

DOI: 10.5281/zenodo.20463412

EXPLAINABLE ARTIFICIAL INTELLIGENCE DRIVEN ORGANIZATIONAL REPUTATION PREDICTION FRAMEWORK USING HR ANALYTICS AND EMPLOYEE ATTRITION INTELLIGENCE

Abdulwahab M. Alsulami^{1*}

¹Department of Public Relations, College of Communication and Media, King Abdulaziz University, Jeddah,
Saudi Arabia

Received: 04/04/2026
Accepted: 20/05/2026

Corresponding Author: Abdulwahab M. Alsulami
(amsulami@kau.edu.sa)

ABSTRACT

Organizational reputation management has become an essential component influencing employee retention, workforce sustainability, and institutional progress in modern organizations. Poor workplace conditions, employee dissatisfaction, and higher attrition rates greatly affect employer branding and organizational reputation in digital and professional settings. Most organizations maintain structured employee advancement strategies, although continual employee attrition can result in considerable losses. Employee growth initiatives involve significant time and financial investment. When there is dissatisfaction within the organization as a result of employee attrition, it influences overall progression and interaction. To address the challenges, this paper introduces an Explainable Hybrid Deep Learning Framework for Organizational Reputation Assessment (XHDL-ORA) using HR analytics and workforce behavior intelligence. The proposed framework utilizes the HR Analytics Case Study dataset containing employee-related behavioral, performance, and workplace attributes suitable for organizational reputation assessment and attrition intelligence modeling. A correlation-based feature selection with Mutual Information analysis is used to identify the most informative and non-redundant workforce attributes influencing employee attrition behavior and institutional reputation. For attrition classification, a Hybrid Transformer-Bidirectional Capsule Network with Attention Mechanism is developed to efficiently learn complex workforce relationships, contextual feature dependencies, and hierarchical employee behavioral representations from heterogeneous HR analytics data. In addition, the Adam optimizer is integrated to improve convergence efficiency during the training process. Furthermore, SHAP-based Explainable Artificial Intelligence analysis is incorporated to improve model transparency by identifying the contribution and influence of workforce features toward employee attrition classification. Experimental analysis demonstrates that the proposed XHDL-ORA approach significantly improves classification performance in terms of accuracy, precision, recall, F1-score, and interpretability compared to existing methods. The developed XHDL-ORA model contributes toward intelligent organizational reputation assessment by enabling proactive workforce analytics, employee retention intelligence, and AI-driven institutional decision support systems.

KEYWORDS: HR Analytics, Employee Attrition Classification, Organizational Reputation Assessment, Hybrid Deep Learning, Capsule Network, Explainable Artificial Intelligence.

1. INTRODUCTION

Recently, Human Resource Management (HRM) has gained greater attention because of its essential contribution in promoting organizational growth and strengthening competitive advantages (Priya et al., 2026). It is broadly recognized that employees are the most valuable asset of every organization, and the overall achievement of the organization is closely associated with the skills, knowledge, and experiences of the employees. Employee attrition turned out to be a key focus area of human resources (HR) specialists, because it acts as a significant measure of organizational competitiveness (Benabou, A. and Touhami, F., 2026). Employee Attrition is the process in which employees exit the organization, thereby creating the need to hire new staff. Various catalysts for staff turnover namely job dissatisfaction, insufficient pay, poor working conditions, personal reasons, and others (Cui et al., 2026). Attrition can be labelled into two different kinds: voluntary and involuntary. Voluntary attrition happens when the best-performing staff members choose to leave the organization, even though the organization tries to retain them (Iorio and Morris, 2026). Conversely, Involuntary attrition refers to a situation where the employee is terminated from the position due to their poor performance or restructuring processes. To estimate the attrition rate, the number of existing employees is split by the average number of present employees during a specific period, which enables company managers to evaluate the company's overall efficiency (Bellora-Bienengräber et al., 2026). The potential to forecast when and why employees may exit an organization can greatly support human resource departments (HRDs) to take favorable policies to keep their valuable staffs by using effective retention strategies. The problem of predicting employee attrition has gained increasing attention, especially due to the financial burden and organizational disruption it can cause (Chakrabarti and Chatterjea, 2026).

Over the previous decades, Machine Learning (ML) has become an effective instrument for predicting employee attrition, providing valuable insight for organizations to implement data-based decisions to reduce attrition problems (Maji and Rajeev, 2026). ML, a subdivision of Artificial Intelligence (AI), has become a significant technology in overcoming real-world problems across various areas (Srivastava and Patnaik, 2025). At present, ML techniques are extensively implemented in human resource management for examining employee-related data and assisting organizational decision-

making. Losing team members is a massive headache for businesses. High turnover not only spikes the costs of hiring and training but also grinds daily operations to a halt (Talebi et al., 2025). Thus, businesses are prioritizing staff retention and improving employee morale (Abdelkarim et al., 2025). ML and deep learning (DL) based prediction approaches can help business recognize employees who are more likely to leave by examining aspects such as job satisfaction, salary, work environment, career growth, and employee performance (Ali et al., 2025). These predictive models help HR anticipate needs and develop retention plans, such as engagement activities, performance assistance, and career development (Sagar et al., 2025).

Although multiple ML and DL methods have been developed for employee attrition prediction, many existing techniques mainly focus on prediction performance and fail to efficiently capture complex workforce behavioral relationships and contextual dependencies from HR analytics data. Besides, many DL models operate as black boxes. This lack of transparency complicates strategic choices and workforce planning for HR leaders. This highlights the need for an efficient, interpretable hybrid DL system that accurately tracks employee retention, illuminates organizational reputation, and empowers workforce decision-making. Motivated by this, this study develops an Explainable Hybrid Deep Learning Framework for Organizational Reputation Assessment (XHDL-ORA) using HR analytics and workforce behavior intelligence. The major contributions of this manuscript are as follows:

- Employ a Correlation-based feature selection (CFS) with Mutual Information (MI) analysis to identify significant workforce attributes associated with employee attrition behavior.
- Develop a hybrid Transformer-Bidirectional Capsule Network with Attention Mechanism (TransBiCaps-AM) to capture complex workforce relationships and contextual behavioral dependencies from HR analytics data.
- Integrate a SHAP-based explainability model for improving model transparency and supporting reliable organizational decision-making.
- Conduct extensive experimental analysis and demonstrate the efficiency of the proposed model compared with conventional systems.

2. LITERATURE REVIEW

In this section, a brief review of employee attrition and HR analytics methods is presented along with the identified research gaps. (Tilokchandani et al., 2026) investigated the purpose of AI in healthcare

brand management. This study offers a data-based view on how healthcare can apply AI methodologies for navigating brand management complexity and enhancing patient experiences, and guaranteeing constant brand loyalty in a highly competitive industry. (Mortezapour et al., 2025) designed a DL approach on the basis of Bidirectional Temporal Convolutional Network (Bi-TCN) to foresee employee attrition. This research demonstrates general experiments on two publicly available datasets and comparing the methods' outcome against traditional ML, DL approaches, and advanced technique under distinct evaluation metrics. (Ma et al., 2025) developed an ML system to predict employee attrition. The presented architecture improves model interpretability, though preserving predictive influence, allowing HR practitioners to apply targeted retention systems. (Vijayan, 2025) presented a new data-driven technique to alleviate employee attrition utilizing data engineering and ML algorithms. The work highlights the significance of model interpretation, applying techniques, including SHAP values, to deliver valuable insights for organizations.

(Shathya et al., 2026) discussed various ML approach like KNearest Neighbor (KNN), Support Vector Machine (SVM), and Random Forest for generating an accurate prediction architecture for employee attrition. For developing a data-driven

retention strategy, the research emphasizes the significance of selecting a suitable predictive model to increase employee engagement and minimize turnover ratings. (Gazi et al., 2024) examined the function of ML, encompassing, more remarkably, Ada-boost classifier, Decision Tree, gradient-boosted classifier, and Random Forest (RF) for predicting employee turnover and its impacts for HR analytics and talent retention methodology. (Qin et al., 2026) introduced a novel tool for predicting employee turnover through big data and ML. This model combines XGBoost and LightGBM to improve stability and accuracy. To enhance approach interpretability, the SHAP strategy is employed to recognize key factors impacting turnover, like position, tenure, and salary. (Benabou et al., 2025) addressed a more effective ML algorithm to forecast employee attrition, thus supporting organizations with a consistent tool to expect turnover and employ proactive retention methodologies. To overcome the limitations of employee attrition by utilizing ML techniques to offer predictive insights that can increase retention strategies. To overcome the limitations of employee attrition by using ML approaches to offer predictive insights that can enhance retention algorithms. The summary of existing related works on employee attrition prediction and HR analytics approaches is presented in Table 1.

Table 1: Existing Studies on employee attrition prediction and HR analytics approaches.

References	Objective	Feature Engineering/ Selection	Models Used
(Tilokchandani et al., 2026)	To study the role of AI in enhancing healthcare reputation management and customer communication strategies	AI-driven customer and communication analysis	AI-based reputation management framework
(Mortezapour et al., 2025)	To design a DL model for forecasting employee turnover utilizing workforce behavioral and organizational data	Workforce behavioral feature analysis	Bi-TCN deep learning model
(Ma et al., 2025)	To enhance employee retention prediction by choosing the most important HR-related features and reducing data dimensionality	Wrapper, filter, and embedded feature selection	ML framework
(Vijayan et al.)	To design a data-driven machine learning framework for reducing employee attrition and supporting retention strategies	Data engineering and SHAP analysis	ML models
(Shathya et al., 2026)	To examine different ML models for precise employee attrition prediction and workforce retention analysis	SMOTE and categorical encoding	KNN, SVM, and Random Forest
(Gazi et al., 2024)	To foresee employee attrition and support HR analytics for refining talent retention and workforce management	HR-related employee attribute analysis	AdaBoost, Decision Tree, Gradient Boosting, Random Forest

(Qin et al., 2026)	To enhance employee turnover prediction accuracy and recognize important workforce factors influencing attrition behavior	SHAP-based feature importance analysis	XGBoost and LightGBM
(Benabou et al., 2025)	To develop ML-based predictive models for employee turnover analysis and proactive employee retention planning	Workforce predictive attribute analysis	Logistic Regression and ML classifiers

2.1. Research Gap Analysis

Although several studies have been carried out for employee attrition prediction using ML and DL techniques, some limitations are still found in the existing works. (Mortezapour Shiri et al., 2025) proposed a Bi-TCN-based DL approach for employee attrition prediction. Still, the study mainly concentrated on prediction accuracy and did not consider organizational-level workforce behavior analysis. (Ma et al., 2025) presented an ML-based feature selection technique for employee retention prediction; however, the work did not comprise hybrid DL methods and explainability analysis for better HR decision support. (Vijayan, 2025) developed a data-driven framework with SHAP analysis for employee attrition prediction, but the approach mainly relied on traditional ML techniques without capturing deeper workforce relationships. Similarly, (Qin et al. 2026) applied XGBoost and LightGBM models with SHAP interpretability for turnover prediction, yet the model did not use advanced transformer and capsule network-based learning for employee behavior analysis. Thus, there is a need for an explainable hybrid DL system that can efficiently examine workforce behavior patterns and enhance employee attrition classification for better organizational decision-making, and the

proposed XHDL-ORA model addresses this need.

3. PROPOSED METHODOLOGY

This section presents the overall workflow of the proposed model, including model architecture, dataset description, classification, optimization, and explainability components.

3.1. Overall Framework Architecture

Fig. 1 shows the overall workflow of the proposed XHDL-ORA model. Primarily, the HR analytics dataset is preprocessed using missing value handling, outlier detection, encoding, normalization, and class balancing techniques to make the dataset fit for analysis. Then, correlation-based feature selection with Mutual Information analysis is applied to select important employee-related features. The selected features are passed to the proposed Hybrid Transformer-Bidirectional Capsule Network with Attention Mechanism (TransBiCaps-AM) for employee attrition classification. The Adam optimizer is used during training to improve model performance. Finally, SHAP-based explainability is integrated to identify the important factors that influences employee attrition and support better organizational decision-making.

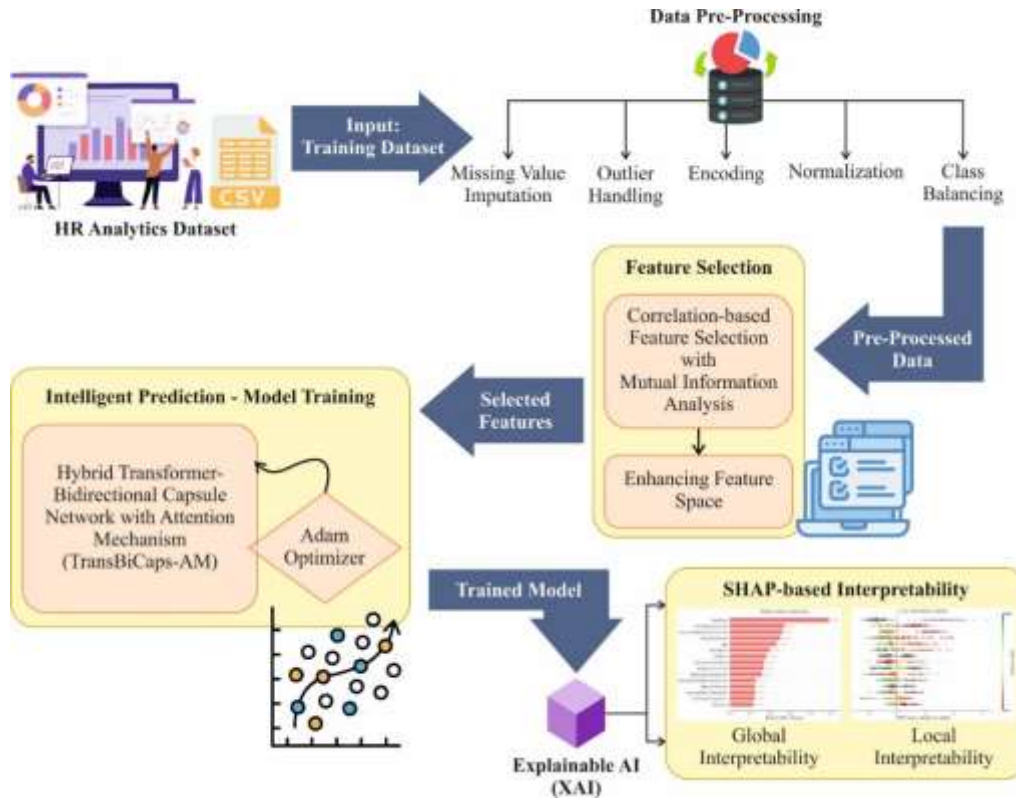


Figure 1. Overall workflow of the proposed XHDL-ORA model.

4. DATA PREPROCESSING

This section explains the preprocessing steps applied to improve data quality and model learning performance.

Missing Value Handling using Median Imputation: Median imputation is applied to handle the missing values in the HR analytics employee attrition dataset (Joel et al., 2022). In this process, missing values are replaced through the median value of every feature. This method is less affected by extreme values and helps enhance data stability during analysis.

$$X_{hk} = \text{"Median"}(X_k) \tag{1}$$

Here, X_{hk} indicates the imputed value, and X_k represents all non-missing values in the k^{th} feature.

Outlier Detection using IQR: The Interquartile Range (IQR) approach is employed for outlier detection to recognize data points that differ notably from most of the observations in the dataset (Mamuriyah, 2025). The IQR indicates the range of the middle 50% of the data and is computed as:

$$IQR = Q_3 - Q_1 \tag{2}$$

Where the 1st and 3rd quartiles are denoted as Q_1 and Q_3 , respectively, data points located far from the normal distribution range are treated as outliers. They can be managed to enhance the data quality and model performance.

One-hot Encoding: The one-hot encoding is employed to transform categorical data into a numerical representation suitable for a DL approach (Yang et al., 2025). In this process, every category is denoted as a binary vector including values of 0 and 1. This process supports enhancing the model performance and preventing problems associated with direct label assignment to categorical features. The logistic regression function after applying encoding can be represented as:

$$h(x) = \frac{1}{1 + e^{-w^T x}} \tag{3}$$

Here, the input vector can be indicated as x , and the weight vector can be represented as w .

Min-max normalization: All the features are standardized through the min-max method, converting feature values to the interval of [0, 1]. The mathematical expression for the normalization is defined as (Hu et al., 2024):

$$X_{norm} = \frac{(X - X_{min})}{(X_{max} - X_{min})} \tag{4}$$

Here X_{min} represents the minimum feature value, X_{max} indicates maximum feature value, and X signifies original feature values.

Data Balancing: The SMOTE-driven adaptive class balancing technique mainly relies on its synthetic sample generation process (Husain et al., 2025). This process reflects the ability of the

algorithm to overcome class disproportion through generating fresh synthetic samples of the minority employee attrition category. For every minority employee instance, N synthetic samples are produced, where N depends on the specified oversampling rate. The step for producing every synthetic employee sample is defined as follows:

$$x_{\text{new}} = x_i + \lambda(x_j - x_i) \quad (5)$$

Here x_i denotes the selected minority employee sample, x_j represents one of the nearest neighboring minority samples, and λ indicates the random number in the interval of $[0, 1]$. Random neighboring employee instances are chosen from the minority attrition category to produce the new synthetic workforce samples.

The SMOTE-driven balancing approach enhances the distribution of employee attrition categories and improves the learning efficiency of the proposed TransBiCaps-AM model in workforce behavior classification and organizational reputation forecasting.

Feature Selection

This section presents the feature selection method employed for choosing the most relevant workforce attributes related to employee attrition.

Correlation-Based Feature Selection

The aim of feature selection (FS) is to choose the most significant attributes while reducing computational time and enhancing predictive outcomes (Alomari et al., 2023). In this paper, a correlation-driven FS method is proposed to decrease data redundancy, minimize the size of employee behavioral features, and choose the most relevant workforce features related to employee turnover and organizational reputation analysis. For the HR analytics data, the correlations among each employee's features and the target class are computed employing Eq. (6).

$$\text{Corr}_{(u,v)} = \frac{(\sum(u_i - \bar{u})(v_i - \bar{v}))}{\sqrt{(\sum(u_i - \bar{u})^2)(\sum(v_i - \bar{v})^2)}} \quad (6)$$

Here, $\text{Corr}_{(u,v)}$ implies the correlation between u_i and v , representing the employee feature and the target variable. \bar{u} and \bar{v} imply the mean values of u and v , respectively.

The correlation-driven FS method classifies major employee factors impacting turnover and eliminates unnecessary information, facilitating the TransBiCaps-AM method to become quicker, steadier, and more precise.

4.1. Mutual Information Analysis

Mutual information (MI) is an assessment of the interdependency among dual features and mostly signifies the dependency among workforce factors

and the target attrition class (Huang et al., 2024). In employee turnover analysis, MI helps in classifying features that contribute considerably to employee retention classification and organizational behavioral analysis.

Assume the joint probability function of dual random features U and V is $p(u,v)$, and their marginal probability functions are $p(u)$ and $p(v)$. Subsequently, the $MI(U, V)$ is expressed as below:

$$MI(U, V) = - \int \int p(u,v) \log \left(\frac{p(u,v)}{p(u)p(v)} \right) du dv \quad (7)$$

As mentioned above, when features are entirely independent, the MI value becomes zero. The higher the MI among features, the more robust the dependency among employee features and the target class. Thus, workforce attributes with greater MI values are chosen to enhance the prediction ability of the proposed structure, while attributes with lesser MI values are removed to minimize feature redundancy.

In the effective application, the probability density approximation process is utilized to estimate the probability function and simplify the MI calculation as follows:

$$MI(U, V) = -1/N \sum_{(i=1)^N} [\log \left(\frac{f_{(u,v)}(u_i, v_i)}{f(u_i)f(v_i)} \right)] \quad (8)$$

In this Eq. (8), $f(u_i)$ and $f(v_i)$ imply the probability density approximations of features U and V , correspondingly, $f_{(u,v)}(u_i, v_i)$ represents the joint distribution, and N denotes the count of employee reports in the data. Based on the above equation, MI requires approximating probability functions, so a kernel density estimate is applied for approximating the marginal and joint functions without accepting any features. Integrating correlation-driven FS with MI support to choose significant employee attributes while eliminating redundant ones, enhancing the TransBiCaps-AM method for better attrition classification and HR analytics data.

4.2. Proposed Classification Model

This section discusses the proposed Hybrid Transformer-Bidirectional Capsule Network with Attention Mechanism (TransBiCaps-AM) for employee attrition categorization. The transformer structure utilized in this research is based on the standard encoding architecture with various fluctuations optimized for organizational reputation analysis and employee attraction classification (Haq et al., 2023). The proposed HR analytics model comprises an input layer, transformer encoder with attention mechanism, bidirectional feature learning

layer, capsule network layer, and final classification layer for employee attrition classification.

The encoding processes the input attributes across several layers of self-attention (SA) and feed-forward networks. The calculation of AM is represented as follows:

$$Q=XW_Q, \quad K=XW_K, \quad V=XW_V \quad (9)$$

In this equation, X denotes the group of input employee attributes, and W_Q , W_K , and W_V imply trainable weight matrices. The SA outcome is evaluated as:

$$\text{Attention}(Q,K,V)=\text{softmax}((QK^T)/\sqrt{d_k})V \quad (10)$$

Here, d_k signifies the key vector dimension. The multi-head attention mechanism (MHA) uses numerous attention heads functioning similarly:

$$\text{MultiHead}(X)=\text{Concat}(\text{head}_1,\dots,\text{head}_h)W_O \quad (11)$$

where every head is calculated as:

$$\text{head}_i=\text{Attention}(QW_Q^i,KW_K^i,VW_V^i) \quad (12)$$

The MHA allows the method to mutually attend to data from diverse feature representation subspaces at various locations. Every attention head acquires several features of employee behavior patterns, where certain heads concentrate on employee performance and satisfaction, whereas others focus on retention-associated dependency. This module processes Q , K , and V , which implies the queries, keys, and values vectors across parallel attention heads before integrating their outcomes.

A different standard transformer structure that

processes each attribute uniformly across SA, the proposed hybrid method first makes employee-specific contextual representations throughout transformer AMs and therefore collects these representations over a bidirectional capsule learning model. It allows the method to take both global organizational interactions and local workforce behavioral patterns, which is a significant distinction from conventional transformers that normally use a unique level of attention through each attribute.

For capturing the contextual dependency particular to employee turnover patterns, the method integrates a feature embedded module. The employee feature embedding E_{fis} is represented as:

$$E_f=\text{Embedding}(X)+P_f \quad (13)$$

Let $\text{Embedding}(X)$ imply the embedded employee feature, and P_f represent the positional feature encoded. The transformer outcome for all employee features is expressed as:

$$H_i=\text{MultiHead}(X_i+P_f) \quad (14)$$

The CapsNet is used for capturing the hierarchical relations between employee features. The fundamental concept of CapsNet is to encode the relations among multiple workforce behavior features, including salary level, employee satisfaction, work-life balance, and performance rating. A different traditional neural network (NN) that considers features self-sufficiently, the CapsNet learns the relations between employee features and maintains significant structural dependency in classification.

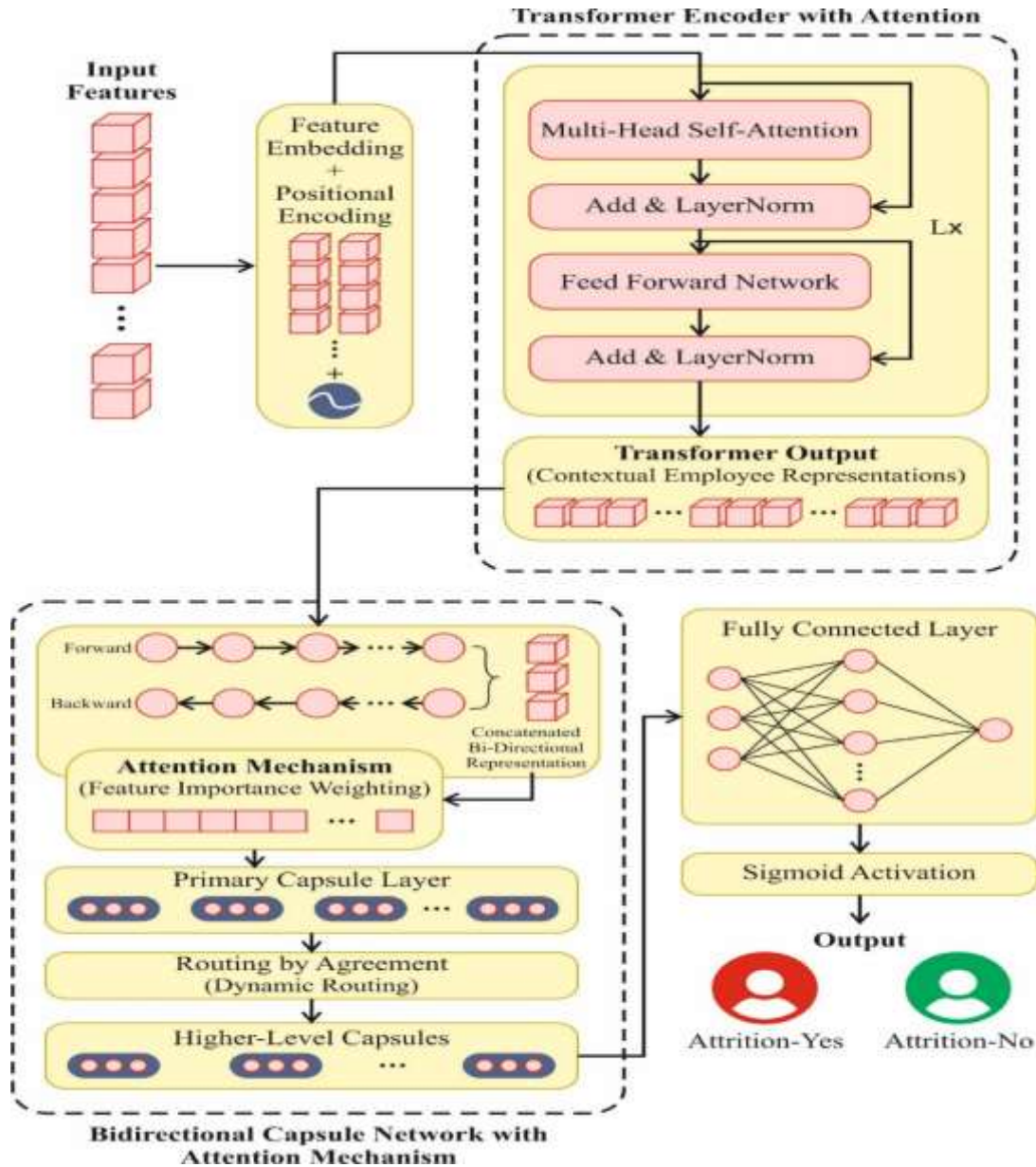


Figure 2. Framework proposed classification model.

Mostly, a CapsNet is a kind of NN that achieves hierarchical feature learning. In employee turnover prediction, workforce features are separated into numerous feature groups, and hierarchical relations are expanded between these feature representations. The application of CapsNet in this research is separated into 3 main regions: input, capsule feature, and outcome classifier layers. The first capsule layer outcome is represented by:

$$V_{(i,j,k)} = \text{Squash} \left(\sum_{l=1}^{L^k} u_{(i,j,l)} W_{(i,j,k,l)} \right) \quad (15)$$

Now, $V_{(i,j,k)}$ denotes the outcome vector of the k^{th} capsule representation, $u_{(i,j,l)}$ implies the vector of input features, and Squash signifies the non-linear activation.

$$S_j = \sum_i c_{ij} u_{(j|i)} \quad (16)$$

In this Eq., S_j denotes the outcomes of the high-level capsule, $u_{(j|i)}$ implies the forecasted feature representation, and c_{ij} signifies the coupled coefficient among minimal and maximal-level capsules.

The AM is combined with the CapsNet to allocate adaptive significance weights to workforce attributes. This module enhances the capability of the method to classify crucial employee retention and organizational behavioral patterns related to turnover prediction.

The resulting classifier output is created through a fully connected (FC) layer with an activation function of sigmoid:

$$y \approx \text{Sigmoid}(W_h H + b) \quad (17)$$

Let y^{\wedge} imply the forecasted employee turnover

categorizes, W_h denotes trainable weights, and H signifies the fused feature representation attained from the capsule and transformer layers.

The proposed TransBiCaps-AM systems allow the method to simultaneously take employee-specific behavioral patterns and intricate organizational feature relations, offering an efficient method for intelligent employee turnover prediction and organizational reputation analysis utilizing HR analytics features. Fig. 2 showcases the proposed XHDL-ORA model for employee attrition classification using a Transformer and a Bidirectional Capsule Network with an Attention Mechanism.

4.3. Model Optimization

This section describes the optimization technique utilized for improving the training efficiency and convergence performance of the proposed approach. The Adam optimization algorithm is applied to adjust the model parameters, a commonly accepted optimization technique in deep learning due to its efficiency in handling adaptive learning rates and momentum capabilities (Ishtiaq et al., 2025). The update rule for parameters θ at iteration t is expressed as:

$$m_t = \beta_1 m_{(t-1)} + (1 - \beta_1) \nabla_{\theta} L_t \tag{18}$$

$$v_t = \beta_2 v_{(t-1)} + (1 - \beta_2) (\nabla_{\theta} L_t)^2 \tag{19}$$

$$\hat{m}_t = m_t / (1 - \beta_1^t) \tag{20}$$

$$\hat{v}_t = v_t / (1 - \beta_2^t) \tag{21}$$

$$\theta_t = \theta_{(t-1)} - \eta (\hat{m}_t / (\sqrt{\hat{v}_t} + \epsilon)) \tag{22}$$

Where η denotes the learning rate, β_1 , β_2 , and ϵ signify numerical stability parameters. The Adam optimization technique enhances the convergence efficiency of the proposed TransBiCaps-AM model through dynamically modifying the learning rates in the training process. This optimization strategy also improves the predictive scalability and classification accuracy of the proposed system for employee attrition analysis, forecasting, and organizational reputation evaluation based on HR analytics information.

4.4. SHAP-based Interpretability

This section defines the SHAP-based explainability analysis leveraged to interpret the contribution of workforce features toward prediction results. DL models are considered black boxes owing to their difficulty and complexity in understanding how specific characteristics impact predictions

(Abdullah and Mustafa, 2025). Therefore, their interpretability makes it complex to learn whether the feature has a negative or positive impact on the model outcome. Shapley Additive exPlanations (SHAP) is one of the better techniques for improving the interpretability of ML and DL systems in HR analytics and offers an explicit approach for estimating feature contributions. SHAP values evaluate the effect of all features on the prediction compared with other attributes and thereby enable the feature significance ranking in the employee turnover prediction method. It is computed utilizing built-in functions to offer an equivalent and objective assessment of workforce-relevant features in the HR databases. The SHAP value computation is represented as Eq. (23).

$$f(X^i) = J_0 + \sum_{i=1}^K \sum_{j \in S} \phi_j(X_i^i) \tag{23}$$

Here, f implies the interpretation method, X^i signifies the input vector of employee features, K denotes the higher dimensionality of features, and J indicates the feature attributions.

Experimental Results and Discussions

This section offers the experimental evaluation of the proposed approach using HR Analytics dataset, available at <https://www.kaggle.com/datasets/bhanupratapbiswas/hr-analytics-case-study/data>. HR analytics otherwise known as people analytics, is an evidence-based approach for handling human resources. It comprises data related to employees, namely recruitment, performance, engagement, and retention to uncover details and make informed decisions. This case study analyzes the application of HR analytics in a hypothetical organization and demonstrates its benefits in optimizing workforce management.

Organization Description: Consider “TechSolutions Inc” a growing, mid-scale tech company. This company focusses on software development and has a varyng crews across various units, including marketing, engineering, sales, and customer support. Number of features in the dataset is 35, out of which 24 features were chosen after the feature selection process. Table 2 shows that the dataset contains two classes, “Attrition Yes” and “Attrition No.” Before augmentation, the dataset had 1470 samples, which increased to 2230 samples after augmentation.

Table 2. Dataset details.

Classes	Before Augmentation	After Augmentation
Attrition Yes	237	997
Attrition No	1233	1233
Number of Samples	1470	2230

Fig. 3 displays the distribution of employee age, monthly income, and years at company. Age is typically focused between 25 to 45 years, with a near-normal distribution pattern. Monthly income is right-skewed, signifying most employees earn in the lower to mid salary range. Years at company is also highly skewed, showing most employees have fewer years of experience in the organization.

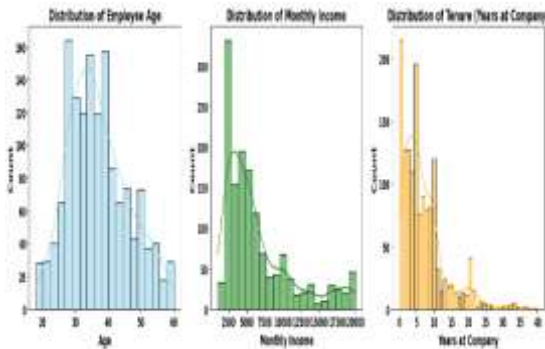


Figure 3. Distribution of key variables

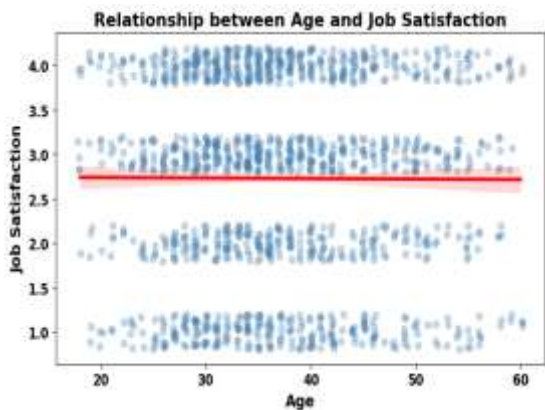


Figure 4. Figure Relationship between Age and Job Satisfaction

Fig. 4 shows the relationship among age and job satisfaction. The plot indicates that job satisfaction is fairly spread across all age groups with no strong linear relationship. The trend line is almost flat, suggesting age has little impact on job satisfaction levels.

Fig. 5 shows the confusion matrix for the XHDL-ORA framework, indicating that the model performs better in classifying attrition of employees. The model correctly classified 173 samples as attrition yes and 245 instances as attrition no. Therefore, the XHDL-ORA model proves higher accuracy and reliability in identifying employee attrition.

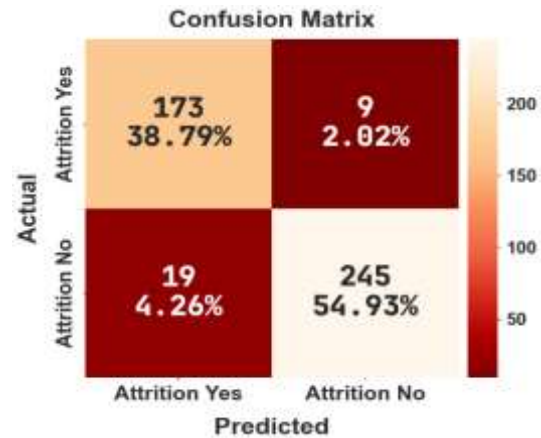


Figure 5. Confusion matrix of the XHDL-ORA method

Table 3 present the attrition classifier outcome of the XHDL-ORA method with 80:20 of train-test split. The outcome implied that the XHDL-ORA approach accurately classified the samples. The XHDL-ORA system obtained average acc $\llbracket ur \rrbracket_y$ of 93.11% in TRAIIP and 93.72% in TESTP, while the $\llbracket F1 \rrbracket_y$ score of 93.05% for TRAIIP and 93.55% for TESTP, showing strong efficiency and consistent performance in predicting both attritions yes and no cases.

Table 3. Attrition classifier result of the XHDL-ORA approach with 80:20.

Metrics	TRAIIP (80%)	TESTP (20%)
Accuracy	93.11	93.72
Precision	93.07	93.28
Recall	93.03	93.93
F1-Score	93.05	93.55
AUC Score	93.03	93.93
MCC	86.10	87.21

Table 4 present the attrition classifier outcome of the XHDL-ORA method with 70:30 of train-test split. The outcome implied that the XHDL-ORA approach accurately classified the samples. The XHDL-ORA system obtained average acc $\llbracket ur \rrbracket_y$ of 91.29% in TRAIIP and 91.78% in TESTP, while the $\llbracket F1 \rrbracket_y$ score of 91.18% for TRAIIP and 91.46% for TESTP, showing strong efficiency and consistent performance in predicting both attritions yes and no cases.

Table 4. Attrition classifier result of the XHDL-ORA approach with 70:30.

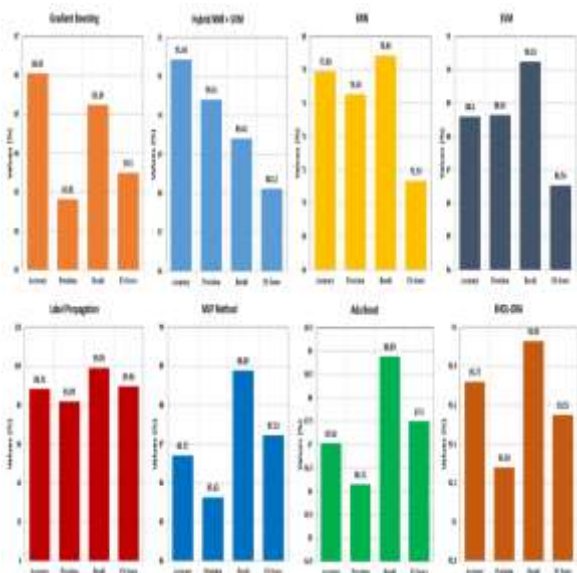
Metrics	TRAIIP (70%)	TESTP (30%)
Accuracy	91.29	91.78
Precision	91.65	91.64
Recall	90.97	91.29
F1-Score	91.18	91.46
AUC Score	90.97	91.29
MCC	82.62	82.93

Table 5 and Fig. 6 show the comparison study of the XHDL-ORA approach with other methodologies

using metrics and Time Complexity (TC) (Al-Ali et al., 2026; Govindarajan et al., 2025; Arqawi et al., 2022). The proposed XHDL-ORA method attains the overall better performance with acc $\llbracket ur \rrbracket$ of 93.72%, $\llbracket n \rrbracket$ of 93.28%, $\llbracket l \rrbracket$ of 93.93%, $\llbracket F1 \rrbracket$ score of 93.55%, and minimal TC value of 3.57sec, while Hybrid NNR + SVM, SVM, Label Propagation, and Ada Boost approaches achieve lesser outcomes with acc $\llbracket ur \rrbracket$ of 91.44%, 88.60%, 88.33%, and 87.03%. At the same time, the MLP Method, Gradient Boosting, and KNN display competitive performance, they are still inferior to the proposed XHDL-ORA method in overall effectiveness.

Table 5. Comparative result of the XHDL-ORA method with existing approaches.

Approach	Accur _y	Preci _n	Recal _l	F1 _{score}	Time Complexity (sec)
Gradient Boosting	86.05	82.82	85.24	83.50	5.81
Hybrid NNR + SVM	91.44	90.41	89.41	88.12	6.93
KNN	77.89	76.54	78.85	71.33	8.98
SVM	88.60	88.65	90.24	86.54	4.70
Label Propagation	88.33	81.89	99.05	89.66	7.82
MLP Method	86.71	85.63	88.89	87.23	6.10
Ada Boost	87.03	86.15	88.89	87.50	5.03
XHDL-ORA	93.72	93.28	93.93	93.55	3.57



An ablation study calculates a model by removing or modifying components to understand their impact on overall performance. Table 6 exhibit the ablation study of the XHDL-ORA approach. The outcome shown that the Without features selection, Correlation only, Mutual Information only, Transformer only, Bidirectional Capsule Network only, Transformer + BiCaps (without attention), and

Transformer + BiCaps + Attention (without Adam optimizer) approaches have reached minimal performance, whereas the XHDL-ORA (Correlation + MI + Transformer + BiCaps + Attention + Adam) method has achieved maximal performance with acc $\llbracket ur \rrbracket$ of 93.72%, $\llbracket n \rrbracket$ of 93.28%, $\llbracket l \rrbracket$ of 93.93%, $\llbracket F1 \rrbracket$ score of 93.55%. Thus, the XHDL-ORA system demonstrated superior and reliable performance through the effective integration of hybrid feature selection and advanced deep learning techniques.

Fig. 7 presents a SHAP analysis showing the most influential features affecting the model’s predictions. The left panel displays global feature importance, where OverTime, StockOptionLevel and EnvironmentSatisfaction have the highest impact on employee attrition prediction. The right panel illustrates how individual feature values influence the model output, with red points signifying maximum feature values and green points denoting minimum values. Positive SHAP values represent a maximum chance of employee attrition, while negative values specify minimum attrition possibility. Employees with overtime work and low job satisfaction showed higher attrition tendency.

Table 6. Ablation study of the XHDL-ORA method with various measures.

Approach	Accu _y	Preci _n	Recal _l	F1 _{score}
Without features selection	89.06	89.00	88.45	88.23
Correlation only	89.77	89.32	90.20	89.71
Mutual Information only	90.30	89.98	90.76	90.22
Transformer only	91.08	90.59	91.30	90.99
Bidirectional Capsule Network only	91.83	91.18	92.10	91.50
Transformer + BiCaps (without attention)	92.47	91.75	92.64	92.16
Transformer + BiCaps + Attention (without Adam optimizer)	93.22	92.55	93.20	92.76
XHDL-ORA (Correlation + MI + Transformer + BiCaps + Attention + Adam)	93.72	93.28	93.93	93.55

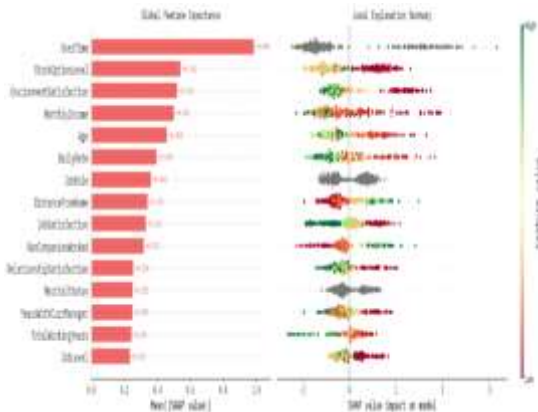


Figure 7. SHAP value for the global and local explanations.

The results showed that the proposed XHDL-ORA approach performed better than existing ML and DL techniques for employee attrition classification. The model obtained 93.72% accuracy and 93.55% F1-score, exhibiting its ability to learn workforce behavior patterns efficiently. The ablation results also indicated that integrating Transformer, Bidirectional Capsule Network, and Attention Mechanism improved the overall classification performance. Furthermore, adaptive class balancing and feature selection stabilized the model and eliminated redundant employee metrics. The SHAP-based interpretability analysis further improved model transparency by recognizing the important employee-related factors influencing attrition behavior. Ultimately, the data demonstrates that the implemented model successfully facilitates the evaluation of workforce attrition and corporate reputation.

5. CONCLUSION AND FUTURE WORK

Data Availability Statement: The data that support the findings of this study are openly available in Kaggle repository at <https://www.kaggle.com/datasets/bhanupratapbiswas/hr-analytics-case-study/data>.

REFERENCES

- ABDELKARIM, K., ALTAYEB, B. and ABDELRAHIM, M., 2025. Predictive Model for Reducing Employee Turnover Using Machine Learning Techniques. *Journal of HRM*, 28(2).
- Abdullah, M.H. and Mustafa, Y.T., 2025. Machine learning and SHAP-based analysis of deforestation and forest degradation dynamics along the Iraq–Turkey border. *Earth*, 6(2), p.49.
- Al-Ali, M., Alwateer, M., Alsaedi, S.A., Balaha, H.M., Badawy, M. and A. Elhosseini, M., 2026. Integrating machine learning and explainable AI for employee attrition prediction in HR analytics: M. AL-Ali et al. *Scientific Reports*, 16(1), p.6344.
- Ali, M.A., Upendra, K., Tejasri, B., Dharani, T. and Narender, B., 2025, October. AI Driven Approaches for Predicting and Preventing Employee Attrition Through Machine Learning Models. In 2025 International Conference on Sustainable Communication Networks and Application (ICSCN) (pp. 1028-1033). IEEE.
- Alomari, E.S., Nuijaa, R.R., Alyasseri, Z.A.A., Mohammed, H.J., Sani, N.S., Esa, M.I. and Musawi, B.A., 2023. Malware detection using deep learning and correlation-based feature selection. *Symmetry*, 15(1), p.123.

This paper has introduced the XHDL-ORA approach for customer attribution classification using the HR Analytics Case Study dataset. A CFS with MI analysis was used to select the most informative and non-redundant workforce attributes influencing employee attrition behavior and institutional reputation. For attrition classification, a Hybrid TransBiCaps-AM was developed for effectively learning complex workforce relationships, contextual feature dependencies, and hierarchical employee behavioral representations from heterogeneous HR analytics data. Besides, the Adam optimizer was incorporated to enhance convergence efficiency during the training process. Additionally, SHAP-based XAI analysis was incorporated to improve model transparency by identifying the contribution and influence of workforce features toward employee attrition classification. Experimental analysis showed that the proposed method significantly enhanced classification performance compared to existing models. Therefore, the proposed model is found to be effective in intelligent organizational reputation assessment by allowing proactive workforce analytics, employee retention intelligence, and AI-driven institutional decision support systems. Although the presented XHDL-ORA model achieved good performance, the study used only one public HR dataset, which may impact its performance in different organizational settings. In future work, larger real-time HR datasets and additional employee behavioral factors can be integrated to further improve the prediction performance and practical usability of the model.

- Arqawi, S.M., Rumman, M.A., Zitawi, E.A., Abunasser, B. and Abu-Naser, S., 2022. Predicting employee attrition and performance using deep learning. *Journal of Theoretical and Applied Information Technology*, 100(21), pp.6526-6536.
- Bellora-Bienengräber, L., Dalton, D.W. and McPhee, G.P., 2026. Enabling Budgeting and Employee Turnover Intentions: The Role of Perceived Organizational Support and Error Climates. *Journal of Management Accounting Research*, pp.1-32.
- Benabou, A. and Touhami, F., 2026. Optimising HRM practices in call centres: Predicting and explaining employee turnover intention using classification and regression trees. *International Journal of Organizational Analysis*, 34(3), pp.870-891.
- Benabou, A., Touhami, F. and Abdelouahed Sabri, M., 2025. Predicting employee turnover using machine learning techniques. *Acta Informatica Pragensia*.
- Chakrabarti, G. and Chatterjea, T., 2026. Emotional Disengagement and Employee Turnover in Workplaces: An Applied Study. Taylor & Francis.
- Cui, Z., Li, C. and Chen, M., 2026. An Empirical Study of the Dynamics of Collective Employee Turnover and Its Implications for Dismissal Policy – Evidence From a Distribution Center. *Human Resource Management Journal*.
- Gazi, M.S., Nasiruddin, M., Dutta, S., Sikder, R., Huda, C.B. and Islam, M.Z., 2024. Employee attrition prediction in the USA: A machine learning approach for HR analytics and talent retention strategies. *Journal of Business and Management Studies*, 6(3), pp.47-59.
- Govindarajan, R., Kumar, N.K. and Reddy, S., 2025. Predicting employee attrition: a comparative analysis of machine learning models using the IBM human resource analytics dataset. *Procedia Computer Science*, 258, pp.4084-4093.
- Haq, M.U., Sethi, M.A.J. and Rehman, A.U., 2023. Capsule network with its limitation, modification, and applications – A survey. *Machine Learning and Knowledge Extraction*, 5(3), pp.891-921.
<https://www.kaggle.com/datasets/bhanupratapbiswas/hr-analytics-case-study/data>
- Hu, H., Dong, W., Yu, J., Guan, S. and Zhu, X., 2024. Utilizing Attention-Enhanced Deep Neural Networks for Large-Scale Preliminary Diabetes Screening in Population Health Data. *Electronics*, 13(21), p.4177.
- Huang, L., Zhou, X., Shi, L. and Gong, L., 2024. Time series feature selection method based on mutual information. *Applied Sciences*, 14(5), p.1960.
- Husain, G., Nasef, D., Jose, R., Mayer, J., Bekbolatova, M., Devine, T. and Toma, M., 2025. SMOTE vs. SMOTEENN: A study on the performance of resampling algorithms for addressing class imbalance in regression models. *Algorithms*, 18(1), p.37.
- Iorio, A. and Morris, S., 2026. Tied in or Tied Up? The Contrasting Effects of Social Ties on Employee Turnover. *Academy of Management Journal*, (ja), pp.amj-2024.
- Ishtiaq, W., Zannat, A., Parvez, A.S., Hossain, M.A., Kanchan, M.H. and Tarek, M.M., 2025. CST-AFNet: A dual attention-based deep learning framework for intrusion detection in IoT networks. *Array*, p.100501.
- Joel, L.O., Doorsamy, W. and Paul, B.S., 2022. A review of missing data handling techniques for machine learning. *International Journal of Innovative Technology and Interdisciplinary Sciences*, 5(3), pp.971-1005.
- Ma, D., Shu, M. and Zhang, H., 2025. Feature selection optimization for employee retention prediction: A machine learning approach for human resource management.
- Maji, S. and Rajeev, D., 2026. Interpersonal microaggression, organizational commitment, turnover intention and mental health: a study on LGBTQA+ employees in India. *Equality, Diversity and Inclusion: An International Journal*, 45(4), pp.751-777.
- Mamuriyah, N., 2025. Implementation Mean Imputation and Outlier Detection for Loan Prediction Using the Random Forest Algorithm. *JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer)*, 10(4), pp.937-944.
- Mortezapour Shiri, F., Yamaguchi, S. and Ahmadon, M.A.B., 2025. A deep learning model based on bidirectional temporal convolutional network (Bi-TCN) for predicting employee attrition. *Applied Sciences*, 15(6), p.2984.
- Priya, T., Pal, Y., Singh, J., Bhadre, A.A., Raju, P.V.M. and Uttamkumar, Y.V., 2026. Employee Retention Strategies Enhanced Through Advanced Machine Learning Models. In *Innovative Computing and Communications: Proceedings of ICICC 2025, Volume 8* (pp. 73-79). Singapore: Springer Nature Singapore.

- Qin, R., Qi, X., Yuan, Y. and Alatas, B., 2026. Employee Turnover Prediction Research of Human Resource Management on Machine Learning Algorithms and Big Data Analysis. *Journal of Organizational and End User Computing (JOEUC)*, 38(1), pp.1-27.
- Sagar, S., Dhumal, R., Kumbhar, V. and Singh, T.P., 2025, September. Comparative Analysis of Deep Learning Models for Employee Attrition Prediction with Explainable AI. In *2025 2nd International Conference on Integration of Computational Intelligent System (ICICIS)* (pp. 1-5). IEEE.
- Shathya, B., Deepika, R. and Keerthana, J., 2026. Prediction of employee attrition using Machine Learning algorithms. *Excel International Journal of Technology, Engineering and Management*, 13(1), pp.1-7.
- Srivastava, A.K. and Patnaik, D., 2025. Data-Driven Insights and Predictive Modelling for Employee Attrition: A Comprehensive Analysis Using Statistical and Machine Learning Techniques. *Journal of Computational Analysis & Applications*, 34(1).
- Talebi, H., Khatibi Bardsiri, A. and Bardsiri, V.K., 2025. Machine learning approaches for predicting employee turnover: A systematic review. *Engineering Reports*, 7(8), p.e70298.
- Tilokchandani, T., Bhatnagar, D. and Kumar, S., 2026. AI in Healthcare Brand Management: Building Trust and Reputation. In *Navigating AI in Business: Strategies and Insights Across Disciplines* (pp. 309-330). Cham: Springer Nature Switzerland.
- Vijayan, N.E., 2025. Mitigating attrition: Data-driven approach using machine learning and data engineering. arXiv preprint arXiv:2502.17865.
- Yang, T., Yu, H., Lu, D., Bai, S., Li, Y., Fan, W. and Liu, K., 2025. Detection and classification of abnormal power load data by combining one-hot encoding and GAN-transformer. *Energies*, 18(5), p.1062.