Employee Attrition Prediction Using Temporal Convolutional Network (TCN) on HR Time Series Dataset to Support Employee Retention Strategies

Muhammad Farrell Saifulloh[1], Ervin Yohannes[2]

[1,2] Informatics Engineering Department, Faculty of Engineering, Universitas Negeri Surabaya

[1]muhammadfarrell.22041@mhs.unesa.ac.id

[2]ervinyohannes@unesa.ac.id

Abstract— Employee attrition continues to be a major challenge for organizations because high turnover increases recruitment costs, disrupts business operations, and affects workforce stability. Predictive analytics provides an opportunity to identify employees who are at risk of leaving before turnover occurs. A Human Resources dataset collected between 2018 and 2024 served as the basis for evaluating the Temporal Convolutional Network (TCN) in employee attrition prediction. The dataset contains 16,050 employee records described by 25 attributes covering demographic information, job characteristics, compensation, and work experience. Data preprocessing included feature selection, categorical encoding, normalization, and sequence preparation before model training. The proposed model was evaluated using three train–validation–test splitting strategies and assessed with Accuracy, Precision, Recall, F1-score, and Area Under the ROC Curve (AUC). Experimental results show that the TCN model consistently achieved high predictive performance across different experimental settings, with the best configuration obtaining an Accuracy of 98.35%, an F1-score of 0.8976, and an AUC of 0.9819. These findings demonstrate that TCN provides stable and reliable employee attrition prediction and has the potential to support Human Resource Analytics through earlier identification of employees with a higher risk of leaving the organization.

Keywords—Employee Attrition, Temporal Convolutional Network, Deep Learning, Human Resource Analytics, Employee Retention.

I. Introduction

Employee attrition remains a significant challenge for organizations because the loss of experienced employees affects productivity, increases recruitment and training costs, and disrupts organizational performance. A recent report estimated that voluntary employee turnover costs businesses in the United States approximately one trillion dollars annually, highlighting the substantial financial impact of employee attrition [1]. Beyond financial losses, retaining skilled employees has become increasingly important as organizations continue to adopt digital transformation and data-driven management practices.

Organizations are also placing greater emphasis on employee experience as an essential component of workforce sustainability. IBM highlights employee experience as an important factor in building workforce engagement and strengthening employees' commitment to the organization. A positive employee experience is also associated with lower turnover, making it an important consideration in modern Human Resource management [2]. Consequently, identifying employees who are at risk of leaving before resignation occurs has become an important objective within Human Resource (HR) Analytics.

Recent advances in artificial intelligence have encouraged the adoption of predictive analytics to support employee retention. Deep learning models have demonstrated promising performance in employee attrition prediction because they are capable of learning complex relationships among demographic, occupational, and behavioral variables without relying on manually defined decision rules [3]. Early attrition predictions give HR managers additional information for planning retention initiatives and prioritizing employees who may need further support.

Temporal Convolutional Network (TCN) has recently emerged as an effective deep learning architecture for sequential prediction problems. By combining causal convolution and dilated convolution, TCN is able to capture long-term temporal dependencies while maintaining efficient parallel computation. Previous research reported that a Bidirectional Temporal Convolutional Network achieved excellent predictive performance for employee attrition prediction while also improving model interpretability through explainable artificial intelligence techniques [5].

Although previous studies have demonstrated the potential of deep learning for employee attrition prediction, relatively few have investigated the robustness of TCN using multiple train–validation–test splitting strategies on a combined Human Resources dataset. Addressing this gap, the present study evaluates the predictive capability of a Temporal Convolutional Network under different experimental configurations to examine its effectiveness and consistency for employee attrition prediction.

II. Literature Review

A. Employee Attrition and HR Analytics

Employee attrition prediction has become an active research topic because organizations increasingly rely on predictive analytics to support workforce planning and retention decisions. Early studies mainly focused on conventional machine learning algorithms, including Logistic Regression, Support Vector Machine, Decision Tree, and Random Forest. These methods generally provide satisfactory classification performance; however, their predictive capability may decrease when complex nonlinear relationships exist among employee attributes [6].

The growing availability of employee data has encouraged the adoption of deep learning techniques for HR analytics. Deep learning models are capable of automatically extracting feature representations from employee records, allowing more comprehensive analysis of relationships among demographic, occupational, and organizational variables. Previous studies have demonstrated that deep learning can improve employee attrition prediction compared with several traditional machine learning approaches [3].

Among recent deep learning architectures, the Temporal Convolutional Network (TCN) has received increasing attention because of its ability to model sequential data efficiently. TCN utilizes causal convolution to preserve chronological information and dilated convolution to expand the receptive field without substantially increasing computational complexity. These characteristics enable the model to learn long-term dependencies while maintaining stable gradient propagation during training [5].

Recent work by Alghosoun *et al.* proposed a Bidirectional Temporal Convolutional Network integrated with explainable artificial intelligence techniques for employee attrition prediction. Their findings demonstrated that TCN achieved excellent predictive performance while also providing interpretable prediction results that support managerial decision-making [5]. In addition, recent studies emphasize that integrating predictive analytics into HR management contributes to more effective employee retention strategies and supports evidence-based organizational policies [7].

Despite these encouraging findings, differences in dataset composition, feature engineering strategies, and evaluation protocols make it difficult to directly compare the reported performance across studies. Furthermore, only limited research has examined the robustness of Temporal Convolutional Networks using multiple data-splitting strategies on combined Human Resources datasets. This limitation provides the motivation for the present study.

B. Temporal Convolutional Network (TCN)

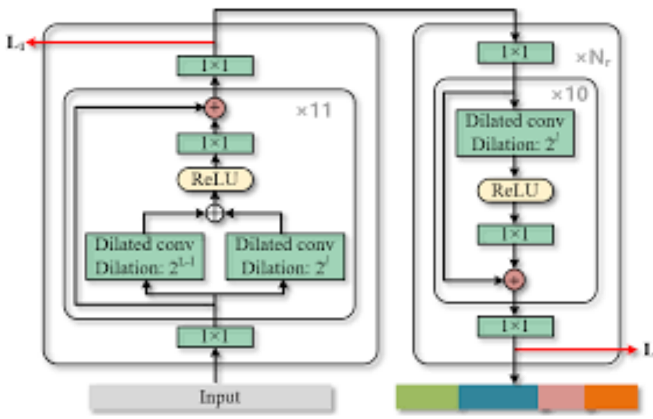

FIG. 1 TCN ARCHITECTURE

TCN is a convolutional architecture designed for sequential data that employs dilated causal convolution and residual connections to process entire sequences in parallel, overcoming the vanishing gradient limitation inherent in recurrent architectures [5]. The output at layer l and time step t is computed as:

$$zt(l) = \sigma(i = 0\sum k - 1\boldsymbol{W}i(l)\cdot \boldsymbol{x}t - d\cdot i + \boldsymbol{b}(l))(1)$$

FIG. 2 TCN FORMULA

where d denotes the dilation factor, K the kernel size, and σ the ReLU activation. Using dilation rates that increase exponentially allows the model to learn temporal patterns over both short and long time intervals without substantially increasing computational cost.

C. Prior Work

TABLE I PRIOR RESEARCH ON EMPLOYEE ATTRITION PREDICTION

Ref	Method	Dataset	Best Result
[3]	DNN	IBM HR	Acc ≈94% (balanced)
[4]	LSTM+BMBOA	IBM HR	Acc 96.68%, F1 96.62%
[5]	Bi-TCN+GAN	IBM+Kaggle	Acc 97.83%, AUC 96%
[6]	Transformer+GAN	IBM+Kaggle	F1 91.58%, AUC 96.32%
This	TCN + Sliding Window	HR Time Series	Acc 97.69%, F1 0.8436

III. RESEARCH METHODOLOGY

A. Research Framework

This study follows a structured workflow that begins with dataset preparation and continues through preprocessing, feature engineering, model training, and performance evaluation. Each stage contributes to ensuring that the TCN model is developed and assessed under consistent experimental conditions. Each stage was designed to ensure that the proposed Temporal Convolutional Network (TCN) model received consistent and representative input data before the learning process.

B. Dataset

A combined Human Resources dataset was used to evaluate the proposed approach. It contains 3,000 employee records described by 15 attributes covering demographic information, job characteristics, compensation, and work-related factors associated with employee attrition. The dataset integrates employee demographic information, job characteristics, compensation, and work-related attributes that are commonly associated with employee attrition. The target variable is Attrition, which represents whether an employee remains in the company or leaves the organization. Before model development, the dataset was examined to ensure data consistency and suitability for sequential learning.

TABLE II DATASET CHARACTERISTICS

Attribute	Value
Total Records	16,050
Year Range	2018–2024 (7 years)
Original Features	25 (13 numeric, 12 categorical)
Target Variable	Yes/No

The study used a raw Human Resources dataset containing 16,050 employee records collected between 2018 and 2024. The dataset consists of 25 attributes describing employee demographics, job characteristics, compensation, work experience, and organizational factors. Employee attrition (Yes/No) was used as the target variable for prediction. Before model development, the raw data underwent preprocessing and feature selection to obtain the variables used in the TCN model.

C. Preprocessing and Feature Engineering

Before training the network, the employee records were examined to improve data quality and ensure consistent model input. The preprocessing stage included checking for variables, and normalizing numerical features. Categorical variables were transformed into numerical representations through encoding, while numerical features were normalized to reduce differences in feature scales. These preprocessing steps provide more stable input for the TCN model and contribute to a more reliable learning process [25].

D. Feature Engineering

Employee attributes associated with attrition were retained and prepared before model training. After preprocessing, the selected variables were organized into sequential inputs so that the TCN model could learn patterns from employee records while maintaining the chronological order of the data.

TABLE III: DATA SPLIT STRATEGIES

Strategy	Train	Val	Test
80:10:10	6,295	787	787
70:15:15	5,508	1,181	1,180
60:20:20	4,721	1,574	1,574

E. Temporal Convolutional Network Model

Employee attrition prediction was carried out using a Temporal Convolutional Network (TCN), which combines causal and dilated convolutions to capture temporal dependencies efficiently. TCN applies causal convolution to ensure that predictions are generated only from current and previous observations, while dilated convolution expands the receptive field without substantially increasing computational complexity. Residual connections were incorporated to improve gradient propagation and support stable training in deeper network architectures [32].

Several training configurations were evaluated by varying the number of epochs and batch sizes. In addition, three train–validation–test splitting strategies (80:10:10, 70:15:15, and 60:20:20) were adopted to examine model robustness under different proportions of training and testing data.

TABLE IV TCN ARCHITECTURE CONFIGURATION

Layer / Block	Parameters	Details
Input Reshape	(batch, 1, n_feat)	Tabular → 3D tensor
Residual Block 1	64 filters, d=1	Causal Conv, ReLU, LN
Residual Block 2	64 filters, d=2	Causal Conv, ReLU, LN
Residual Block 3	32 filters, d=4	Causal Conv, ReLU, LN
GlobalAvgPool1D	—	Temporal aggregation
Dense + Dropout	32 units, ReLU / 0.20	Hidden classification layer
Output Dense	1 unit, sigmoid	Binary attrition probability

TABLE V TRAINING HYPERPARAMETER SCENARIOS

Scenario	Epochs	Batch	LR	Optimizer
Test 1	100	32	0.001	Adam
Test 2	50	16	0.001	Adam
Test 3	150	64	0.001	Adam

The same training strategy was used in all experimental scenarios to ensure a fair comparison. EarlyStopping (patience = 10) was employed to prevent unnecessary training after convergence while restoring the best-performing weights. The learning rate was adjusted automatically using ReduceLROnPlateau (factor = 0.5, patience = 5, minimum learning rate = 1×10^{-6}), and the classification threshold was determined from the validation set. Model performance was then assessed using Accuracy, Precision, Recall, F1-score, and the Area Under the ROC Curve (AUC).

IV. RESULTS AND DISCUSSION

A. Learning Curves

Figures 1 and 2 illustrate the training and validation accuracy and loss obtained by the TCN model using the 80:10:10 data split. The learning curves show a steady convergence throughout the training process, with only a small gap between the training and validation results. This behavior suggests that the model learned effectively without noticeable overfitting, while the combination of EarlyStopping and ReduceLROnPlateau helped maintain stable optimization during training..

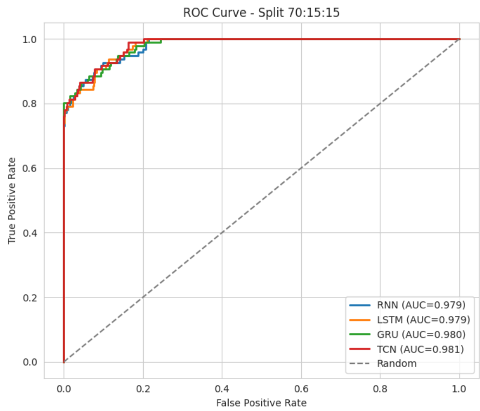

FIG. 3 TCN TRAINING CURVE (80:10:10) — ACCURACY

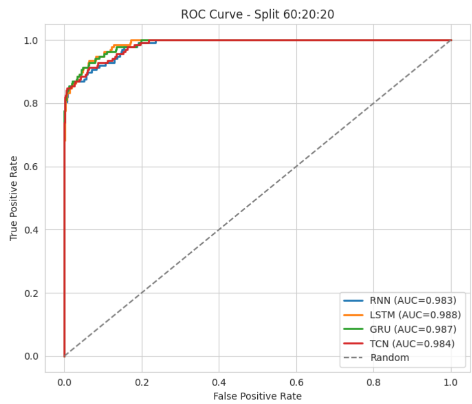

FIG. 4 TCN TRAINING CURVE (80:10:10) — LOSS

B. Confusion Matrix Analysis

Figures 3–5 show the confusion matrices of the TCN model obtained from the three train–validation–test splitting strategies. Across all experiments, the model correctly classified most employees in the non-attrition class, resulting in a high number of True Negatives (TN) and relatively few False Positives (FP). Most employees who remained with the organization were classified correctly, resulting in a large number of True Negative predictions. Identifying employees who eventually resigned proved to be more difficult because the attrition class accounted for only about 8.4% of the available records, leading to a small number of False Negative cases.

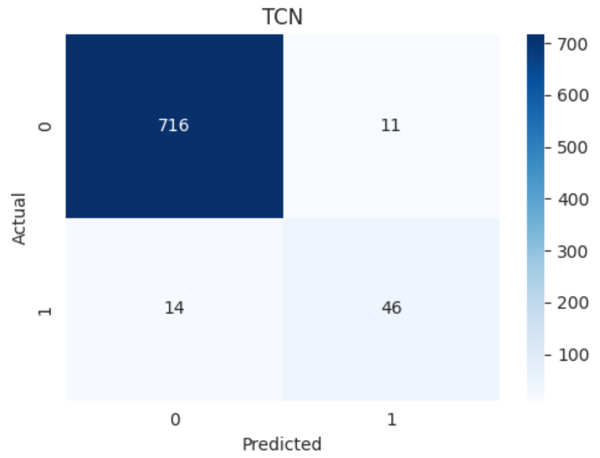

FIG. 5 CONFUSION MATRIX — SPLIT 80:10:10

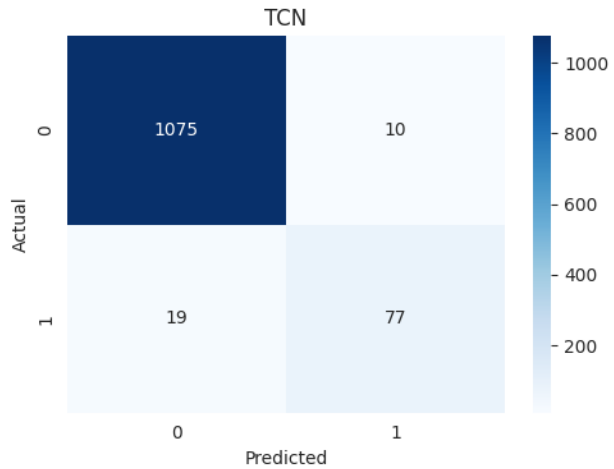

FIG. 6 CONFUSION MATRIX — SPLIT 70:15:15

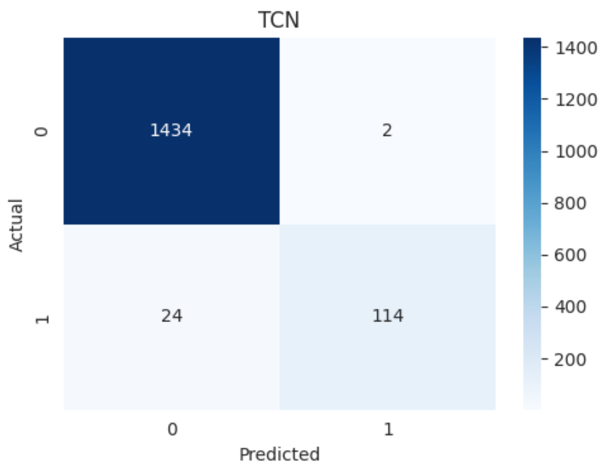

FIG. 7 CONFUSION MATRIX — SPLIT 60:20:20

C. ROC Curve Analysis

Figures 6–8 present the ROC curves of the TCN model for the three train–validation–test splitting strategies. In each experiment, the ROC curve remains well above the diagonal reference line, indicating that the model effectively distinguishes employees with attrition from those who remain in the organization. The AUC values, ranging from 0.980 to 0.984, show only minor variation across the different data splits, suggesting that the predictive performance of the TCN model remains stable under different training and testing configurations.

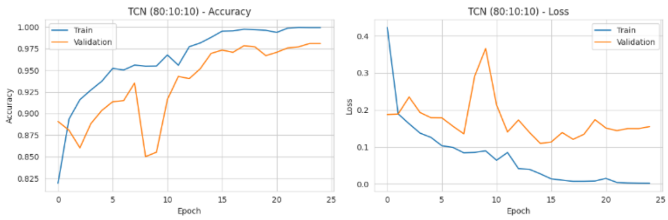

FIG. 8 ROC CURVE — SPLIT 80:10:10 (TCN AUC=0.980)

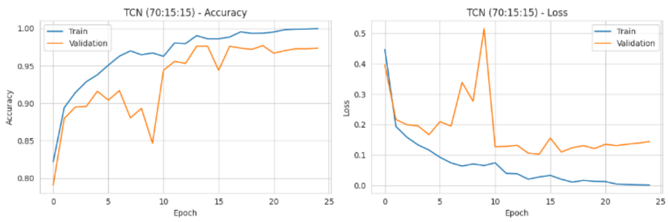

FIG. 9 ROC CURVE — SPLIT 70:15:15 (TCN AUC=0.981)

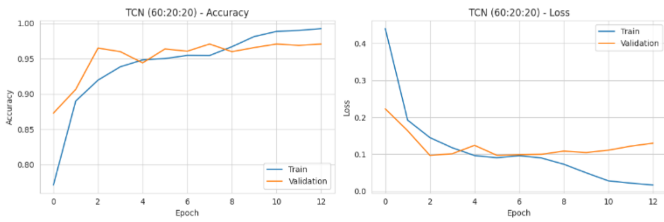

FIG. 10 ROC CURVE — SPLIT 60:20:20 (TCN AUC=0.984)

D. Quantitative Evaluation Results

TABLE VI TCN RESULTS — TEST 1 (100 EPOCHS, BS=32)

Split	Thr	Acc (%)	Precision	Recall	F1	AUC

80:10:10	0.94	97.97	0.9583	0.7667	0.8519	0.9802
70:15:15	0.93	97.29	0.8556	0.8021	0.8280	0.9810
60:20:20	0.72	97.14	0.8298	0.8478	0.8387	0.9844
Average	—	97.47	0.8812	0.8055	0.8395	0.9819

TABLE VII TCN RESULTS — TEST 2 (50 EPOCHS, BS=16)

Split	Thr	Acc (%)	Precision	Recall	F1	AUC
80:10:10	0.91	97.97	0.9583	0.7667	0.8519	0.9796
70:15:15	0.94	97.46	0.8667	0.8125	0.8387	0.9796
60:20:20	0.69	96.57	0.7800	0.8478	0.8125	0.9850
Average	—	97.33	0.8683	0.8090	0.8344	0.9814

TABLE VIII TCN RESULTS — TEST 3 (150 EPOCHS, BS=64)

Split	Thr	Acc (%)	Precision	Recall	F1	AUC
80:10:10	0.89	96.82	0.8070	0.7667	0.7863	0.9816
70:15:15	0.93	97.54	0.8851	0.8021	0.8415	0.9795
60:20:20	0.84	98.35	0.9828	0.8261	0.8976	0.9845
Average	—	97.57	0.8916	0.7983	0.8418	0.9819

E. Cross-Scenario Comparison

TABLE IX TCN PERFORMANCE SUMMARY ACROSS ALL SCENARIOS

Scenario	Ep/BS	Avg Acc (%)	Avg F1	Avg AUC
Test 1	100/32	97.47	0.8395	0.9819
Test 2	50/16	97.33	0.8344	0.9814
Test 3	150/64	97.57	0.8418	0.9819
Overall Avg	—	97.46	0.8386	0.9817

Performance differences among the three experimental settings were relatively small. Comparable Accuracy, F1-score, and AUC values were obtained under each train–validation–test split, showing that changes in the data partition had only a limited effect on prediction performance. The highest Accuracy and F1-score were recorded in Test 3 using the 60:20:20 split, although the improvement over the other configurations was modest. Similar results across all experiments reflect the robustness of the TCN model when applied to employee attrition prediction.

F. Dominant Attrition Factors and Retention Recommendations

TABLE X DOMINANT ATTRITION FACTORS

Factor	Significance
Panel Position	47.2% attrition rate on terminal year vs. 0% in preceding years
MonthlyIncome	Significant income gap between departing and retained employees
YearsAtCompany	Tenure < 3 years associated with substantially higher attrition risk
Age	Employees aged 20–35 exhibit higher career mobility and attrition rates
JobSatisfaction	Lower satisfaction scores correlate with increased resignation probability

Based on the factors identified as having the strongest association with employee attrition, several practical retention strategies can be considered by organizations. First, compensation policies may be reviewed regularly through salary benchmarking, performance-based incentives, and financial support for employees whose earnings are below the organizational average. Second, organizations can strengthen onboarding programs by providing mentoring during the first months of employment and establishing clear career development pathways with measurable objectives. Third, work–life balance may be improved by monitoring overtime practices, ensuring fair compensation for additional working hours, and expanding employee well-being initiatives. In addition, employee satisfaction can be enhanced through regular satisfaction surveys, recognition programs, and continuous improvements to the working environment. Finally, the trained TCN model can be incorporated into an HR analytics dashboard to provide continuous monitoring of employee attrition risk, allowing managers to identify high-risk employees earlier and implement appropriate retention actions before turnover occurs.

V. CONCLUSION

A Temporal Convolutional Network (TCN) was employed to predict employee attrition using Human Resources data collected over multiple years. Before model training, the employee records were prepared through preprocessing and feature selection, followed by evaluation using three train–validation–test split configurations. Comparable results across all experiments show that the model learned meaningful temporal information while maintaining stable prediction performance.

The findings indicate that the proposed model can effectively distinguish employees with a higher likelihood of attrition while maintaining reliable performance based on Accuracy, Precision, Recall, F1-score, and AUC. In addition, variables such as **MonthlyIncome**, **TotalWorkingYears**, **Age**, **JobSatisfaction**, and **OverTime** were identified as important factors associated with employee attrition. These variables provide valuable information for understanding employee turnover and may assist organizations in developing more targeted retention strategies.

Results obtained in this work show that the TCN model can be applied to employee attrition prediction with consistent performance across different experimental settings. Early prediction of employees with a higher risk of leaving gives HR managers additional information when planning retention programs and workforce allocation. Further investigation may involve data collected from different organizations, richer employee-related variables, and alternative deep learning models to examine whether similar performance can be maintained under broader conditions.

ACKNOWLEDGMENT

The authors express sincere gratitude to Mr. Ervin Yohannes, S.Kom., M.Kom., M.Sc., Ph.D. as thesis supervisor, and to Mrs. Paramitha Nerisafitra, S.ST., M.Kom. and Mrs. Anita Qoiriah, S.Kom., M.Kom. as thesis examiners, along with all faculty and staff of the Informatics Engineering Department, Universitas Negeri Surabaya, for their invaluable guidance and support.

REFERENCES

[1] S. McFeely and B. Wigert, "This Fixable Problem Costs U.S. Businesses $1 Trillion," Gallup Workplace, 2019.

[2] IBM Institute for Business Value, "The Employee Experience Index," IBM Corp., 2024.

[3] I. H. Najm, M. F. Abed, and G. M. Abdulsahib, "Employee Attrition Prediction Using Deep Neural Networks," Computers, vol. 10, no. 11, p. 141, 2021, doi: 10.3390/computers10110141.

[4] S. Ganapathisamy and V. Narayan, "LSTM-RNN with BMBOA for Employee Attrition Prediction," 2023.

[5] H. A. Alghosoun, M. F. I. Othman, and A. M. Ali, "A deep learning model based on bidirectional temporal convolutional network (Bi-TCN) for predicting employee attrition," Appl. Sci., vol. 15, no. 6, p. 2984, 2025, doi: 10.3390/app15062984.

[6] M. Türk, A. T. Gürdal, and A. Ö. Yildiz, "Predicting employee attrition: XAI-powered models for managerial decision-making," Inf. Dyn. Appl., vol. 13, no. 7, 2025.

[7] V. Nayak, S. Chaudhary, and C. Kumar, "The impact of HR analytics usage frequency on retention effectiveness," J. Innov. Entrepreneurship Res., vol. 7, no. 3, 2025.

[8] R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training recurrent neural networks," Proc. ICML, vol. 28, pp. 1310–1318, 2013.

[9] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.

[10] K. Cho et al., "Learning phrase representations using RNN encoder-decoder for statistical machine translation," Proc. EMNLP, pp. 1724–1734, 2014.

[11] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," JAIR, vol. 16, pp. 321–357, 2002.