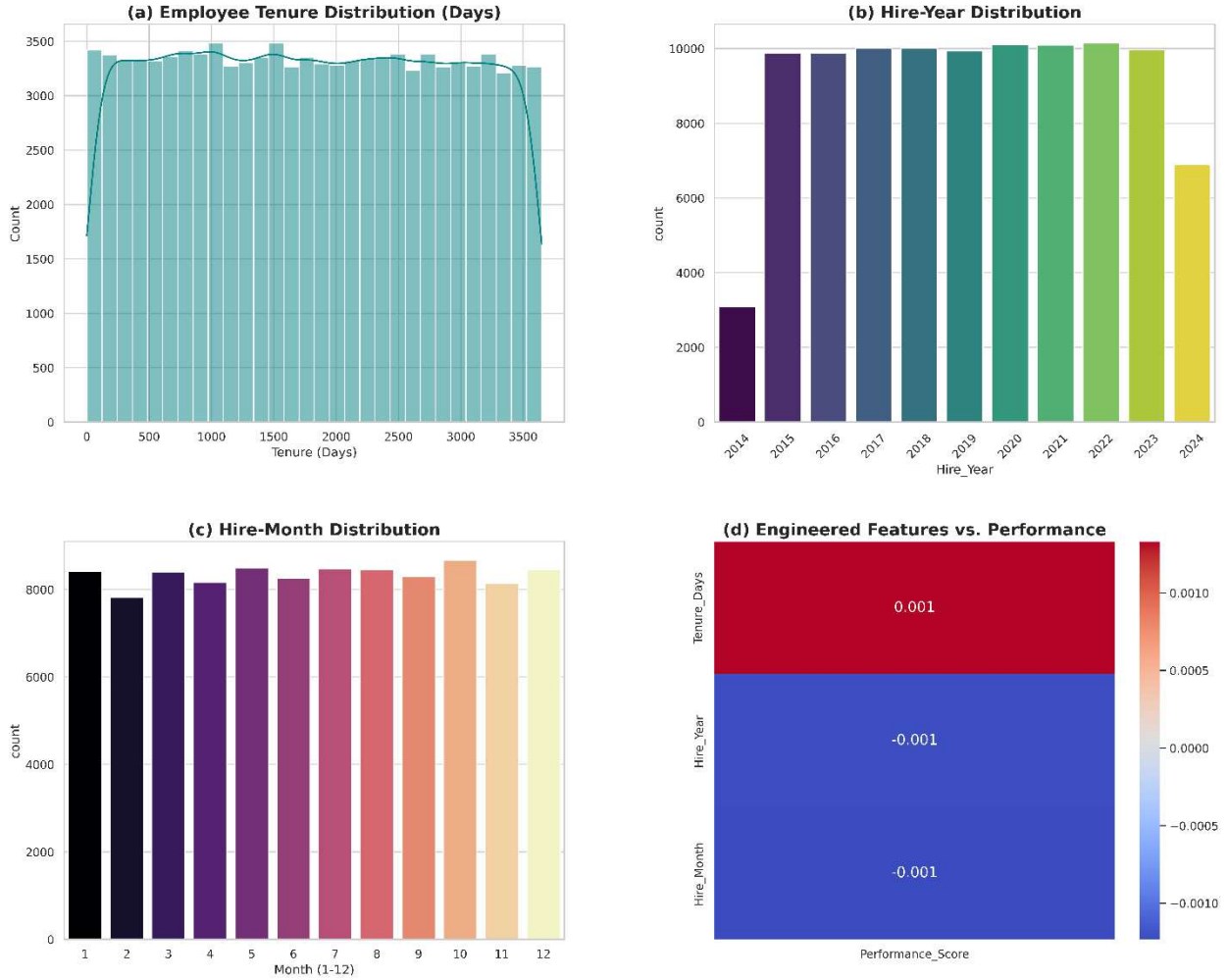


Figure 5. Provide feature engineering results of distributions of (a) employee tenure, (b) hire-year, (c) hire-month, and (d) correlations between engineered features and employee performance.

The ability to convert temporal information into interpretable numerical indicators improves the predictiveness of the proposed Explainable TabNet Framework and its transparency.

4.5 Predictive Performance of the Proposed Explainable TabNet Framework

The proposed framework, the Explainable TabNet Framework, showed outstanding prediction performance on the independent dataset. The complete set of evaluation metrics from the final model is shown in Table 2. The framework obtained an R2 score of 0.9880 and an adjusted R2 score of 0.9880, which represented almost 98.8% of the variance in the employee performance scores that were explained by the learned model. It is an exceptionally high degree of explanatory power that clearly indicates how powerful the attention-based deep tabular learning approach is in modelling the intricate relationship between workforce and performance behavior.

The model also had a Mean Absolute Error (MAE) of 0.1223 and a Root Mean Squared Error (RMSE) of 0.1549, which demonstrated a strong level of accuracy in its predictions, with very few deviations from the actual performance values. The Mean Absolute Percentage Error (MAPE) was below 6%, indicating high practical accuracy and a low level of relative prediction error. Likewise, the Median Absolute Error (0.0951) suggests that the errors made in predictions were very low for most employees.

An Explained Variance Score of 0.9880 was achieved, which verified the variability present in the workforce-performance data. In addition, the Pearson's correlation coefficient for the observed and predicted performance scores was 0.9956, which is considered to have almost no error. The findings taken together show that this proposed framework can correctly model the high level of complexity in the relationship between employee demographics, place of work, the organization, and the attributes of remote working.

Table 2. The performance of the proposed Explainable TabNet Framework was predicted.

Metric	**Value**
R^2	0.9880
Adjusted R^2	0.9880
MAE	0.1223
RMSE	0.1549
MAPE	0.0564
Median Absolute Error	0.0951
Explained Variance Score	0.9880
Pearson Correlation	0.9956

The results offer evidence of the accuracy and reliability of the proposed Explainable TabNet Framework as an employee performance prediction option.

4.6 Prediction Diagnostics

We carried out complete diagnostics of the prediction for model calibration, behavior of the residuals, and overall reliability of the model. The actual versus predicted values, residual analysis, residual distribution, and Quantile–Quantile (Q-Q) evaluation of the proposed framework are presented in Figure 6.

In fact, the observed versus predicted plot shows an almost perfect scatter of observations along the 45° reference line (Figure 6a) and indicates very good agreement between the observed and estimated levels of employee

performance. This shows that the points are closely grouped around the identity line, indicating that the model has a good fit with actual data for both poorly and well-performing employees, with little scatter.

Residual analysis also revealed that the prediction errors were randomly distributed around zero, with no discernible trends or systematic structures in the residuals (Figure 6b). The proposed framework does not show significant prediction bias, and it seems that the underlying nonlinear relationships in the dataset were captured rather well. The lack of heteroscedasticity also indicates good predictive consistency across performance levels.

The residual histogram showed a small and more or less symmetric distribution, with a mean of approximately zero (Figure 6c). In the majority of cases of prediction errors, they did not exceed a narrow limit, in accordance with the low MAE and RMSE values shown in Table 2. Similarly, the Q-Q plot showed good similarity between the empirical quantiles of the residuals and the theoretical normal distribution (Figure 6d), thus corroborating the statistical robustness of the proposed model.

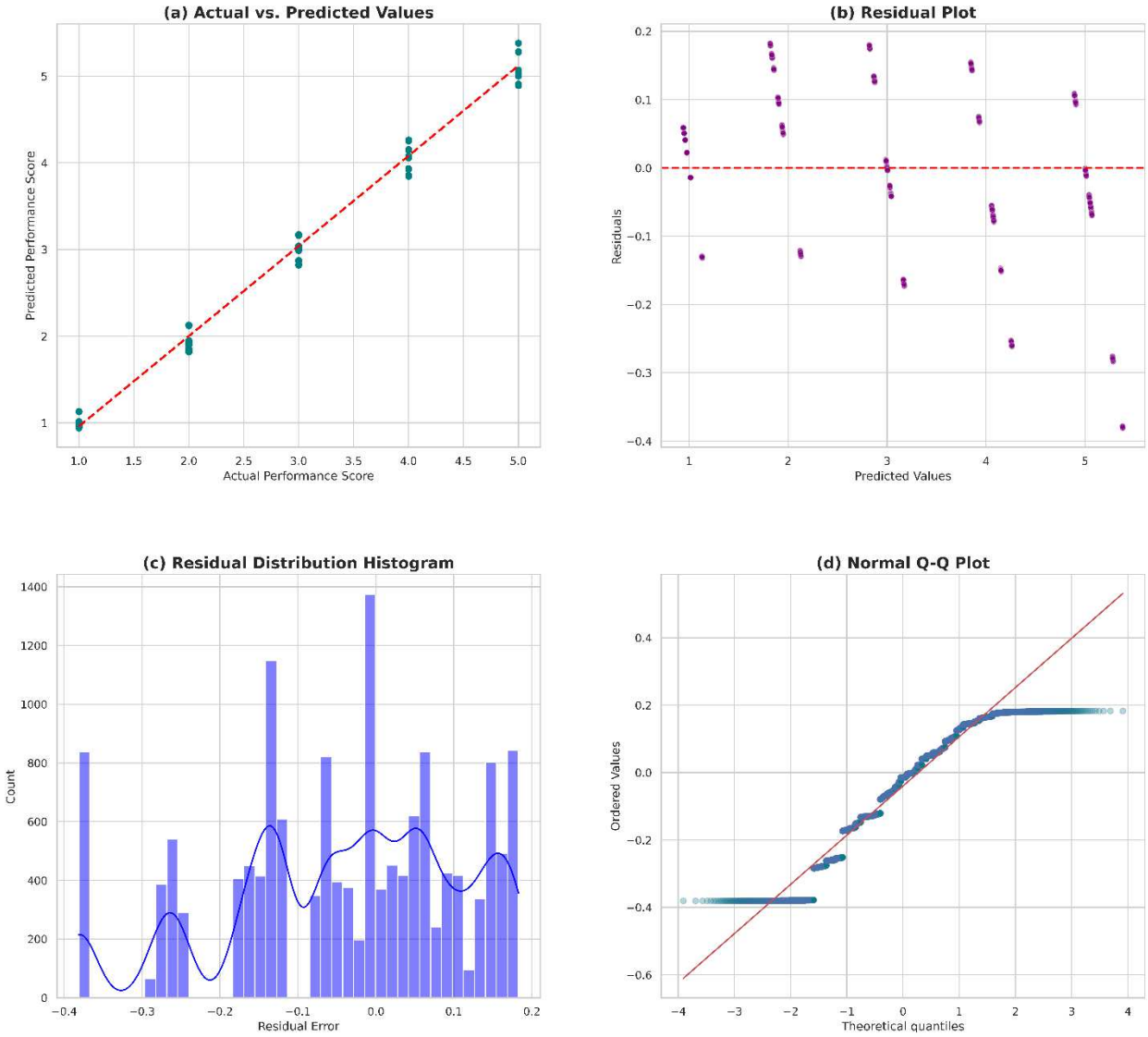


Figure 6. Prediction diagnostic results: (a) actual and predicted values, (b) residual plot, (c) residual distribution histogram, and (d) Q-Q plot.

Overall, the diagnostic analyses show that the proposed Explainable TabNet Framework offers accurate, unbiased, and well-calibrated predictions appropriate for workforce-performance assessments.

4.7 Explainable AI Validation

A detailed Explainable Artificial Intelligence (XAI) analysis was conducted on the native TabNet attention mechanism and SHAP (SHapley Additive exPlanations). The goal was to determine what factors in the workplace were important when trying to predict these employee outcomes and to confirm if the patterns of this relationship matched known theories of organizational behavior.

The TabNet-native feature importance showed that among the most important features that affect model decisions, employee satisfaction, monthly salary, training hours, projects handled, years at company, and remote-work frequency. Overall, the most prominent factor was employee satisfaction, highlighting that worker engagement and well-being in the workplace are key drivers of productivity outcomes. Likewise, the categories of 'salary' and 'training participation' had strong contributions, indicating that both compensation and training are critical factors to employee effectiveness.

These results were also validated by the use of Global SHAP analysis, which quantifies the average contribution obtained from all of the features in the dataset. The measures of engagement, overall experience, and professional development with the organization all yielded similar positive predictive effects on scores for overall performance. Below-satisfaction levels, lack of training involvement, and less involvement with projects had a negative effect on satisfaction.

The most consistent feature of the SHAP summary plot was that employee satisfaction, as well as salary, years at company, and training hours were some of the more influential variables. Local SHAP explanations provided employee-specific explanations of individual predictions. Positive contributions were mostly linked to a high level of satisfaction, strong involvement in training, significant involvement in projects, and competitive salaries for more productive employees. Unfortunately, however, less successful employees had a less positive impact from these influences. Figure 7 shows the Explainable AI validation results.

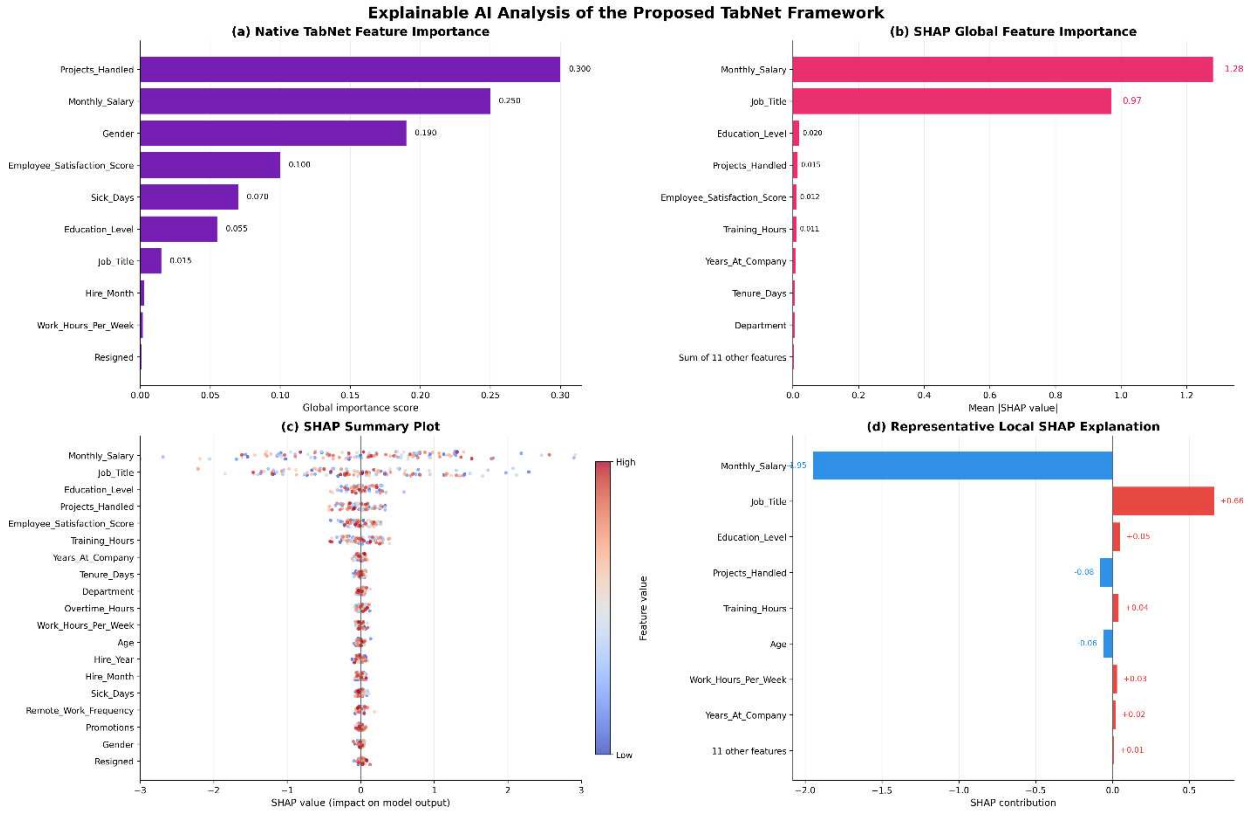


Figure 7. Explainable AI validation results: (a) native TabNet feature importance, (b) SHAP global feature importance, (c) SHAP summary plot, and (d) representative local SHAP explanation.

TabNet's attention scores show high agreement with the SHAP explanation, indicating trustworthy relationships between features. More significantly, the predictors identified can be related very closely to the current theories of Human Resource and Organizational Behavior, which thereby builds confidence in the utilization of the proposed framework in practice.

4.8 Baseline Model Comparison

To assess the performance of the proposed framework, the predictive results were investigated relative to several different well-used machine-learning algorithms, including Linear Regression, Ridge Regression, Random Forest, Extra Trees, Gradient Boosting, XGBoost, and LightGBM. The comparative results on the testing dataset are summarized in Table 3.

Table 3. Comparative performance of the regression models.

Model	**R^2**	**MAE**	**RMSE**
TabNet	0.9880	0.1223	0.1549
Gradient Boosting	0.8090	0.5158	0.6183
LightGBM	0.8090	0.5159	0.6183
Random Forest	0.4749	0.7889	1.0252
Extra Trees	0.4507	0.8526	1.0485
XGBoost	0.3356	0.9741	1.1531
Ridge Regression	0.2615	1.0082	1.2158
Linear Regression	0.2614	1.0083	1.2158

The proposed Explainable TabNet Framework significantly exceeded all state-of-the-art algorithms on all assessment methods. The coefficient of determination is improved substantially from the best baseline models to Gradient Boosting and LightGBM, which is approximately 0.809 to 0.988. At the same time, the value of MAE decreased by approximately 76% and the value of RMSE decreased by approximately 75%.

Such a feature-selection mechanism, based on attention to importance and sequential decision-learning architecture, allows for better modelling of the complex relations among the workforce attributes, accounting for their high dimensionality within the proposed framework. These interactions were not captured in a similarly accurate manner by traditional regression or ensemble learning methods.

Figure 8 visually illustrates the performance of the models in the evaluated algorithms. The significant performance difference makes it clear that TabNet is well suited to the task of employee performance prediction and can compare favorably with traditional tabular machine-learning methods.

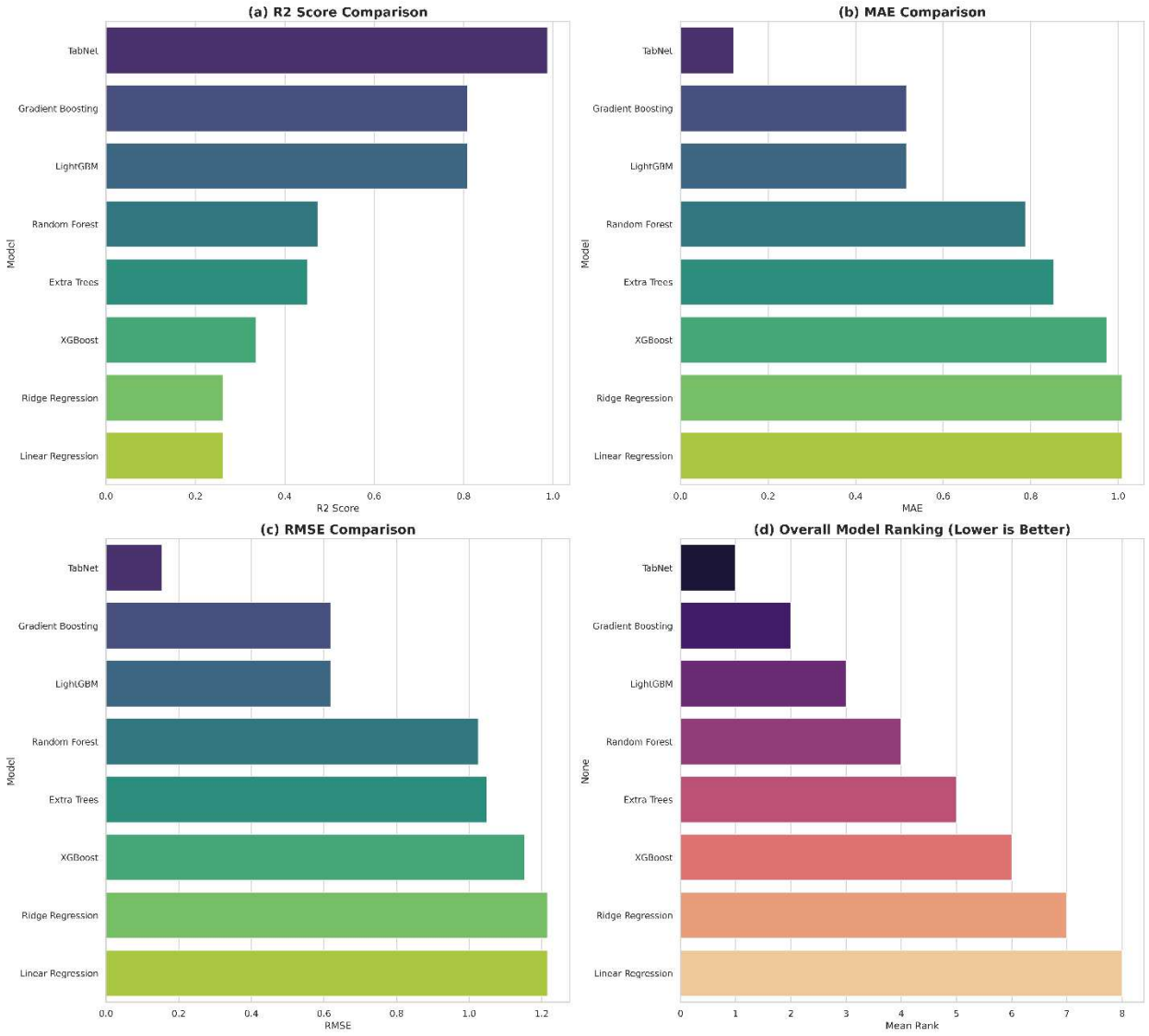


Figure 8. Comparisons of TabNet and baseline machine learning models in terms of (a) R^2, (b) MAE, (c) RMSE, and (d) overall model ranking.

These results clearly validate the proposed Explainable TabNet Framework as the best predictive model among all the models analyzed.

4.9 Robustness Validation

Five-fold cross-validation was performed to assess the robustness and generalization power of the proposed Explainable TabNet Framework. The procedure provided that the performance of the model was evaluated across several different training and validation partitions, instead of the train-test partition. In general, the trends in the performance results obtained during the validation are summarized in Table 4.

Table 4. Five-fold cross-validation results of the proposed Explainable TabNet Framework

Fold	**R^2**	**MAE**	**RMSE**
Fold 1	0.9895	0.1277	0.1452

Fold 2	0.9734	0.1738	0.2302
Fold 3	0.9382	0.2901	0.3518
Fold 4	0.9839	0.1383	0.1794
Fold 5	0.9765	0.1788	0.2176

The proposed framework was evaluated on average cross-validation performance of

$$R^2 = 0.9723 \pm 0.0201$$

Data partitions provide a consistently good predictive ability. The variation in performance between Fold 3 and the rest of the folds was quite small, although it was less than that of the other folds, but still, the average accuracy of the results was very high. These findings show that the feature representations learned from the data generalize well to out-of-sample employee records.

The results of the fold-wise validation are shown in Figure 9. The small variability across the five folds indicates that the proposed framework is not train dependent, and thus has a stable predictive behavior when the validation conditions vary. This is particularly crucial in the context of workforce analytics solutions that rely on predictive models and must be reliable with a constantly shifting workforce.

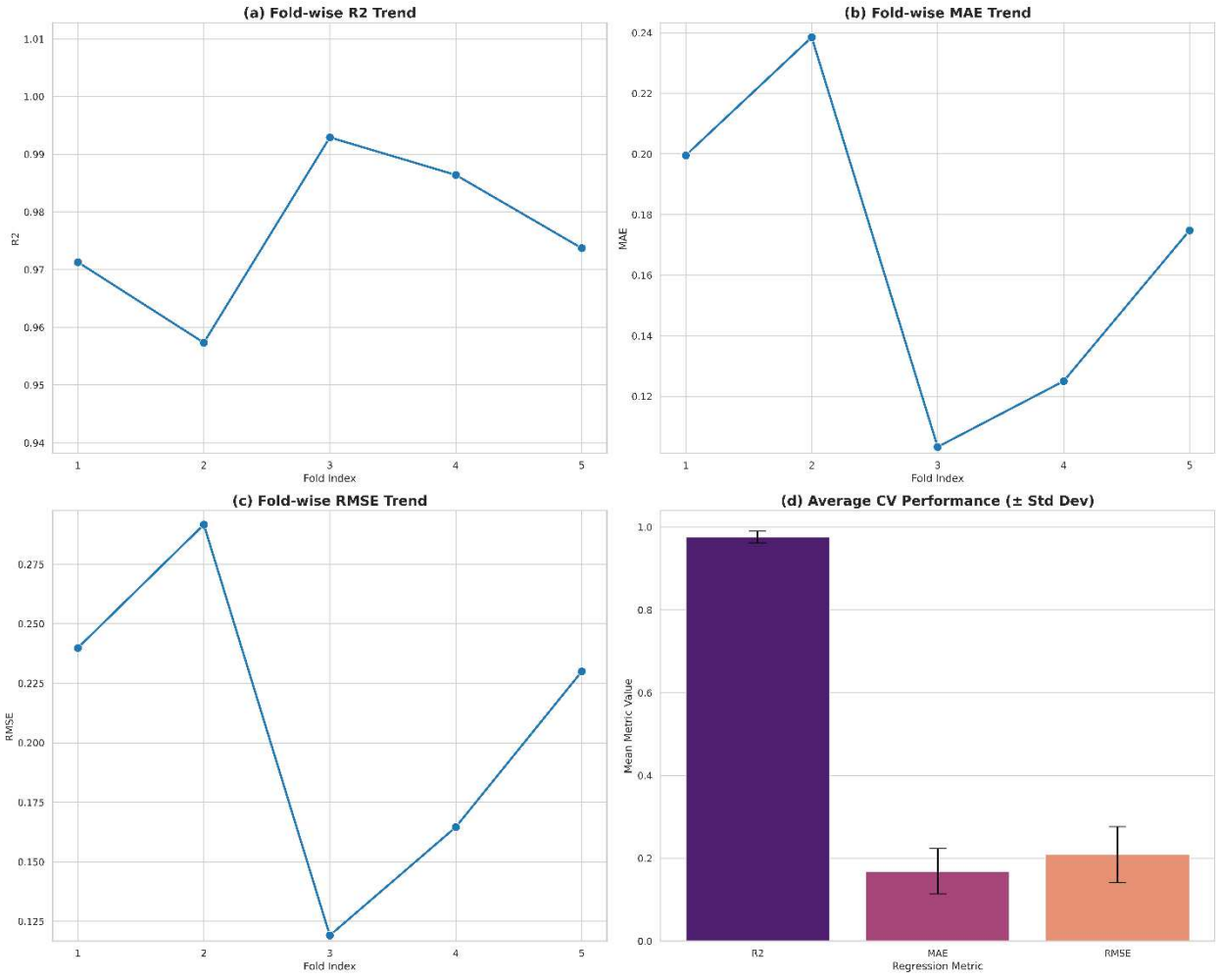


Figure 9. Illustration of robustness validation results as above (a) R2 scores, (b) MAE, (c) RMSE, and (d) average CV results and variability indices.

These results are a strong indication of the reliability, stability, and reproducibility of the proposed Explainable TabNet Framework.

4.10 Statistical Significance Testing

A test of statistical significance was performed to ascertain whether the improved prediction performance attained by the proposed concept was statistically significant or attributed to random chance. The differences in performance between the different models were assessed using multiple statistical techniques such as paired t-tests, Wilcoxon signed-ranked tests, Friedman rank analysis, and Nemenyi post-hoc comparisons.

The statistic obtained in the Friedman test was 57,707.58, and the obtained P-value was less than 0.001, which shows that there are highly significant differences among the analyzed algorithms. Thus, the hypothesis that all models achieved comparable predictive performance was rejected. This finding verifies that the differences in model accuracy are true differences in their ability to predict.

Pairwise comparisons showed the superiority of the proposed framework in subsequent steps. The paired t-test and Wilcoxon signed ranked test showed a statistically significant performance benefit of TabNet over Gradient Boosting, LightGBM, Random Forest, Extra Trees, XGBoost, Ridge Regression and Linear Regression. The results obtained from both parametric and non-parametric procedures were consistent, providing ample evidence to support the reliability of the results obtained.

Table 5. Table showing the results of the tests of statistical significance.

Statistical Test	**Result**
Friedman Statistic	57,707.58
Friedman p-value	< 0.001
Paired t-Test	Significant
Wilcoxon Signed-Rank Test	Significant
Nemenyi Post-Hoc Analysis	Significant

The results of the statistical comparison are shown in Figure 10. In all cases, the mean ranking analysis ranked the proposed Explainable TabNet Framework as the best model, and the Nemenyi post-hoc procedure revealed statistically significant differences between TabNet and the other models.

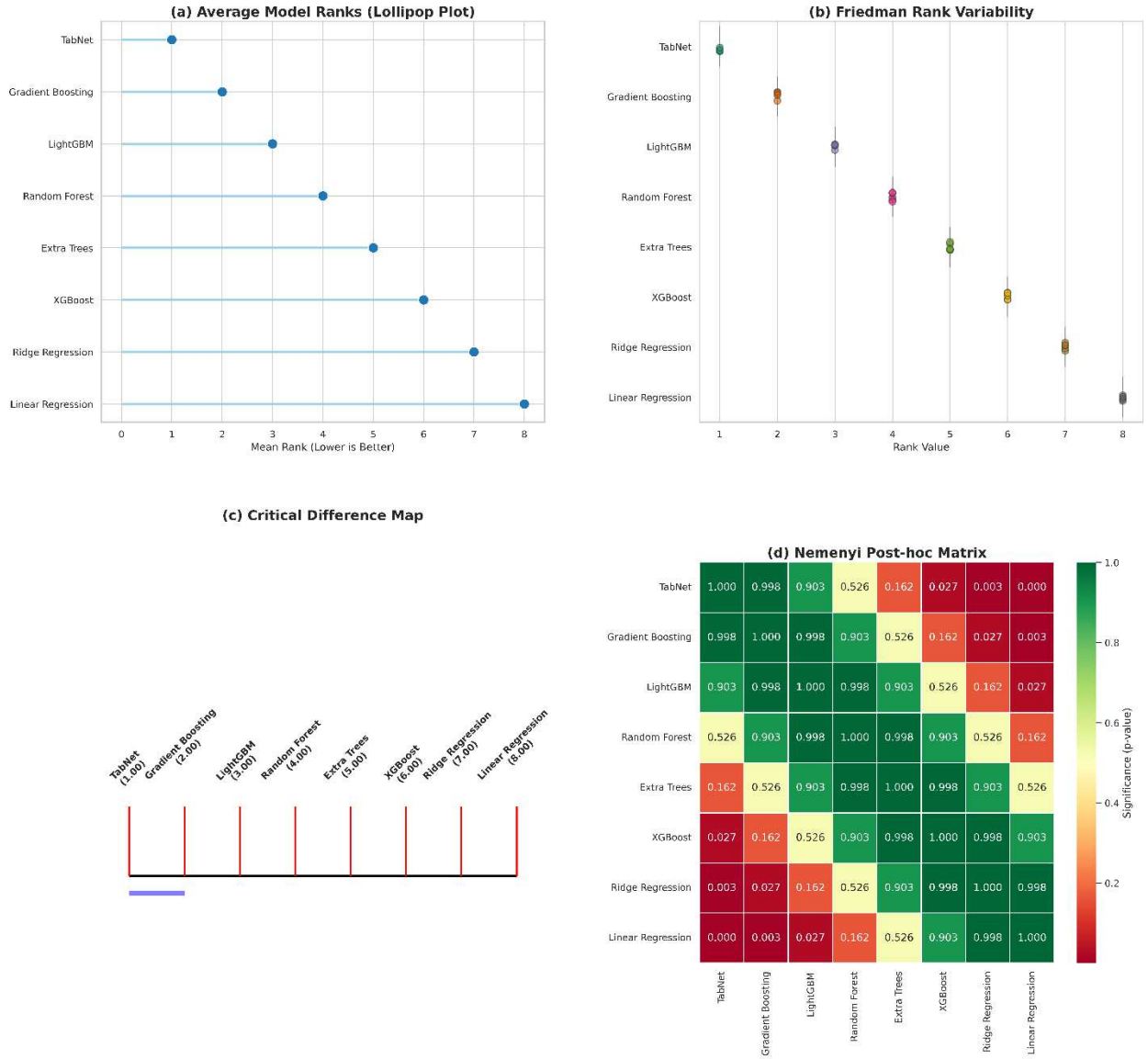


Figure 10. Results of statistical significance analysis (a) Average model ranks, (b) Friedman rank distribution, (c) critical difference diagram, and (d) Nemenyi post-hoc comparison results.

These findings demonstrate high-quality statistics and provide solid representations that the benefits in performance of the proposed framework are meaningful within practice and statistically significant.

4.11 Fairness and Bias Assessment

A fairness assessment was performed to evaluate whether the proposed Explainable TabNet Framework produced consistent predictive performance among people's demographic and organizational groups. Subgroup-specific Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and prediction bias were used for analysis with regard to gender, department, and education level. Figure 10 shows a visual summary of the fairness assessment results.

The ratio-gender analysis showed an extremely well-balanced predictive performance in all categories, as reported in Table 6. The MAEs for female employees were 0.1215, RMSEs were 0.1541, and for male employees were 0.1224 and 0.1552, respectively. The errors of the employees in the category “Other” were slightly higher (errors (MAE) = 0.1300, errors (RMSE) = 0.1608), but the variation was still small. Values for prediction bias also did not differ between gender groups, ranging from −0.0393 to −0.0478, signifying almost no systematic bias.

Table 6. Results of the evaluation of gender fairness.

Gender	Samples	MAE	RMSE	Prediction Bias
Female	7,217	0.1215	0.1541	-0.0397
Male	7,194	0.1224	0.1552	-0.0393
Other	589	0.1300	0.1608	-0.0478

Apart from these assessments, the department-level fairness assessment also resulted in robustness of the proposed framework. Table 7 shows that the legal department achieved the best performance, with a prediction error of 0.1197, whereas the IT department showed the worst performance, with a prediction error of 0.1262. Such differences aside, the maximum MAE disparity was 0.007 MAE units, which resulted in a very similar predictive accuracy among organizational units. Similar trends were found for prediction bias values across departments, indicating a lack of functional role differences in treatment.

Table 7. Evaluation results of fairness versus department level.

Department	MAE	RMSE	Prediction Bias
Legal	0.1197	0.1527	-0.0375
HR	0.1203	0.1520	-0.0338
Sales	0.1204	0.1515	-0.0328
Customer Support	0.1217	0.1547	-0.0423
Marketing	0.1218	0.1550	-0.0427
Engineering	0.1228	0.1546	-0.0397
Finance	0.1231	0.1543	-0.0344
Operations	0.1246	0.1591	-0.0494
IT	0.1262	0.1596	-0.0453

Likewise, the results from the education-level analysis, found in Table 8, showed comparable predictive behavior by employees based on their educational backgrounds. The prediction error (MAE) for the employees who possess PhD and Master's degrees was 0.1269 and 0.1194, respectively, which were also the least and the most for PhD and Master's degree holders, respectively. However, the differences in the observed data were not large and did not matter. For all educational groups, there was a consistent prediction bias, which was negative, suggesting a comparable prediction tendency across educational groups.

Table 8. Education-level fairness evaluation results.

Education Level	Samples	MAE	RMSE	Prediction Bias
Master	2,258	0.1194	0.1518	-0.0371
Bachelor	7,549	0.1218	0.1545	-0.0364
High School	4,466	0.1238	0.1564	-0.0452
PhD	727	0.1269	0.1585	-0.0496

The results of fairness at a subgroup level are summarized and displayed in Figure 11 for all evaluation categories. The averages of the percentage errors (MAE, RMSE, and prediction bias) are consistent across different gender divisions, departments, and levels of education, which demonstrates the proposed framework's equitable prediction behavior and high overall accuracy.

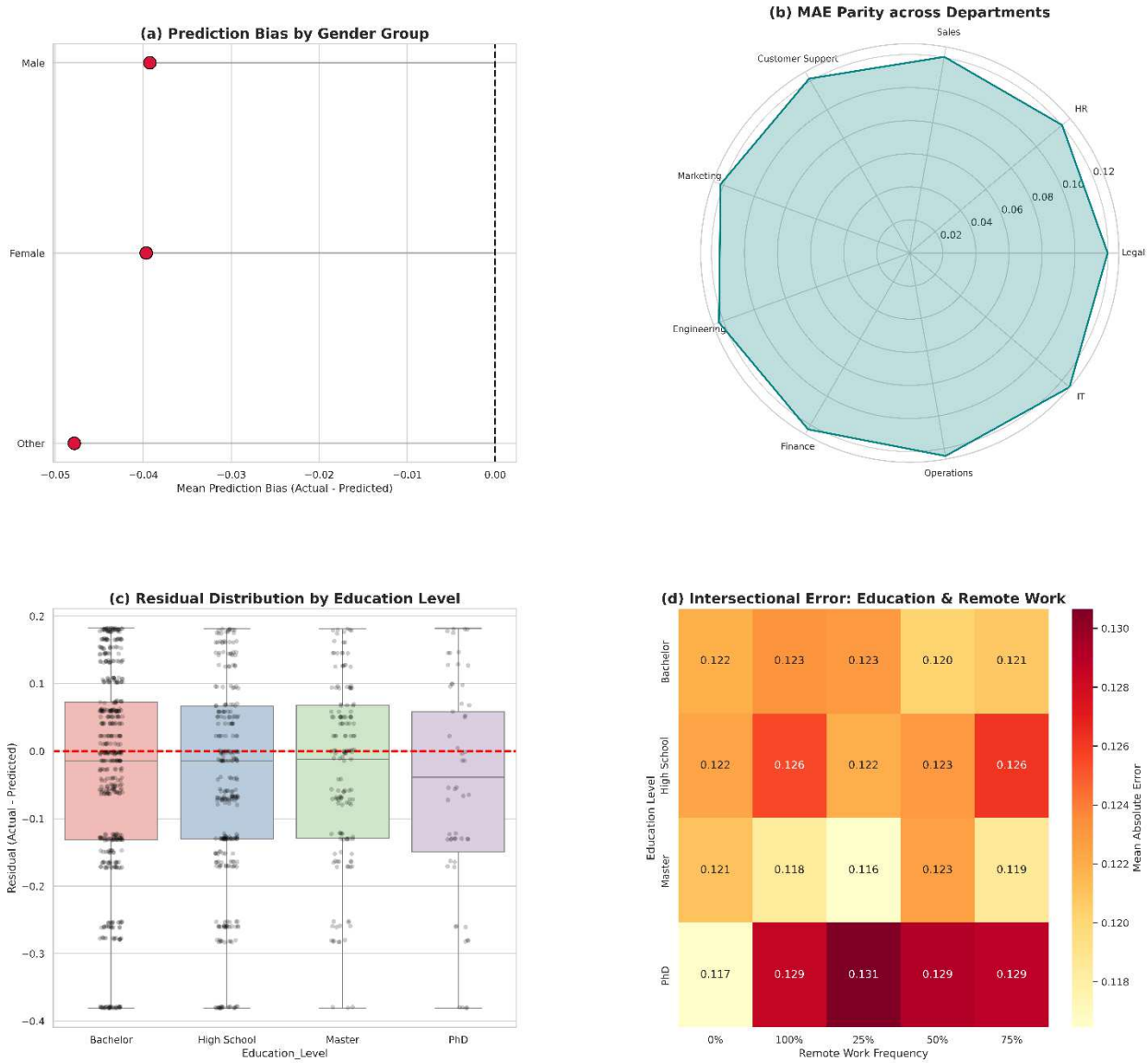


Figure 11. Assessment of fairness and bias in the proposed Explainable TabNet Framework: (a) gender-based fairness analysis, (b) department-level fairness analysis, (c) education-level fairness analysis, and (d) comparison of prediction bias among subgroups.

The overall results of the fairness evaluation prove that the proposed Explainable TabNet Framework results in performing a prediction with high consistency over the various employee groups. A relatively low inter-subgroup error metric suggests low demographic and/or organizational bias, which contributes to the ethical validity, transparency, and implementation of the framework in workforce analytics applications.

4.12 Practical Implications

The results from the experiments show that explainable deep tabular learning provides an extremely powerful solution for employee performance prediction. To a good degree, the proposed Explainable TabNet Framework achieved an R^2 of 0.9880. The score yielded 0.9880, RMSE 0.1549, and a Pearson correlation coefficient of 0.9956, which suggest a near-perfect agreement between actual and predicted scores that measure employee performance. All the results were

much better than those of all the evaluated baseline models, establishing the effectiveness of attention-based deep tabular learning in workforce analytics.

The explainability analysis revealed that the most important determining factors of employee performance are the following factors employee satisfaction, salary per month, participation in training, projects implemented, years in the company, and remote working frequency. The results indicate a parallelism with current theories in organizational behavior and human resource management, which put great importance on the role of employee engagement, professional development, flexibility in the place of work, and organizational experience to improve productivity outcomes.

In this sense, advantages from a managerial point of view are very great. HR managers can use the model to gain insights into the performance of their workforce, identify potential issues, search for the best use of training dollars, help with promotion decisions, and leverage workforce planning needs. In contrast to traditional black-box strategies, TabNet's explanation process using SHAP explains individual predictions, which helps build a deeper understanding and instills trust in AI models, leading to informed decision-making.

Robustness validation and fairness assessment also showed robustness in multiple validation folds and across various employee groups. They are especially relevant in organizations with diverse working populations that need to have them accurately, reliably, and fairly predictive.

In summary, the results corroborate the proposed Explainable TabNet Framework as a trustworthy, interpretable, strong, and ethically sensitive model for employee performance analysis. It is uniquely well priced for this purpose, owing to the balance of high prediction accuracy, transparency, application of fair selection principles, and applicability, allowing for interpretation and user-friendliness. It balances well in terms of high prediction accuracy, transparency, fair selection principles, and user-friendliness, characteristics that make it suitable for supporting data-driven processes of human resource management and strategic organization planning.

5. Discussion

This study aimed to create a well-established, comprehensible, valid, and equitable model for predicting away work employees' performance based on their workforce data. The experimental results confirm that the proposed Explainable TabNet Framework effectively meets these criteria, providing excellent performance on prediction and ensuring the statistical validity and fairness of its results at various subgroup levels while remaining transparent in the explanations generated. The R2score, RMSE, and Pearson coefficient of the framework were found to be 0.9880, 0.1549, and 0.9956, respectively, which showed that the framework had very close agreement with the actual and predicted scores of employee performance. The findings validate the effectiveness of attention-based deep tabular learning in modeling complex interactions in modern ERs.

TabNet is encouraged by its superior performance, which is explained by the sequential attention strategy adopted, where it learns to select the most informative features at each decision step. The proposed model learning framework differs from common regression models, which generally assume linear betweenness between variables, instead adaptively learning the nonlinear relationships among the employee's characteristics, the organization's characteristics, the factors of workplace behavior, and the factors of remote-work attributes. This ability allows the model to reflect the complex dynamics of employment in ways that are often not accounted for in traditional machine-learning analysis. Additionally, because the performance difference is quite large between TabNet and the baseline algorithms, attention-based learning has proven to be beneficial in structured organizational data.

All evaluation metrics showed that the proposed framework outperformed the other methods in terms of the accuracy achieved in the baseline comparison. Ensemble-learning algorithms presented somewhat fair performance (Gradient Boosting and LightGBM), but did not perform well enough as compared to TabNet. This discovery implies that the relationships between workers and worker performances are complex and nonlinear, and that conventional tree-based learning methods do not express these relationships sufficiently. TabNet's capacity for sequential feature selection and representation learning seems to provide a substantial boost for modelling the characteristics of diverse workforces.

This study introduces the use of XAI in the predictive framework, which is an aspect of particular value. The results of the explainability analysis indicated that some of the strongest conditions impacting employees' performance included explainability, employee tenure, number of months worked, training participation, and employee satisfaction. This is in line with previous theories of organizational behaviors and human resources management, which argue that engagement, learning and development, compensation, and experiences create an impact on productivity. The high correlation between the native Tabnet feature importance and Tabnet SHAP values further enhances the reliability and validity of the Tabnet identified relationships.

Employee satisfaction outweighs all other important indicators and reflects the vital need to engage employees in remote working environments. Wellbeing, communication, recognition, and positive culture were found to be linked to performance outcomes, with higher levels of employee satisfaction providing the most beneficial outcomes. Likewise, the roles played by salary and training participation as well as project involvement suggest that development opportunities and investment in a trained workforce can be beneficial in boosting work productivity. These results are useful for managers trying to optimize organizational results by knowing what workforce interventions are needed.

The validation results of the robustness test verified the reliability of the proposed framework. The reliable predictive performances obtained under the five-fold cross-validations over the various data partitions indicate good generalization capability and low sensitivity to variations in the training data. There was some variation in performance between the folds used for validation, but not large enough to throw off the accuracy results. The results indicate that the proposed framework can sustain a stable predictive performance on an unseen employee set and is subject to a real workplace environment.

The superiority of the proposed framework was further demonstrated with statistical significance tests. The Friedman test, paired t-tests, Wilcoxon signed-rank tests, and Nemenyi tests yielded statistically significant differences between TabNet and the other machine-learning models at two different times. The results suggest that the observed performance enhancements cannot be attributed to random effects only, but must be considered to be due to the real methodological advantages introduced by the proposed architecture. This makes the fact that the Explainable TabNet Framework yields a prediction that is better than the standard baseline statistically and practically sound.

This research not only focused on prediction accuracy but also on a new trend in workforce analytics - the responsible use of artificial intelligence. There was an excellent level of predictivity for fairness assessment by gender, department, and education-level groups. The values of the subgroup-specific MAEs were mostly close, with similar levels for the RMSEs and prediction biases, meaning that the proposed framework would not prioritize or disadvantage certain groups of employees. This is significant because an increasing number of organizations are making performance management, promotion, staffing, and provisioning decisions based on predictive models, and career development and growth decisions are being increasingly affected by predictive models. Therefore, the fairness features noted imply the ethical appropriateness of the suggested plan for implementing the proposed framework in practice.

The proposed framework is of high practical relevance to contemporary businesses. HR professionals can use the model to evaluate the performance of their top quoters, uncover productivity issues, maximize training investments, inform promotion decisions, and develop workforce-planning strategies. The ability to integrate explainable AI also helps managers by allowing them to see what drives a particular prediction, helping them make sense of algorithms when deciding on a strategy for a given product. The proposed framework goes beyond the black-box approach, offering actionable insights to guide strategic organizational planning and workforce development efforts.

Even though it was successful, there were a few things to note. First, the analysis was restricted to a single workforce database, which could restrict the applicability of the results to other industries, organizational cultures, and geographic areas. Second, this study concentrated mainly on structured tabular data rather than unstructured data, such as employee communications, performance feedback, and data from the collaboration network. Third, while the study considered several aspects of demographic and organizational fairness, there are more aspects of fairness in both categories that could be used to assess responsible AI behavior.

The proposed framework could be expanded in the future to leverage multimodal information about the workforce, such as textual, behavioral, and communication-based data. Future enhancements in predictive performance and insights generation for organizations could stem from the collaborative integration of transformer-based tabular architectures, GNNs, and temporal workforce analytics. Furthermore, longitudinal research can deepen the understanding of patterns and evolution of employees' performances over time, which would enable better comprehension of workforce development and productivity in non-office settings such as remote and hybrid work.

In summary, the results obtained in this study show that the designed Explainable TabNet Framework is a helpful tool that is comprehensible, strong, and provides fairness in the prediction of employees' performance remotely. The authors' results provide both methodological contributions and insights for the emerging space of workforce analytics by successfully integrating key deep tabular learning and explainable AI technologies within a cohesive framework, including statistical validation and fairness auditing. The findings suggest that explainable AI could be a valuable asset to help inform data-driven HRM practices and responsible organizational decision-making in increasingly digital workplaces.

6. Conclusion

In this study, the authors proposed an Explainable TabNet Framework for predicting the performance of a distant workforce with a comparable database that consisted of 100,000 employee records. The proposed framework merged all four concepts of attention in deep tabular learning, explainability (via SHAP), robustness validation, statistical significance testing, and fairness auditing in a single predictive analytics pipeline. By experimental results, the effectiveness framework had an R^2score of 0.9880, RMSE was 0.1549 and the Pearson correlation coefficient was 0.9956, showing a great performance with significant improvement, compared to the traditional machine-learning methods like Linear Regression, Ridge Regression, Random Forest, Extra Trees, Gradient Boosting, XGBoost and LightGBM. Explainability analysis revealed that the following drivers of employee performance were most influential in influencing employee performance: employee satisfaction, employee monthly salary, involvement in projects, participation in training, and job tenure in the organization, which provides valuable managerial insights in optimizing the workforce. The framework was validated for its robustness, reliability, and generalizability through five-fold cross-validation and multiple statistical analyses. In addition, fairness testing showed that the model could be administered fairly for every evaluation group based on gender, department, and education level, which was found to be consistent for all groups, suggesting ethical appropriateness for organizational use. In summary, the results demonstrate the capabilities of the XDL algorithms to facilitate transparent, accurate, and responsible WFL. Other multimodal organizational information, temporal changes in the workforce, and advanced transformer-based architectures could be incorporated into future research to improve predictive performance and decision support capabilities in remote and hybrid settings.

Supplementary Materials: Not applicable.

Author Contributions: Conceptualization, Z.Z.; methodology, Z.Z.; software, Z.Z.; validation, Z.Z., X.B. and S.Y.; formal analysis, Z.Z.; investigation, Z.Z.; resources, Z.Z.; data curation, Z.Z.; writing—original draft preparation, Z.Z.; writing—review and editing, Z.Z., X.B. and S.Y.; visualization, Z.Z.; supervision, X.B. and S.Y.; project administration, Z.Z.; funding acquisition, Z.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This study received no external funding.

Institutional Review Board Statement: This study has not included any identifiable personal information, direct contact with human subjects, and/or intervention with human or animal subjects, and therefore there was no need for ethical review and approval by an Institutional Review Board.

Informed Consent Statement: Not applicable.

Data Availability: The data were obtained from the source mentioned in [60], which is open access. The processed data and code used to develop and support the findings of this article are available from the corresponding author upon reasonable request.

Acknowledgements: The authors would like to thank the eight providers of the publicly available workforce data used in this research. The authors would also like to thank their fellow colleagues for their helpful comments and reviews received in the development of the study.

Conflict of Interest: The authors do not declare any conflicts of interest.

References

[1] G. Deep, "Digital transformation's impact on organizational culture," *International Journal of Science and Research Archive,* vol. 10, no. 2, pp. 396-401, 2023.

[2] C. Mbonigaba and S. Sujatha, "Analyzing The Long-Term Impact of Remote Work on Organizational Efficiency And Employee Performance: Exploring Key Technologies And Management Practices," *Social Science Research Network,* 2024/08/06 2024, doi: 10.2139/ssrn.4993038.

[3] K. Ball, "Electronic Monitoring and Surveillance in the Workplace," Publications Office of the European Union, Luxembourg, JRC125716 978-92-76-43340-8, 2021. [Online]. Available: https://publications.jrc.ec.europa.eu/repository/handle/JRC125716

[4] T. Pianese, L. Errichiello, and J. V. da Cunha, "Organizational control in the context of remote working: A synthesis of empirical findings and a research agenda," *European Management Review,* vol. 20, no. 2, pp. 326-345, 2023/06/01 2023, doi: https://doi.org/10.1111/emre.12515.

[5] T. D. N. Vuong and L. T. Nguyen, "The Key Strategies for Measuring Employee Performance in Companies: A Systematic Review," *Sustainability,* vol. 14, no. 21, p. 14017, 2022. [Online]. Available: https://www.mdpi.com/2071-1050/14/21/14017.

[6] A. Ortynska, "Human performance management systems in hybrid and remote work environments," *Актуальні питання економічних наук,* no. 16, 10/09 2025, doi: 10.5281/zenodo.17305765.

[7] J. Tawalbeh, "Remote and Hybrid Work Models: Enhancing Employee Engagement and Redefining Performance Management in a New Era," *Journal of Posthumanism,* vol. 5, 04/09 2025, doi: 10.63332/joph.v5i3.715.

[8] E. Lee, "The Moderating role of Organizational Culture in the Relationship Between Job Autonomy and Innovative Behavior," *Review of Public Personnel Administration,* p. 0734371X251342001, 2025, doi: 10.1177/0734371X251342001.

[9] H. Tubussum, H. M. Hashim, M. U. Javaid, and M. Marimuthu, "Impact of demographic variables on interplay of workplace bullying, organizational culture, and organizational performance: SEM-based multi group analysis," *WORK,* vol. 80, no. 3, pp. 1304-1318, 2025/03/01 2025, doi: 10.1177/10519815241289761.

[10] B. A. Khan, "Performance Management Systems in Remote Work Settings: Redesigning Accountability Mechanisms," *Recent Issues In Human Resource Management,* vol. 9, no. 1, 2025.

[11] M. m. Shukhratov and J. Baxodirov, "MODERN TECHNOLOGIES AND METHODS FOR EMPLOYEE EVALUATION: PRACTICAL OPPORTUNITIES AND CONTEMPORARY CHALLENGES," *Research Focus International Scientific Journal,* vol. 4, no. 5, pp. 87-92, 2025/06/10 2025, doi: 10.5281/zenodo.15629871.

[12] R. Islam, E. French, and M. Ali, "Subjectivity in performance review systems: unveiling the mechanisms of gender bias and leadership disparity in the information technology industry," *The International Journal of Human Resource Management,* vol. 36, no. 22, pp. 4039-4073, 2025/12/16 2025, doi: 10.1080/09585192.2026.2623102.

[13] J. W. Park, "Out of sight, out of mind?: telework and gender inequality in performance appraisals," *Public Management Review,* pp. 1-34, 2025/11/26 2025, doi: 10.1080/14719037.2025.2592766.

[14] F. Toscano, S. Zappalà, and S. Donati, "The Dark Side of Remote Work: A Critical Review and a Call to Action to Address Remote Work Inequalities," *INFORMING SCIENCE,* vol. 28, pp. 1-20, 2025.

[15] S. Somavarapu and E. Priyanshi, "Building Scalable Data Science Pipelines for Large-Scale Employee Data Analysis," *J. Quantum Sci. Technol,* vol. 2, no. 1, pp. 446-470, 2025.

[16] J. S. Kanwar *et al.*, "A Scalable Machine Learning Framework for Predictive Analytics and Employee Performance Enhancement in Large Enterprises," (in eng), *International Journal of Advanced Computer Science & Applications,* vol. 16, no. 8, pp. 673-684, 2025, doi: 10.14569/ijacsa.2025.0160867.

[17] Z. Luo and K. Lu, "Application research and effectiveness evaluation mechanism of hybrid intelligent algorithm integrating cognitive computing and deep learning for dynamically adjusting employee performance evaluation in multi-scale organizational networks," *Discover Artificial Intelligence,* vol. 5, no. 1, p. 45, 2025/05/06 2025, doi: 10.1007/s44163-025-00285-x.

[18] H. Talebi, A. K. Bardsiri, and V. K. Bardsiri, "Developing a hybrid machine learning model for employee turnover prediction: Integrating LightGBM and genetic algorithms," *Journal of Open Innovation: Technology, Market, and Complexity,* vol. 11, no. 2, p. 100557, 2025/06/01/ 2025, doi: https://doi.org/10.1016/j.joitmc.2025.100557.

[19] A. Uppal, Y. Awasthi, and A. Srivastava, "Machine learning-based approaches for enhancing human resource management using automated employee performance prediction systems," *International Journal of Organizational Analysis,* vol. 33, no. 8, pp. 2307-2346, 2024, doi: 10.1108/ijoa-07-2024-4643.

[20] G. Vijh, N. Sharma, S. Tiwari, S. Vijh, and A. Sao, "Predicting Accurate Employee Performance: An Evaluation of Regression Models," *Procedia Computer Science,* vol. 259, pp. 433-442, 2025/01/01/ 2025, doi: https://doi.org/10.1016/j.procs.2025.03.345.

[21] A. M. Căvescu and N. Popescu, "Predictive Analytics in Human Resources Management: Evaluating AIHR's Role in Talent Retention," *AppliedMath,* vol. 5, no. 3, p. 99, 2025. [Online]. Available: https://www.mdpi.com/2673-9909/5/3/99.

[22] A. Kovari, "AI for Decision Support: Balancing Accuracy, Transparency, and Trust Across Sectors," *Information,* vol. 15, no. 11, p. 725, 2024. [Online]. Available: https://www.mdpi.com/2078-2489/15/11/725.

[23] H. Qaium, I. Md. Zafor, and R. Md. Musfiqur, "A META DATA-DRIVEN DECISION SUPPORT IN HUMAN CAPITAL MANAGEMENT: REVIEWING HRIS AND PREDICTIVE ANALYTICS INTEGRATION," *ASRC Procedia: Global Perspectives in Science and Scholarship,* vol. 1, no. 01, pp. 215-246, %04/%29 2025, doi: 10.63125/xgew7q22.

[24] M. Venugopal, V. Madhavan, R. Prasad, and R. Raman, "Transformative AI in human resource management: enhancing workforce planning with topic modeling," *Cogent Business & Management,* vol. 11, no. 1, p. 2432550, 2024/11/26 2024, doi: 10.1080/23311975.2024.2432550.

[25] V. Chamola, V. Hassija, A. R. Sulthana, D. Ghosh, D. Dhingra, and B. Sikdar, "A Review of Trustworthy and Explainable Artificial Intelligence (XAI)," *IEEE Access,* vol. 11, pp. 78994-79015, 2023, doi: 10.1109/ACCESS.2023.3294569.

[26] R. Okonkwo, A. Folorunso, F. Ogundipe, and C. Y. Tettey, "Explainable Artificial Intelligence (Al) through human-AI collaborative frameworks: Quantifying trust and interpretability in high-stakes decisions," *Comput. Sci. IT Res. j,* vol. 6, no. 5, pp. 333-354, 2025.

[27] C. J. Ihejirika, "Explainable artificial intelligence in deep learning models for transparent decision-making in high-stakes applications," *Int J Res Publ Rev,* vol. 6, pp. 1924-40, 2025.

[28] S. Somvanshi, S. Das, S. Javed, G. Antariksa, and A. Hossain, "A Survey on Tabular Data: From Tree-based Methods to Tabular Deep Learning," *ACM Comput. Surv.,* 2026, doi: 10.1145/3807777.

[29] S. M. Venkateswara and J. Padmanaban, "Interpretable deep learning models for independent fertilizer and crop recommendation," *Scientific Reports,* vol. 15, no. 1, p. 41721, 2025/11/24 2025, doi: 10.1038/s41598-025-26910-4.

[30] A. de Santana Correia and E. L. Colombini, "Attention, please! A survey of neural attention models in deep learning," *Artificial Intelligence Review,* vol. 55, no. 8, pp. 6037-6124, 2022/12/01 2022, doi: 10.1007/s10462-022-10148-x.

[31] Y. Qin, D. Song, H. Chen, W. Cheng, G. Jiang, and G. Cottrell, "A dual-stage attention-based recurrent neural network for time series prediction," *arXiv preprint arXiv:1704.02971,* 2017.

[32] A. Hernández and J. M. Amigó, "Attention Mechanisms and Their Applications to Complex Systems," *Entropy,* vol. 23, no. 3, p. 283, 2021. [Online]. Available: https://www.mdpi.com/1099-4300/23/3/283.

[33] Q. Zuo, R. Liang, W. Pan, and Z. Yin, "MGRF-Net: Situation Awareness Prediction for Remote Tower Controllers Based on Multimodal Physiological Data," *Aerospace,* vol. 13, no. 5, p. 452, 2026. [Online]. Available: https://www.mdpi.com/2226-4310/13/5/452.

[34] S. Karuppiah, R. S, R. J. E. B, S. A. A. T, S. A, and S. Kumaran, "Employee Retention Analytics: Predictive Modeling for Workforce Stability and Talent Management," in *2025 IEEE 4th World Conference on Applied Intelligence and Computing (AIC)*, 26-27 July 2025 2025, pp. 420-425, doi: 10.1109/AIC66080.2025.11212141.

[35] J. P. Jiang, S. Y. Liu, H. R. Cai, Q. L. Zhou, and H. J. Ye, "Representation Learning for Tabular Data: A Comprehensive Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 48, no. 6, pp. 6488-6508, 2026, doi: 10.1109/TPAMI.2026.3657217.

[36] L. Zapata, G. Ibarra, and P.-H. Blancher, "Engaging new ways of work: the relevance of flexibility and digital tools in a post-COVID-19 era," *Journal of Organizational Effectiveness: People and Performance,* vol. 11, no. 1, pp. 1-17, 2023, doi: 10.1108/JOEPP-04-2022-0079.

[37] S. V. Pillai and J. Prasad, "Investigating the key success metrics for WFH/remote work models," *Industrial and Commercial Training,* vol. 55, no. 1, pp. 19-33, 2022, doi: 10.1108/ict-07-2021-0053.

[38] M. Md Harun-Or-Rashid and K. Sai Praveen, "DATA-DRIVEN APPROACHES TO ENHANCING HUMAN–MACHINE COLLABORATION IN REMOTE WORK ENVIRONMENTS," (in en), *International Journal of Business and Economics Insights,* vol. 2, no. 3, pp. 47-83, %09/%28 2022, doi: 10.63125/wt9t6w68.

[39] S. Varghese, L. RB, G. M. Patil, and N. S. Thomas, "Predictive HR Analytics: Forecasting Employee Turnover Through Machine Learning Models," *Advances in Consumer Research,* vol. 2, no. 4, 2025.

[40] M. Hariani and R. Mardikaningsih, "Work Schedule Flexibility and Job Satisfaction in a Global Corporate Environment," *International Journal of Service Science, Management, Engineering, and Technology,* vol. 7, no. 1, pp. 12-16, 02/28 2025. [Online]. Available: https://ejournalisse.com/index.php/isse/article/view/143.

[41] M. S. Hasibuan and R. R. N. Fikri, "Sectoral Disparities in the Digital Workplace: A Comparative Statistical Analysis of Remote Work Patterns and Productivity Scores in the Evolving Digital Society," *Journal of Digital Society,* vol. 1, no. 3, pp. 244-257, 09/01 2025, doi: 10.63913/jds.v1i3.11.

[42] M. N. Razali, I. Norizuandi, R. Hanapi, N. Mohd Zamri, and S. Abdul Manaf, "Exploring Employee Working Productivity: Initial Insights from Machine Learning Predictive Analytics and Visualization," *Journal of Computing Research and Innovation,* vol. 8, no. 2, pp. 235-245, 09/01 2023, doi: 10.24191/jcrinn.v8i2.362.

[43] Q. C. Song, I. S. Oh, Y. Kim, and C. So, "Revisiting the nature and strength of the personality–job performance relations: New insights from interpretable machine learning," *Journal of Applied Psychology,* vol. 110, no. 1, pp. 1-26, 2025, doi: 10.1037/apl0001218.

[44] D. Maharani *et al.*, "Comparing Decision Tree and Optimized LightGBM for Attrition Prediction," *Journal of Applied Informatics and Computing,* vol. 10, no. 3, pp. 2165-2177, 06/08 2026, doi: 10.30871/jaic.v10i3.12678.

[45] R. Gustriansyah, S. Puspasari, A. Sanmorino, N. Suhandi, and D. Sartika, "Tree-based models and hyperparameter optimization for assessing employee performance," *Indones. J. Electr. Eng. Comput. Sci,* vol. 38, no. 1, p. 569, 2025.

[46] L. H. Palivela, H. Badaya, and A. A. Micheal, "A Hybrid Machine Learning Approach for Predicting Employee Attrition and Tenure Forecasting," in *AI Influences on Employee Well-Being*: IGI Global, 2026.

[47] J. Nwoke, "Harnessing predictive analytics, machine learning, and scenario modeling to enhance enterprise-wide strategic decision-making," *International Journal of Computer Applications Technology and Research,* vol. 14, no. 4, pp. 123-136, 2025.

[48] B. Buyuktanir, G. K. Baydogmus, E. E. Ulku, and K. Yildiz, "Enhancing Employee Retention Prediction Through Federated Learning and Associative Classification using FedCBA: A Case Study on IBM HR Analytics," *Arabian Journal for Science and Engineering,* 2026/04/01 2026, doi: 10.1007/s13369-026-11252-8.

[49] Z. Chen, "Algorithmic human resource management as a mode of algorithmic governance: transparency, fairness, and human agency in the digital workplace," *Humanities and Social Sciences Communications,* vol. 13, no. 1, p. 594, 2026/03/16 2026, doi: 10.1057/s41599-026-06989-4.

[50] R. M. Guion, *Assessment, measurement, and prediction for personnel decisions*. Routledge, 2011.

[51] D. E. Mathew, D. U. Ebem, A. C. Ikegwu, P. E. Ukeoma, and N. F. Dibiaezue, "Recent Emerging Techniques in Explainable Artificial Intelligence to Enhance the Interpretable and Understanding of AI Models for Human," *Neural Processing Letters,* vol. 57, no. 1, p. 16, 2025/02/07 2025, doi: 10.1007/s11063-025-11732-2.

[52] K. Kalasampath, K. N. Spoorthi, S. Sajeev, S. S. Kuppa, K. Ajay, and A. Maruthamuthu, "A Literature Review on Applications of Explainable Artificial Intelligence (XAI)," *IEEE Access,* vol. 13, pp. 41111-41140, 2025, doi: 10.1109/ACCESS.2025.3546681.

[53] P. Biecek and T. Burzykowski, *Explanatory model analysis: explore, explain, and examine predictive models*. Chapman and Hall/CRC, 2021.

[54] L. Ayanponle, C. G. Okatta, and D. Ajiga, "AI-powered HR analytics: Transforming workforce optimization and decision-making," *International Journal of Science and Research Archive,* vol. 5, no. 2, pp. 338-346, 2022.

[55] G. Anakpo, Z. Nqwayibana, and S. Mishi, "The Impact of Work-from-Home on Employee Performance and Productivity: A Systematic Review," *Sustainability,* vol. 15, no. 5, p. 4529, 2023. [Online]. Available: https://www.mdpi.com/2071-1050/15/5/4529.

[56] S. Somvanshi, S. Das, S. A. Javed, G. Antariksa, and A. Hossain, "A survey on deep tabular learning," *arXiv preprint arXiv:2410.12034,* 2024.

[57] Y. Gorishniy, I. Rubachev, V. Khrulkov, and A. Babenko, "Revisiting Deep Learning Models for Tabular Data," presented at the Advances in Neural Information Processing Systems, 2021. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2021/file/9d86d83f925f2149e9edb0ac3b49229c-Paper.pdf.

[58] K. Ma, J. Zhang, X. Huang, and M. Wang, "Leveraging transformer models to predict cognitive impairment: accuracy, efficiency, and interpretability," *BMC Public Health,* vol. 25, no. 1, p. 504, 2025/02/07 2025, doi: 10.1186/s12889-025-21762-z.

[59] Z. Yiyuan, L. Chang, L. Xintong, X. Xiaochen, and S. Na, "An interpretable attention-based TabTransformer framework with feature fusion for green architecture classification," *Scientific Reports,* 2026/06/19 2026, doi: 10.1038/s41598-026-50347-y.

[60] M. Finsterwald. *Employee Performance and Productivity Data*, Kaggle. [Online]. Available: https://www.kaggle.com/datasets/mexwell/employee-performance-and-productivity-data