An integrated model to evaluate the transparency in predicting employee churn using explainable artificial intelligence

Meenu Chaudhary [a], Loveleen Gaur [b], Amlan Chakrabarti [c], Gurmeet Singh [d], Paul Jones [e], Sascha Kraus [f,g,*]

[a] Human Resource Analytics from Amity University Noida, India
[b] Amity International Business School, Amity University, Noida, India
[c] A.K. Choudhury School of Information Technology at the University of Calcutta, India
[d] School of Business and Management, The University of the South Pacific, Fiji
[e] School of Management, Swansea University, Swansea, UK
[f] Free University of Bozen-Bolzano, Faculty of Economics & Management, Piazza Università 1, 39100 Bolzano, Italy
[g] University of Johannesburg, Department of Business Management, Johannesburg, South Africa

ARTICLE INFO

ABSTRACT

Recent studies focus on machine learning (ML) algorithms for predicting employee churn (ECn) to save probable economic loss, technology leakage, and customer and knowledge transference. However, can human resource professionals rely on algorithms for prediction? Can they decide when the process of prediction is not known? Due to the lack of interpretability, ML models' exclusive nature and growing intricacy make it challenging for field experts to comprehend these multifaceted black boxes. To address the concern of interpretability, trust and transparency of black-box predictions, this study explores the application of explainable artificial intelligence (XAI) in identifying the factors that escalate the ECn, analysing the negative impact on productivity, employee morale and financial stability. We propose a predictive model that compares the best two top-performing algorithms based on the performance metrics. Thereafter, we suggest applying an explainable artificial intelligence based on Shapley values, i.e., the Shapley Additive exPlanations approach (SHAP), to identify and compare the feature importance of top-performing algorithms logistic regression and random forest analysis on our dataset. The interpretability of the predictive outcome unboxes the predictions, enhancing trust and facilitating retention strategies.

Introduction

The biggest challenge in today's dynamic and competitive business environment is the effective retention of human capital. Due to globalisation, organisations need assistance to retain employees to overcome two key challenges: the loss of skilled and knowledgeable employees and synchronisation with trends in the skilled market. The transition to employee churn has impacted businesses in many sectors and economies. It gathered speed throughout the global epidemic and witnessed a significant increase during the phenomenon known as the 'Great Resignation', which began in the United States in 2021.

High costs are incurred from inducting and training a new employee; organisational stability and productivity are key challenges that

organisations encounter (Siong et al., 2006). According to the EY 2021 Work Reimagined Employee Survey, around 54 % of surveyed employees would consider leaving their current job post-COVID-21. Society for Human Resource Management (SHRM) suggests that employee replacement costs 90 % to 200 % (Alsheref et al., 2022). Deloitte reported that employee costs are up to 1.5 to 2 times the yearly salary (Punnoose & Ajit, 2016).

Employee churn can be voluntary and involuntary (Gelbard et al., 2018). The voluntary churn cost can range by up to five times the average employee's annual salary depending on the criticality of the job and problems in adequately replacing an employee (Liborius & Kiewitz, 2022). Voluntary churn occurs when employees leave for personal reasons, while involuntary churn refers to the wilful release of

employees by the organisation (Chalutz Ben-Gal, 2019). Organisations typically focus on voluntary churn, which arises when an employee resigns seeking enhanced financial compensation, perks, or an improved working environment (Anh et al., 2020). By contrast, unfavourable reasons for departure include disagreements with supervisors, lack of promotion opportunities, employees themselves, organisational culture, or a repetitive, unchallenging job. This necessitates for HR professionals, managers, and leaders to understand the key factors understanding and contributing to employee churn.

Academicians and practitioners have endeavoured to predict employee churn. This has led to including data-driven predictions using machine learning (ML) algorithms (Fallucchi et al., 2020). Various data mining methods were used to build ML models. Existing studies have used ML models to predict employee churn with high accuracy (Sisodia et al., 2017; Chaudhary et al., 2022; AbdElminaam et al., 2023; Bandyopadhyay & Jadhav, 2021; Mhatre et al., 2020; Thompson et al., 2022). However, ML prediction needs to provide an understanding of a complex data structure, and it is difficult to extract valuable insights and connections among employee features (El-Rayes et al., 2020).

Though studies reveal that ML predicts employee churn using various algorithms, minimal research has focused on explaining these predictions, limiting ML algorithms to opaque black boxes. The rapidly developing discipline of artificial intelligence (AI) offers an intriguing answer to the current problem of high ECn. By utilising sophisticated AI algorithms and adopting an explainable AI (XAI) approach, it is possible to discover complex patterns and indicators that point to an employee's likelihood of leaving an organisation (Marín Díaz et al., 2022). XAI not only reveals the decision-making process of the predictive model but also identifies the key features or key determinants with a particularly significant effect on the predicted outcome (Linardatos et al., 2020).

In addition, businesses equipped with this information might modify their recruitment procedures. Employers should proactively search for candidates who meet the specific criteria to increase employee retention and make staff more likely to stay with the firm for an extended period. The inclusion of AI, emphasising the ability to provide explanations, presents an unprecedented method to combat employee turnover. By analysing the complex network of factors that lead to employee churn, firms may strategically develop plans to retain current employees and attract new individuals who will likely form long-lasting connections with the organisation. This shift in perspective can significantly influence an organisation's stability, productivity, and overall development.

This requires extracting features and establishing connections to bolster the efficacy of predictive models using explainable AI (XAI). Shapley Additive exPlanations (SHAP), proposed by Lundberg and Lee in 2017, provide the solution to this limitation. SHAP benefits in balancing the precision and interpretability of results, which is otherwise a challenging endeavour. The study, therefore, clarifies crucial aspects of employee churn and the importance of using artificial intelligence in churn analysis. Overall, our research contributions are:

- Experiment: Comparable ML algorithms based on performance metrics for employee churn analysis.
- Development of a framework integrating ML and explainable AI for employee churn to address the limitation of algorithms on the scale of explainability, transparency and interpretability.
- Assess the applicability and implications of the framework in the organisation for further evaluation of employee churn.

Fig. 1 describes the research gap identified by the authors. Researchers focus on churn predictions using data-driven methods but have ignored the fundamental aspects of predicted results, i.e., not all decision-makers can understand these complex outputs. The explainability of the model is essential when output-based decisions are intertwined with ethical, security, confidentiality, and lawful concerns for the applicability of opaque models (Crawford, 2013). This study's novelty is developing a systematic approach to performing a study on interpretable employee churn that can be easily understood. This study identifies beneficial factors that impact EC that may be utilised to make well-informed decisions regarding employee retention and recruitment tactics.

By embracing the unique approach suggested in this study, businesses can examine and explain the various factors contributing to ECn. Properly comprehending the model is crucial for obtaining practical insights and understanding the reasoning behind predictions. Understanding this information is vital for individuals in positions of authority since it enables businesses to tailor their tactics and interventions precisely. This article encourages a fundamental change in how corporations see and handle ECn. By leveraging an interpretable approach, firms can strategically apply methods that improve employee retention and augment recruitment. The conclusions of this research provide valuable guidance for achieving a stable and motivated staff, promoting sustained organisational growth. The following study is organised in various sections. Section 2 summarises the literature on employee churn regarding ML and artificial AI. Section 3 elaborates on the proposed framework and preprocessing of the data and the research methods (logistic regression, random forest algorithms and SHAP) involved in the prediction and interpretability of the model. The analysis of the result is presented in Section 4, followed by a discussion of managerial and theoretical implications in Section 5. Section 6 presents the conclusion, followed by limitations and directions for future research in the last
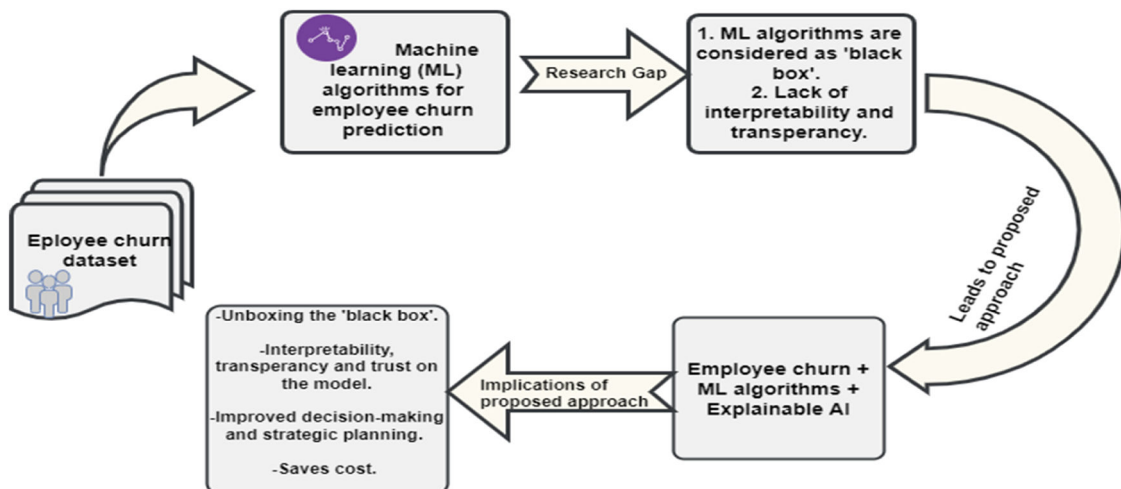

Fig. 1. Research Gap identification.

section.

Literature review

Employee churn

In recent decades there has been an increase in Human Resource Management academic literature due to the belief that organisation human resources were considered a valued resource for potentially achieving competitive advantage. Due to technological advancement and globalisation, there has been an upsurge in retaining talent through operative human resource practices (Du Plessis, 2006). Fig. 2 illustrates a primary search of publications related to ECn in the Scopus database. The authors identified 8394 documents using the combination of keywords ("Attrition" OR "Turnover" OR "Churn" AND "Employee"). The publications by year are graphically shown in Fig. 2 thus highlighting the growth of academic interest in this area in recent years.

Employee churn encompasses direct and indirect impacts ranging from replacement cost to low employee morale (Boushey & Glynn, 2012). Table 1 illustrates that existing research has demonstrated many factors that lead to employee churn. These factors analysis has contributed to the development of human resource analytics (Chalutz Ben-Gal et al., 2021). Demographic factors, namely age, marital status, and department, are determinants of employee churn (Chowdhury, 2015). Millennials prefer meaningful employment over well-paying professions (DeVaney, 2015). Chen (2020) developed a structural equation model that included confirmatory factor analysis and path analysis to assess the impact of leadership on job satisfaction and employee churn. The result confirms that working conditions have a mediating result on job satisfaction and leadership. Table 1 has compiled the studies focused on the factors/features that lead to employee churn in.

Machine learning (ML) and employee churn (ECn)

Human Resource Predictive Analysis has accelerated technological advancement and can achieve 100 % accuracy in human resource decision-making. Over the last decade, as analytics implementation has increased, it has entered human resources, but it has yet to gain significant momentum. The current market situation necessitates that the HR function goes beyond reporting to accurate forecasting. Data mining is vital for scrutinising the relationship between the data variables (Jayanthi et al., 2008). Gartner notes the adoption and integration of artificial intelligence (AI) increased by 270 % between 2015 and 2019.

Table 1

Literature review of factors contributing to employee churn.

Features	Authors
Age	Frian & Mulyani, 2018; Kabungaidze et al., 2013; Chih et al., 2016; Guha & Chakrabarti, 2014; Bose & Subha, 2021
Business travel	Setiawan et al., 2020; Srivastava & Tiwari, 2020
Department	Zhang, 2016; Sisodia et al., 2017
Distance from home	Ghazi et al., 2021; Gao et al., 2019; Jain & Nayyar, 2018
Education	Grzenda & Buczynski, 2015; Ju & Li, 2019; Ribes et al., 2017
Environment satisfaction	Setiawan et al., 2020; Lee at al., 2015; Edirisinghe & Manuel, 2015
Gender	Grzenda & Buczynski, 2015; Bose & Subha, 2021; Mölders et al., 2019
Job Involvement	Craig et al., 2013; Bugis et al., 2021; Demirtaş et al., 2017
Job Level	Yingjie et al., 2019; Ju & Li, 2019; Dwesini, 2019; Yin et al., 2021
Job role	Yingjie et al., 2019; Bamfo et al., 2018; Stamolampros et al., 2019
Job satisfaction	Edirisinghe & Manuel, 2015; Oosthuizen et al., 2016
Marital status	Ghazi et al., 2021; Madariaga et al., 2018; Tharani & Raj, 2020
Monthly income	Ghazi et al., 2021; Grzenda & Buczynski, 2015; Rybnicek et al., 2019
Number of companies worked to date	Bhuva & Srivastava, 2018; Yang & Islam, 2020; Setiawan et al., 2020; Al-Darraji et al., 2021
Overtime	Al-Darraji et al., 2021; Alhashmi, 2019
Percentage salary hike	Mate et al., 2022; Banerjee et al., 2017
Relationship satisfaction	Prasman & Ariyanto, 2020; Grund & Titz, 2022
Stock option level	Srivastava & Tiwari, 2020; Vignesh et al., 2018; Jain, 2017
Total working years	Setiawan et al., 2020
Training time last year	Srivastava & Tiwari, 2020
Work-life balance	Walsh & Taylor, 2007; Abdien, 2019; Kaushalya & Perera, 2018; Adame-Sánchez et al., 2018
Years at company	Grzenda & Buczynski, 2015; Ju & Li, 2019; Ribes et al., 2017
Years since the last promotion	Al-Darraji et al., 2021
Years with current manager	Khera and Divya 2019

Around 37 % of organisations will use AI in the workplace in 2019 (Costello, 2019). Cognitive technologies such as AI, robotics, and ML will replace 16 % of jobs by 2025. It is vital to comprehend the number of publications addressing ML methods in forecasting ECn during the research process. Furthermore, understanding the implementation of
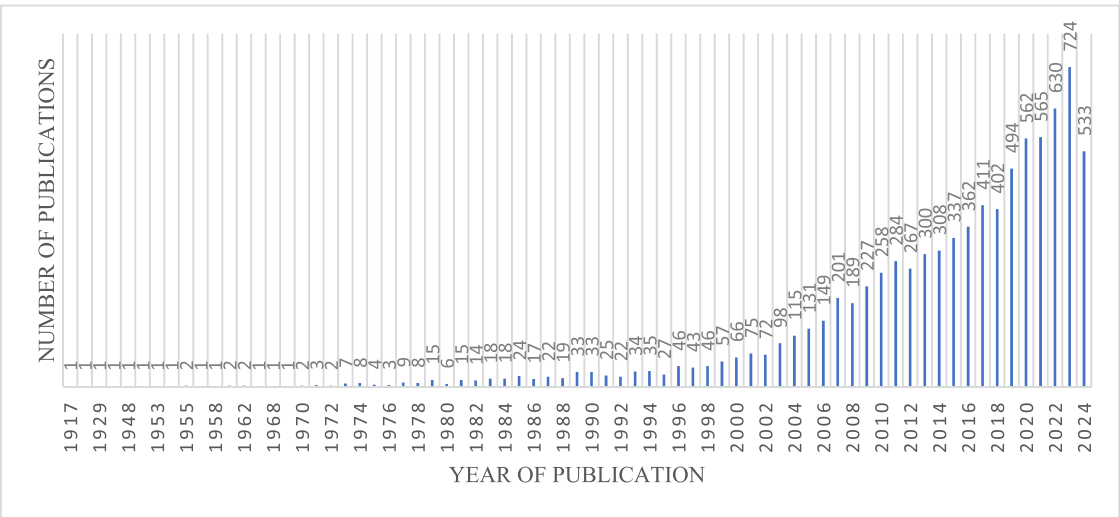

Fig. 2. Publications on employee churn.

strategies designed to reduce attrition is important.

Fig. 3 provides a graphical illustration showing the decrease in the publication of research articles to 90 with the introduction of this additional variable. Gao et al. (2019) used an improved random forest algorithm to predict employee turnover. Nagadevara et al., 2008 implemented artificial neural networks (ANN), classification trees (C5.0), classification and regression trees (CART) and discriminant analysis to study the association between employee demographics, absenteeism, tenure and employee churn. Saradhi and Palshikar (2011) studied the comparison of Naïve Bayes (NB), Support Vector Machines (SVM), LR, Decision Trees (DT) and RF algorithms for employee churn prediction.

Table 2 emphasises the employee churn predictions based on various ML algorithms, but the authors need to look into the explainability of these models based on artificial intelligence.

Explainable ai (XAI) and HRM

In recent years, ML algorithms have begun to concentrate on interpretability. Previously, the main emphasis was on algorithmic accuracy, which often resulted in a compromise between accuracy and interpretability. Increased precision is often associated with decreased interpretability (Stiglic et al., 2020). The very existence of this trade-off has contributed to the increasing importance of XAI as an essential tool. XAI strives to achieve improved prediction accuracy using opaque algorithms and tackles the crucial requirement of interpreting AI decisions. When assessing the interpretability of ML algorithms, researchers have classified them into two main categories (Gilpin et al., 2018). White-box models strive to establish a direct relationship between input factors and resulting outputs. By contrast, 'black-box models' lack decision rules that are simply comprehensible. It is pertinent to mention that interpretability is still being discussed, even among white-box models, where issues about their interpretability have been highlighted (Loyola-Gonzalez, 2019). The increased appreciation for XAI highlights the trend towards understandable AI systems, particularly when decisions have significant implications for individuals or society. Establishing the optimal balance between prediction accuracy and interpretability remains a constant challenge in the business. Striking a balance is crucial for cultivating confidence and ensuring the adoption of AI systems in several practical domains, like healthcare, banking, and more. Fig. 3 depicts the temporal evolution of the research quantity and algorithms' interpretability. Despite the widespread use of AI in HRM, understanding its legitimacy and workers' views is presently in the initial phase of development (Lu, 2021). Table 3 summarises the prominent research on XAI in human resource management.

The six papers shown in Fig. 4 apply to this study, as they were identified using a combination of keywords to identify the relevant documents on explainable AI in human resource management.

Recent studies have focused on multivariate ML algorithms for

Table 2

Previous studies on employee churn and machine learning.

Authors	Algorithms used	Explainable AI
Sisodia et al. (2017)	SVM, C5.0 Decision Tree Classifier, RF, KNN and NB	No
Alaskar et al. (2016)	LR, DT, NB, SVM, AdaBoost	No
Chaudhary et al. (2022)	CatBoost, SVM, DT, RF, XGBoost	No
Mhatre et al. (2020)	LR, DT, KNN, SVM, XGBoost, NB	No
Anh et al. (2020)	SVM, LR, RF	No
Bandyopadhyay and Jadhav (2021)	VM, RF and NB	No

Table 3

Publications related to XAI in HRM.

| Authors	Source	Cited by	Method used
Harl, Weinzierl, Stierle and Matzner (2020)	Journal of Decision Systems	51	Gated graph neural networks (GGNN)
Yuan et al. (2022)	Science Robotics	39	Bidirectional human-robot value alignment framework
Langer and König (2023)	Human Resource Management Review	31	Comprehensive analysis
Berman, de Fine Licht and Carlsson (2024).	Technology in Society	4	Case-study research design
Abonamah, La Torre, Poulin and Repetto (2022)	IEEE Xplore	2	Model-agnostic technique
Bhattacharya, Zuhair, Roy, Prasad and Savaliya (2022)	IEEE Xplore	2	SHAP
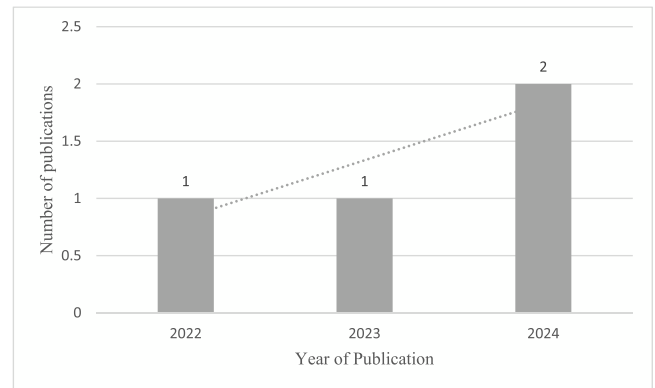
Fig. 4. Graphical representation of publications on XAI and Employee Churn.
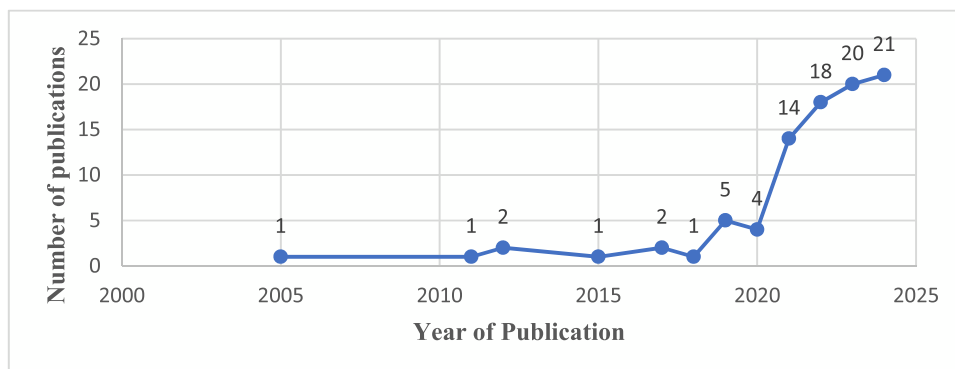

Fig. 3. Publications on employee churn and machine learning.

predicting employee churn, but very few have focused on explaining these predictions. The explainability of the model is essential when output-based decisions are intertwined with ethical, security, confidentiality, and lawful concerns for the applicability of opaque models (Crawford, 2013). However, the definition of explainability and interpretability are domain dependent. There is currently a dominant trend to prioritise using interpretability tools not exclusive to any one model. Automating interpretability becomes significantly simpler when we detangle the interpretation technique from the utilised model. By employing agnostic methodologies, we can substitute both the learning model and interpretation mechanism, resulting in significant scalability potential.

XAI explains to the end-users who depend on recommendations and decisions given by the AI system, and these end-users may not be the developers but policymakers, operators or managers (Arrieta, 2020). Complex ML models emphasise the importance of input variables for prediction. These variable attributions can be global, assessing variable relevance over the entire dataset, or local, focusing on a single variable wherein importance is measured at individual-level observations (Kazemitabar et al., 2017). Individual predictions are decomposed into variable contributions using local approaches (Ribeiro et al., 2018). Local methods provide the main advantage of presenting the functional form of the correlation between a variable and the output determined by the classifier. Independent of predictive accuracy, Shapley values show which variables impact a predicted value. As described in Fig. 4 and Table 4, it is apparent that studies on explainable AI in human resource management are limited and in the early stage of exploration. The authors have identified the research gap regarding XAI application in interpreting the ECn prediction. The novel contribution of this research lies in filling the research gap.

Methodology

Proposed framework

The methodology used in this study aligns with the knowledge discovery in databases (KDD) approach, and the established method for data mining, known as the cross-industry standard procedure for data mining (CRISP-DM) in the framework (Fig. 5). The IBM dataset from Kaggle is sourced. After the data source selection, we proceeded to progress to the next steps that constitute the model systematically.

The first step of CRISP-DM is to understand the domain, referring to understanding churn analysis. It is important to understand the HR processes, policies, terminologies, and challenges related to employee churn. Based on this understanding, the next step is to clarify the objectives. This involves identifying the critical drivers that lead to employee churn and establishing success criteria like prediction accuracy. We utilised the IBM HR dataset for churn analysis. To increase the

Table 4
Studies on interpretable models.

Authors	Model	Description
Kovalev, Utkin and Kasimov (2020)	Modified LIME	Solve an unconstrained convex optimisation problem
Zhang, Cho and Vasarhelyi (2022)	LIME and SHAP	Auditing the likelihood of significant errors in financial statements.
Tsoka et al. (2022)	LIME and SHAP	Classification of building energy performance certificates (EPC)
De Lange, Melsom, Vennerød and Westgaard (2022).	SHAP	Predicting credit default
Chen, Ke, Han, Gupta and Sivarajah (2024)	SHAP	The effect of product descriptions on sales forecasting
Wasilefsky, Caballero, Johnstone, Gaw and Jenkins (2024)	SHAP	Interpretable model for pilot candidate selection

effectiveness of the ML algorithm, data is pre-processed to increase the data quality, including processing the missing data/values, constant feature processing and label coding. Next, data transformation is making raw data into a structure that improves quality by making it easier to analyse. The dataset is randomly split into training and testing datasets in an 80:20 ratio. The training data are used to build the prediction model, while the unknown test data are used to evaluate the model's performance. The training dataset is subjected to 10-fold cross-validation. ML algorithms, namely LR, SVM, KNN, RF, decision tree classifier and Gaussian NB, are applied using Python. Therefore, LR and RF, which are performed with the highest accuracy, are used for further evaluation. In the next phase, SHAP is integrated to derive feature values and compare the SHAP results of both algorithms. The proposed framework used in this study is represented in Fig. 5.

Data demographics

To validate the study's objectives, the dataset from Kaggle is used for analysis. The dataset contains details of 1470 employees and 31 features, out of which 14 are continuous, and 17 are categorical variables. The dataset includes 882 males and 588 females, and the number of employees with marital statuses of married, single, and divorced are 673, 470, and 327, respectively. The data consists of thirteen categorical, thirteen continuous, and two flag variables. The 'churn' variable is the target variable.

Research methods

During the process's start, a series of methodologies will be used to carry out a preliminary exploratory data analysis. Subsequently, a set of ML algorithms will be used to create a predictive model with the highest potential accuracy for identifying ECn. Python is used for implementing many methods, including logistic regression (LR), support vector machine (SVM), K-nearest neighbours (KNN) random forest (RF), decision tree classifier, and Gaussian native Bayes (GNB). The LR and RF models with the highest accuracy are selected for further evaluation. During the next phase, the SHAP model is incorporated to calculate feature value and compare the SHAP outcomes for both algorithms. The structure utilised in this investigation is depicted in Fig. 5.

Logistic regression (LR) and random forest (RF)

LR states the connection between distinct variables, giving improved results in numerical data values (Gladence et al., 2015). LR defines the probability of realising dependent variable Y (Yan & Su, 2009) based on observation of input variable X and is more useful when the dependent variable is categorical (Yedida et al., 2018). The RF algorithm was proposed by Breiman in 2001 and is very popular as a classification and regression model. Random forest combines tree classifiers and offers to correct over-fitting for training datasets (Chaudhary et al., 2022). It is a non-parametric method that uses an ensemble of tree topologies. Besides being simple, it is recognised for its accuracy and capability to deal with a wide range of features. It estimates more complex patterns than the traditional ones and cannot be characterised by smoothness and standard sparseness (Biau & Scornet, 2016). The codes for top-performing RF and LR algorithms based on the dataset we used are mentioned below.

SHAP

SHAP determines each player's contribution to success based on the game theory. SHAP can explain multiple supervised learning models and assign importance values to each input variable for each prediction. Furthermore, recent research demonstrated that SHAP is the only method for identifying a new class of additive feature significance measures. Shapley's value defines the average marginal role of the eigenvalue with the consistency of local interpretation with global interpretation, which effectively overpowers the limitation of the LIME
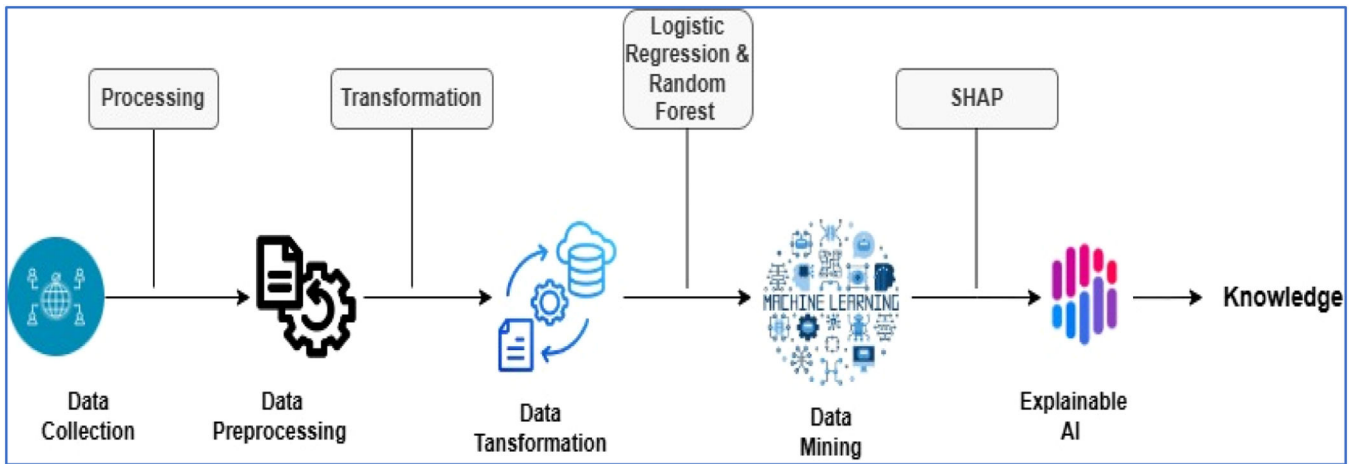
Fig. 5. Proposed framework.

method (Lundberg et al., 2020). The input variables are ranked in order of relevance; the highest mean SHAP value determines the most essential feature.

SHAP values measure feature contribution value to explain the black-box model. SHAP has an advantage over LIME as it can be used for any tree-based model, and each feature has its own set of SHAP values. Explainable models provide explanations made by machine learning models, which are feature-based and rule-based models. Feature-based models' output presents the top features to elucidate the machine learning prediction and associated weights (Ribeiro et al., 2016). However, the rule-based model functions on if-then-else rules (Oxborough et al., 2018). It functions on three properties, i.e. local accuracy that refers to the outcome, which is the total of the feature attributions and necessitates the model to match the output for the simplified input; missingness assures that no significance value is attributed to the missing features. Consistency ensures that change in more prominent impact features will maintain the attribution ascribed to that feature.

SHAP values aim to simplify the result of a function f into the sum of the individual impacts ϕ_i of each independently entered feature. We must recognise that the sequence in which the attributes are introduced holds significance for nonlinear functions. The SHAP values are determined by calculating the average of all attainable values, where $\sum_{i=0}^{M} \phi_i = f(x)$. The additive feature attribution methods utilise a model with interpretation g, represented by a linear function of binary variables.

$$g(z') = \phi_0 \sum_{i=1}^{M} \phi_i z'_i = f(\mathbf{x})$$

Where $z'_i \in \{0,1\}^M$, M is the number of input features, and $\phi_i \in$ R.

The concept of local precision refers to the sum of the feature attributions equivalent to the output of the function we are attempting to explicate. The lack of features suggests that z0 equals zero; no significance is assigned to them. Consistency refers to the principle that modifying a model to boost the influence of a particular feature will not result in a fall in the importance attributed to that characteristic.

Compared to LIME, SHAP is more computation-intensive, theoretically sound, robust and globally consistent for complex models and large datasets.

Analysis

The study aims to strengthen the existing literature by investigating the interpretability mechanisms employed by the machine learning models utilised in generating predictions (Sekaran & Shanmugam, 2022). This analysis will enable us to identify the critical variables in the

churn process, enhancing the business's internal processes to manage employee retention effectively. Based on the content presented in Section 3 of this study, the author examines the complete HR process by utilising the KDD and CRISP methodologies (Shafique & Qaiser, 2014) (as seen in Fig. 5. This examination will involve the identification of the following subprocesses:

The training dataset is subjected to 10-fold cross-validation, meaning that the dataset is divided into ten subsets of datasets and performs training and model evaluation on different datasets each time. Cross-validation aims to enhance performance and avoid model generalisation, i.e., overfitting and underfitting. The authors have used model parameter tuning for the given dataset. LR, SVM, KNN, RF, decision tree classifier, and Gaussian NB are evaluated on accuracy parameters. As per the dataset analysis, logistic regression outperforms random forest when several explanatory factors are fewer than or equal to the number of noisy variables. LR and RF gained the highest accuracy, 87.96 % and 87.29 %, respectively.

Table 5 demonstrates the result of five ML algorithms based on accuracy and accumulated AUC. As the number of explanatory factors in a dataset grows, the random forest's true and false positive rate increases. In binary classification, the class is categorised as positive and negative classes. A confusion matrix is a combination of four possible outcomes are true positive, symbolised as *TP* (the predicted positive value is true); true-negative, symbolised as *TN* (the predicted negative value is actual); and false-positive, symbolised as *FP* (the negative value is predicted as positive). False-negative symbolised as *FN* (positive value is predicted as negative). The confusion matrix measures the model's accuracy by comparing predicted values with the absolute ones. The performance of the RF and LR algorithms is measured using three performance metrics: accuracy (ACC), precision (PREC), F1 score, and area under the curve (AUC). These metrics are generated for all ten folds as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN}, \ \text{Precision} = \frac{TP}{TP + FP}, \ Recall$$
$$= \frac{TP}{TP + FN}$$

AUC measures the total area under the Receiver Operating

Table 5
Result based on comparison of five selected classifiers/algorithms.

Classifier/algorithm	Accuracy (in %)	Accumulated AUC
Logistic Regression	87.96	0.894
Support Vector Machine	83.61	0.685
Neural Network	86.96	0.867
XGBoost	86.62	0.885
Random Forest	87.29	0.882

Characteristic (ROC) curve, one of the criteria used to assess model performance. The AUC value varies from 0 to 1, with a value closer to 1 suggesting a more accurate model. The AUC score is 0 when 100 % of predictions are wrong; one whose predictions are 100 % correct has an AUC of 1.

We have used logistic regression and random forest algorithms for further analysis since the calculated accuracy is 87.96 % and 87.29 %, respectively (Table 6). The AUC - ROC curve measures the performance of classification problems at several threshold configurations. ROC is a likelihood and monotonic increasing curve, while the AUC quantifies the level or extent of distinguishability. It indicates how well the model can differentiate between classes. When the area under the ROC curve is high, true negative and accurate positive distributions do not intersect, reflecting that the classes have been appropriately segregated. AUC shows the performance of the model.

As a result, LR shows improved performance compared to the RF algorithm. The fine-tuned LR model has a better AUC score than the RF classifier. The variables with the most positive correlation are performance rating, monthly rate, number of companies worked, and distance from home. In contrast, the variables with the most negative correlation with churn are total working years, job level, years in a current role, monthly income, and age. Our study used a heatmap to demonstrate a correlation between variables/features; see Fig. 6.

Heatmap depicts a data matrix with a colour gradient indicating numerical disparities in Fig. 4. Diverging palettes fix colours at both the lower and upper ends of the data and in the middle; it is ideal for data that goes in both negative and positive directions. It provides an excellent picture of the matrix's most significant and smallest values. Heatmaps describe relationships between variables in the form of colours instead of numbers. In simpler terms, it represents a correlation matrix to show a correlation between multiple variables. Closer to zero values indicate that the two variables have no linear relationship. The correlation is more positively connected when its value is close to 1. SHAP balances the accuracy and comprehensibility of black box models generated by machine learning. SHAP values show the distribution of each variable's impact. The feature is represented on the y-axis, while the Shapely values are mentioned on the x-axis. SHAP summary plots depict the mean effect of a variable on the complete result and display the knowledge using bars of appropriate length or dimension.

SHAP demonstrates the values and features more understandably. As shown in Fig. 7, it can be interpreted that the overtime feature is the highest predictor of employee churn, followed by monthly income, stock option, job level and so on, while the job role of sales executive has the most negligible chances of being churned. However, compared to the SHAP result of logistic regression, the overtime feature/variable is the highest predictor and contributor of employee churn, followed by the employees whose marital status is single. The result is shown in Fig. 8. SHAP plots differ for different models; however, the general features trend remains constant in both models.

SHAP presents substantial logical analysis by discovering the most influential features that enabled the model to effectively forecast whether or not an employee will leave the organisation.

Discussion

With the introduction of advanced technology, human resource management is going through a paradigm change (McCartney & Fu, 2022). The decision-making process in the field of human resources (HR), specifically with employee retention or hiring procedures, is the

essential foundation of every company's business model. Existing literature has primarily focused on mechanism-driven predictions and identifying factors contributing to employee churn (Tao et al., 2021); however, churn prediction is complex due to diverse parameters. Therefore, assessing and forecasting employee turnover are crucial in HR procedures.

The study aims to establish a thorough methodology to execute interpretability in a context such as human resources management. This will create a baseline for formulating guidelines derived from the study to conduct recruitment procedures that align with the drivers of attrition and serve as a complement to retention operations. The proposed method facilitates comprehension and prioritisation of the necessity for using AI in business applications, particularly in areas about Management of employees, which is considered the most essential and significant asset for the organisation. The interpretability of the acquired results supports the findings derived based on the preliminary data analysis. Interpreting the forecasts allows us to figure out the variables that influence turnover risk and take proactive steps to mitigate them. The novel contributions of this study are outlined below:

- Inclusive Methodological Framework: This paper proposes an integrated and comprehensive methodology approach to incorporating interpretability into the ML pipeline, particularly emphasising the HR field. The framework includes data preprocessing, model selection, hyperparameter tuning, and a model-agnostic interpretability tool (SHAP). It provides an organised approach to improving the transparency and applicability of predictive models.
- Improved Decision-Making in HR: The research aims to enhance HR professionals' capabilities in making well-informed and impactful decisions about ECn. Analysing the significant elements contributing to attrition will inform HR initiatives to help retain talented employees and create a more efficient and productive workplace.
- Enhancing Trust through Model-Agnostic Interpretability: Adopting the interpretability model for complex algorithms like LR and RF enhances trust and transparency in prediction. Consistently interpreting and illustrating model predictions to stakeholders increases comprehension and approval of ML results, particularly in HR decision-making.
- Exploratory data analysis (EDA) enables an initial comprehension of how variables interact regarding the outcome variable, churn. Before initiating the ML process, it is imperative to proactively identify any potential biases present in the dataset. Moreover, the approach used ought to be flexible and capable of adapting, continually gaining knowledge from employee data to prevent the introduction of bias in a proactive manner.

These novel aspects strengthen the comprehensive ECn prediction. Implementing HRM will facilitate informed managerial, administrative, and adequate staff retention strategies.

Theoretical contribution

This paper has illustrated the importance of employee churn prediction and the explainability of results using XAI. To validate, we relied on the existing dataset with variables proven to contribute to employee churn in the literature (Table 1). We have identified the gap in the existing ML models used in employee churn and focused on filling the gap in the existing literature (Table 2).

Regardless of the theoretical account for human decision-making, people, including experts, only sometimes make the best choices. To overcome these limitations, predictive analytics scrutinise the hidden patterns in the HR database to derive predictions. This approach decreases the associated business risk through churn analysis. However, the decision-makers cannot understand the algorithms used in these predictions due to a lack of interpretability, making it difficult to trust these models, leading to ML being considered a black box. Explainable

Table 6
Performance metrics of top-performing classifiers.

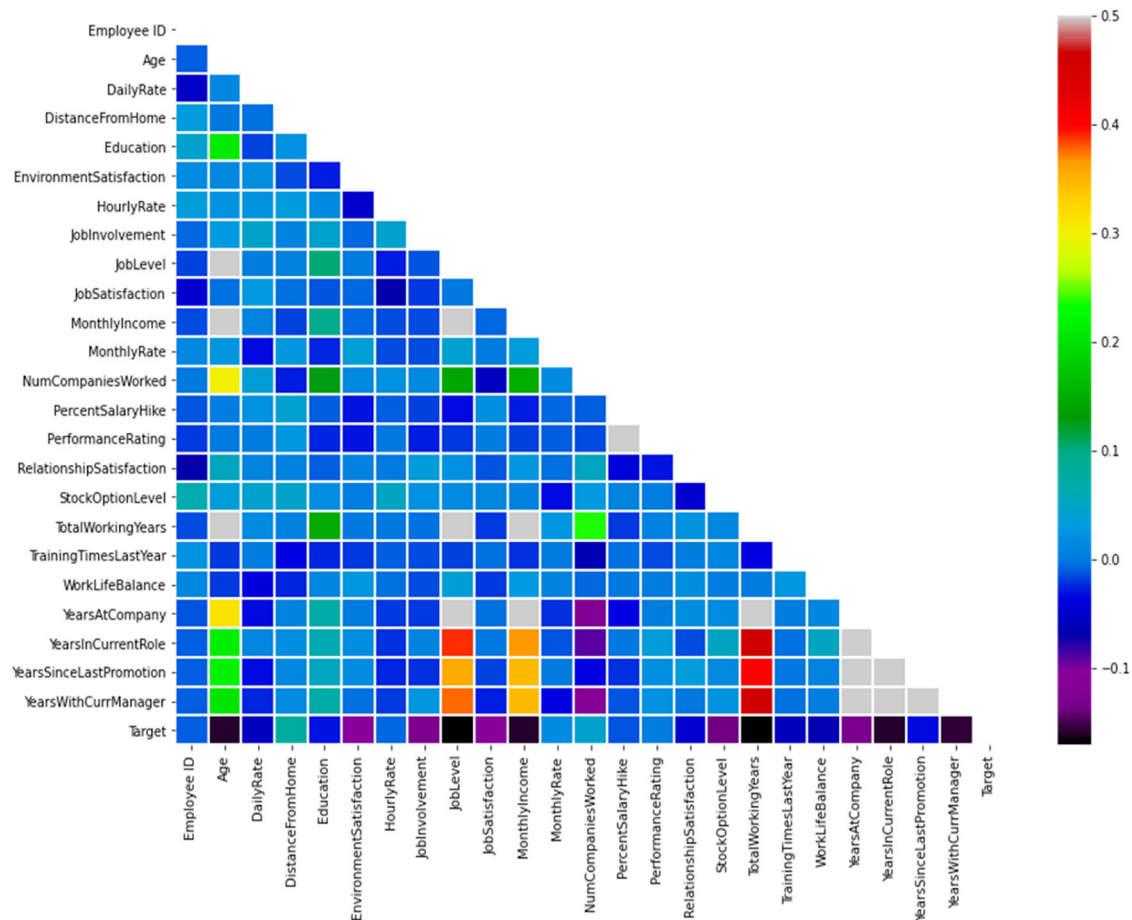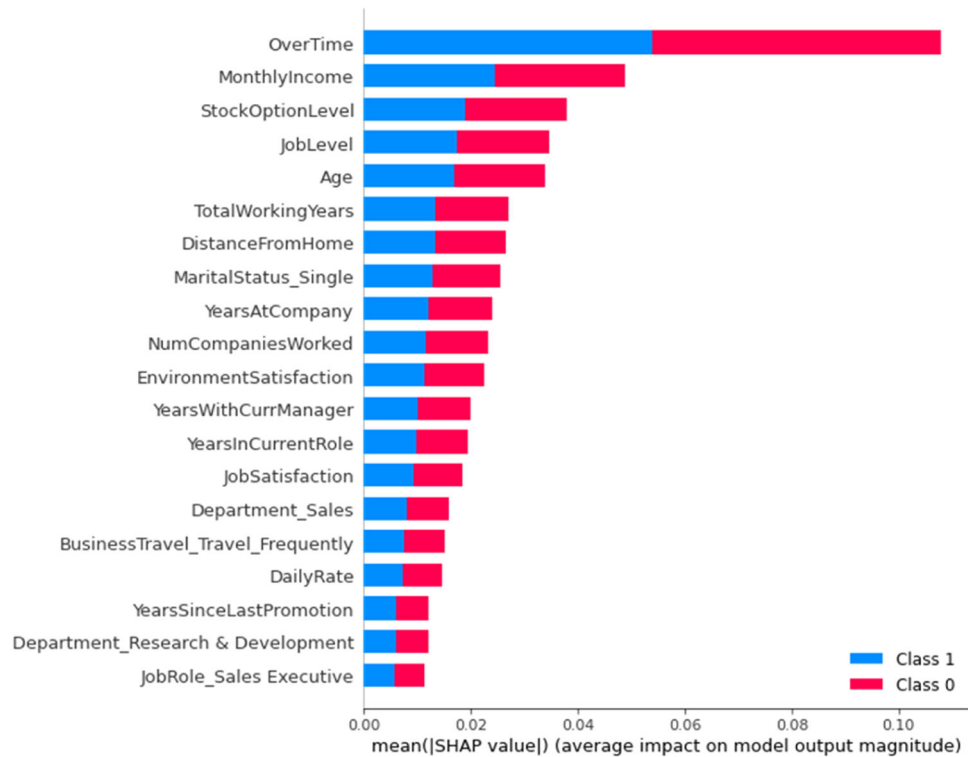| Classifier/Algorithm	Accuracy	Precision	Recall	AUC
Logistic Regression	0.879	0.86	0.75	0.894
Random Forest	0.872	0.84	0.86	0.882

Fig. 6. Heatmap for feature comparison.

Fig. 7. Summary plot representation for Random Forest for data instances.
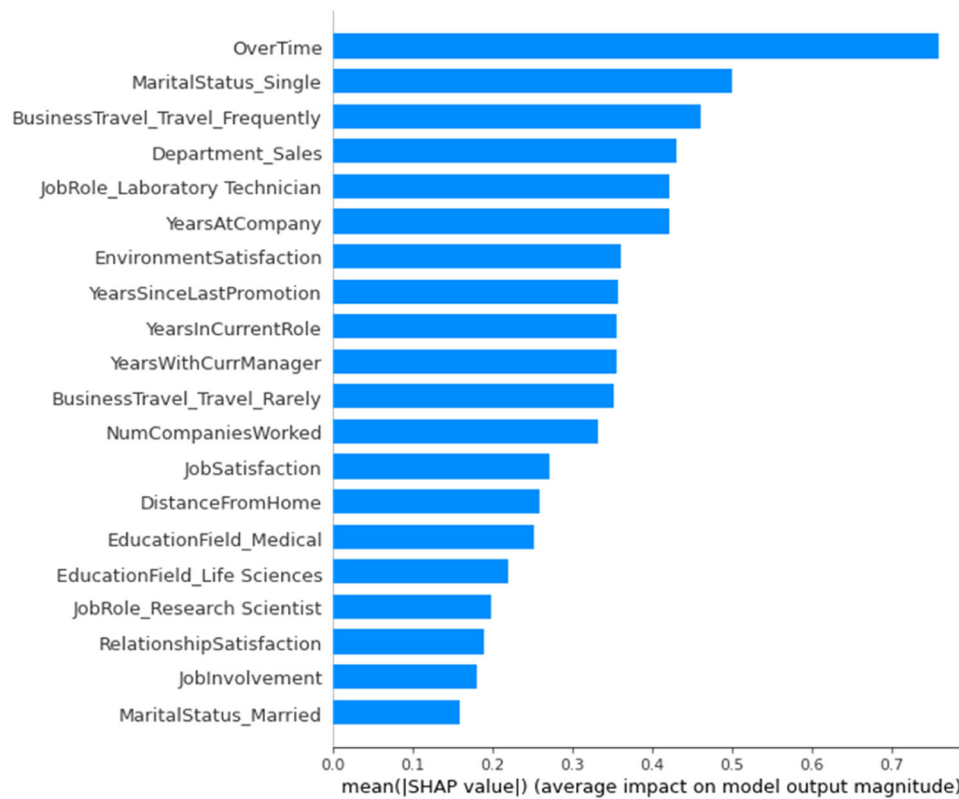
Fig. 8. Graphic depiction of Logistic Regression for data instances.

AI techniques answer the questions of transparency, interpretability, and trust. It focuses on the why aspect of prediction results. The constant increase in business pressures and success stories of AI has paved the way for the inclusion of AI in different aspects of business, proving to be a competitive advantage.

Transparency and explainability have accelerated the adoption of AI, ensuring ethics and compliance and providing strategic insights into business (Oxborough et al., 2018) and subsequently improving decision-making and strategic planning. Explainable AI discovers hidden information related to different facets of business, collecting information from an employee database, social media, and transactions. It works on the cause-and-effect model to identify patterns. Such correlations are seen utilising explainable algorithms, showing the main drivers behind them. Transparency supports the assessment of the precision of output forecasting, identifying the potential risks associated with model utilisation, and anticipating situations in which the model might not perform as expected. It can also discourage adversarial assaults by educating business users on manipulating model inputs to impact outputs. Individuals in charge of a model can anticipate instances of failure and take action accordingly by obtaining an intuitive awareness of its behaviour.

Industrial Revolution 4.0 is centred on advanced technologies that convert information to gain competitiveness, growth, efficiency, and effectiveness (Park et al., 2019; Neumeyer et al., 2020). This era counts on cutting-edge technology like AI and equipment for increased production and wealth creation. Businesses use robotics and AI to automate repetitive and simple operations while projecting algorithms to aid sophisticated decision-making. Providing explanations for human decisions is required to improve human predictions or decisions. Decisions often are made based on biases and heuristics, sometimes based on the rationale of the decision-maker. In contrast to heuristics and biases, bounded rationality in human decision-making using complying with limitations is an alternative theory.

Customers, regulators, and industry confederations pressure

businesses to guarantee that their AI systems adhere to ethical standards and function within publicly reasonable bounds. Securing vulnerable consumers, ensuring data privacy, fostering ethical behaviour, the performance of accounting firms (Oberoi et al., 2022), environmental concerns (Gaur et al., 2021) and preventing bias are all regulatory concerns. Using explainable models is one technique to check for bias and make decisions that do not breach business ethics or harm a company's reputation. Considering the factors mentioned above, the proposed framework ensures the explainability and interpretability of the prediction model. It lets the decision-makers identify the factors that may lead to employee churn and develop individual preventive plans.

Managerial implications

Though employee churn is like customer churn, in some organisations, the cost of ECn is comparatively higher and requires increased attention from the researchers. The high employee churn rate has unfavourable consequences for the company, affecting ongoing initiatives and causing customer discontent. Replacing individuals with specialised skill sets can be challenging. Considering the increasing employee churn rate, identifying factors leading to churn has become paramount. Existing literature has worked on the factors mentioned in Table 1. However, more than merely identifying and representing facts is required in the dynamic business environment. The inability to understand why churn occurs paralyses a company's ability to prevent it. Decision-makers must identify the factors and make their predictions. Therefore, HR professionals need to keep abreast in the industry to keep their organisations updated about the change.

Organisation leaders can better understand person-organisation fit to create or improvise internal process design to increase job satisfaction and employee retention, thus minimising employee churn. Employees can influence client satisfaction favourably, but newly hired employees may not. This demonstrates how high employee turnover impacts a company's efficiency since client needs may need to be appropriately

handled.

Industrial Revolution 4.0 has transitioned from 'employee engagement' to 'employee experience', where employees want to contribute. Design thinking can improve the employee experience by integrating the three ways experience mapping means identifying the actual experiences and needs, touchpoint simplification referring to improvising emotional points of contact between employees and the organisation and rapid prototyping. Recreational programmes can be opted to break the monotony of work, and employees must be awarded and appreciated for their contribution and hard work. To reduce the repercussions of employee churn, efforts must be made to reduce it. A positive employee workplace facilitates firm profitability and potentially increases employee satisfaction (Storey, 2016).

Summarising the above points, it is crucial to retain employees and take proactive steps to predict employee churn, which can be a competitive advantage. Thus, ML is being integrated, but we must remember that all decision-makers and HR professionals may need to be more comfortable understanding the 'black-box' models.

Explainability ensures that AI-based decisions are based correctly. Rather, the transparency of model predictions makes the results more interpretable. XAI may have substantial social roles in the future that will include learning and explaining individuals and connecting knowledge to get inter-disciplinary insights, connecting previous knowledge for further application. Ozcan et al. (2020) summarised that HR mining can be used for academic, industrial, and governmental needs.

Conclusions

Employee churn is the most prominent issue that decision-makers of organisations have to deal with, as adverse effects range from low morale to disruption in productivity and long-term growth strategies. This paves the way for the prediction of the risk factors that prompt employee churn using ML algorithms. Though analytics has contributed to the business in many ways, it also faces criticism regarding the 'black-box' method and lacks actionable matters (Gelbard et al., 2018). A critical component of our research is the interpretability of predictions, providing an improved awareness of the factors which contribute to employee churn. Besides the accuracy of the prediction model, users' trust directly impacts the usability of the model (Gelbard et al., 2018). Data mining techniques are an inherent part of industry revolution 4.0. Due to the diverse applicability of ML, it is not limited to specific industries but also across FMCG (Tarallo et al., 2019), hospitality (Gaur et al., 2021), healthcare (Gaur et al., 2021; Mahbub et al., 2022; Saeed et al., 2022), and many others. Machine learning models must achieve high AUC values.

Our study focuses on most of the features that contribute to employee churn and ranks them based on their importance and correlation to churn. Implementing machine learning algorithms can make decision-making tasks easy, but adding SHAP to the system can act as a catalyst for non-technical decision-makers to understand and interpret the result. Based on the SHAP summary plots of logistic regression and random forest, the features that contributed positively and negatively to employee churn are approximately the same. However, the order and weight differ in both models. Businesses can modify institutional policies to address the grievances that contribute to employee churn. Resources, such as money and time, can be invested in more profitable ventures. Assessing the influence of each variable on attrition permits management to determine the underlying reasons for the issue and concentrate their efforts on addressing these problems proactively before they worsen. In addition, it assists firms in prioritising concerns and addressing the most critical ones if they cannot handle all the issues simultaneously.

Limitations and direction for future research

We have used the proposed framework (Fig. 5) on the dataset available on Kaggle, but this can be used on real-time data for further research to increase the efficiency and effectiveness of business. The hyperparameter tuning can be executed for real-time dataset for higher accuracy. Academics have focused on theoretical aspects of explainable AI; however, more focus can be laid on integrating XAI in different possible ways. Our study addressed employee churn; however, future research can focus on problem statements related to HRM and cross-functional issues in business. One of the limitations of XAI implementation is the requirement of technical expertise to derive the results of the problem statement; then only it can help decision-makers. XAI may have substantial social roles that include learning interpretability and coordination with other agents to develop cross-disciplinary understanding based on previous knowledge to enhance the future of XAI.

Funding

CRediT authorship contribution statement

Meenu Chaudhary: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Loveleen Gaur:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Amlan Chakrabarti:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Gurmeet Singh:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Paul Jones:** Writing – review & editing, Visualization, Validation, Supervision, Resources, Project administration, Investigation, Formal analysis, Data curation, Conceptualization. **Sascha Kraus:** Validation, Writing – original draft, Writing – review & editing.

Disclosure

The authors have no pertinent financial or non-financial conflicts of interest to disclose.

References

AbdElminaam, D. S., Maged, M., Mousa, M. K., Younis, A. O., Abdelsalam, M. S., Hisham, Y., & Talaat, T. (2023). EmpTurnoverML: An efficient model for employee turnover and customer churn prediction using machine learning algorithms. In *2023 International Mobile, Intelligent, and Ubiquitous Computing Conference* (pp. 1–8). IEEE.

Abdien, M. K. (2019). Impact of communication satisfaction and work-life balance on employee turnover intention. *Journal of Tourism Theory and Research, 5*, 228–238. https://doi.org/10.24288/jttr.526678

Abonamah, A., La Torre, D., Poulin, M., & Repetto, M. (2022). Explainable artificial intelligence in Human resources: A computational study. In *2022 International Conference on Data Analytics for Business and Industry (ICDABI)* (pp. 1–6). IEEE.

Adame-Sánchez, C., Caplliure, E. M., & Miquel-Romero, M. J. (2018). Paving the way for competition: Drivers for work-life balance policy implementation. *Review of Managerial Science, 12*, 519–533. https://doi.org/10.1007/s11846-017-0271-y

Alaskar, L., Crane, M., & Alduailij, M. (2016). Employee turnover prediction using machine learning. In A. Alfaries, H. Mengash, A. Yasar, & E. Shakshuki (Eds.), *Advances in Data Science, Cyber Security and IT Applications. ICC 2019. Communications in Computer and Information Science* (pp. 301–316). Springer. https://doi.org/10.1007/978-3-030-36365-9_25.

Al-Darraji, S., Honi, D. G., Fallucchi, F., Abdulsada, A. I., Giuliano, R., & Abdulmalik, H. A. (2021). Employee attrition prediction using deep neural networks. *Computers, 10*, 141. https://doi.org/10.3390/computers10110141

Alhashmi, S. M. (2019). Towards understanding employee attrition using a decision tree approach. In *2019 International Conference on Digitization* (pp. 44–47).

Alsheref, F. K., Fattoh, I. E., & Ead, M. W. (2022). Automated prediction of employee attrition using ensemble model based on machine learning algorithms. *Computational Intelligence and Neuroscience, 2022*(1), Article 7728668.

Anh, N. T. N., Tu, N. D., Solanki, V. K., Giang, N. L., Thu, V. H., Son, L. N., Loc, N. D., & Nam, V. T. (2020). Integrating employee value model with churn prediction. *International Journal of Sensors Wireless Communications and Control, 10*(4), 484–493.

Arrieta, A. B. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion, 58*, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

Bamfo, B. A., Dogbe, C. S. K., Mingle, H., & Wright, L. T. (2018). Abusive customer behaviour and frontline employee turnover intentions in the banking industry: The mediating role of employee satisfaction. *Cogent Business & Management, 5*, 1–15. https://doi.org/10.1080/23311975.2018.1522753

Bandyopadhyay, N., & Jadhav, A. (2021). Churn prediction of employees using machine learning techniques. *Tehnički Glasnik, 15*, 51–59.

Banerjee, A., Ghosh, R. K., & Ghosh, M. (2017). A study on the factors influencing the rate of attrition in IT sector: Based on Indian scenario. *Pacific Business Review International, 9*, 1–13. http://www.pbr.co.in/2017/2017_month/Jan/01.pdf.

Berman, A., de Fine Licht, K., & Carlsson, V. (2024). Trustworthy AI in the public sector: An empirical analysis of a Swedish labor market decision-support system. *Technology in Society, 76*, Article 102471.

Bhattacharya, P., Zuhair, M., Roy, D., Prasad, V. K., & Savaliya, D. (2022). Aajeevika: Trusted explainable AI based recruitment scheme in smart organisations. In *2022 5th International Conference on Contemporary Computing and Informatics (IC3I)* (pp. 1002–1008). IEEE.

Bhuva, K., & Srivastava, K. (2018). Comparative study of the machine learning techniques for predicting the employee attrition. *International Journal of Research and Analytical Reviews, 5*, 568–577. https://ijrar.org/papers/IJRAR1903202.pdf.

Biau, G., & Scornet, E. (2016). A random forest guided tour. *Test, 25*, 197–227. https://doi.org/10.1007/s11749-016-0481-7

Bose, M. R., & Subha, K. (2021). A study on the application of HR analytics on talent acquisition, compensation & benefits and employee turnover in the Indian IT industry. *Sambodhi, 44*, 66–76.

Boushey, H., & Glynn, S. J. (2012). There are significant business costs to replacing employees. *Center for American Progress, 16*, 1–9.

Bugis, M., Dedi, P. E. S., & Saparuddin, S. (2021). The effect of job involvement and work stress on turnover intention with organizational commitment as an intervening variable PT. Perkebunan Minanga Ogan. *International Journal of Multicultural and Multireligious Understanding, 8*(9), 421–431. https://doi.org/10.18415/ijmmu.v8i9.3033

Chalutz Ben-Gal, H. (2019). An ROI-based review of HR analytics: Practical implementation tools. *Personnel Review, 48*(6), 1429–1448. https://doi.org/10.1108/PR-11-2017-0362

Chalutz Ben-Gal, H., Avrahami, D., Pessach, D., & Singer, G. (2021). A machine learning examination of turnover: Hidden patterns and new insights. In *, 2021. Academy of Management Proceedings* (p. 12152). https://doi.org/10.5465/AMBPP.2021.158

Chaudhary, M., Gaur, L., Jhanji, N. Z., Masud, M., & Aljahdali, S. (2022). Envisaging employee churn using MCDM and machine learning. *Intelligent Automation & Soft Computing, 33*, 1009–1024.

Chen, J. K. C. (2020). Perspective on the influence of leadership on job satisfaction and lower employee turnover in the mineral industry. *Sustainability, 12*, 1–16. https://doi.org/10.3390/su12145690

Chen, S., Ke, S., Han, S., Gupta, S., & Sivarajah, U. (2024). Which product description phrases affect sales forecasting? An explainable AI framework by integrating WaveNet neural network models with multiple regression. *Decision Support Systems, 176*, Article 114065. https://doi.org/10.1016/j.dss.2023.114065

Chih, Y.-Y., Kiazad, K., Zhou, L., Capezio, A., Li, M., & Restubog, S. L. D. (2016). Investigating employee turnover in the construction industry: A psychological contract perspective. *Journal of Construction Engineering and Management, 142*(6), Article 04016006. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001101

Chowdhury, F. (2015). Demographic factors impacting employee turnover in the private banking sector of Bangladesh. *International Journal of Business and Management Invention, 4*, 46–54.

Costello, K. (2019). *Gartner survey shows 37 percent of organisations have implemented ai in some form.* Gartner. https://www.gartner.com/en/newsroom/press-releases/2019-01-21-gartner-survey-shows-37-percent-of-organizations-have.

Craig, C. A., Allen, M. W., Reid, M. F., Riemenschneider, C. K., & Armstrong, D. J. (2013). The impact of career mentoring and psychosocial mentoring on affective organisational commitment, job involvement, and turnover intention. *Administration & Society, 45*, 949–973.

Crawford, K. (2013). The hidden biases of big data. *Harvard Business Review, l*. https://hbr.org/2013/04/the-hidden-biases-in-big-data.

De Lange, P. E., Melsom, B., Vennerød, C. B., & Westgaard, S. (2022). Explainable AI for credit assessment in banks. *Journal of Risk and Financial Management, 15*(12), 556.

Demirtaş, Ö., Arslan, A., & Karaca, M. (2017). Why perceived organisational and supervisory family support is important for organisations? Evidence from the field. *Review of Managerial Science, 13*, 841–869. https://doi.org/10.1007/s11846-017-0264-x

DeVaney, S. A. (2015). Understanding the millennial generation. *Journal of Financial Service Professionals, 69*, 11–15.

Du Plessis, A. (2006). An overview of the influence of globalisation and internationalisation on domestic human resource management in New Zealand. In *, 5. International conference of Australian and New Zealand Academy of Management* (pp. 1–21).

Dwesini, N. F. (2019). Causes and prevention of high employee turnover within the hospitality industry: A literature review. *African Journal of Hospitality, Tourism and Leisure, 8*(3), 1–15. https://www.ajhtl.com/uploads/7/1/6/3/7163688/article_38_vol_8_3_2019.pdf.

Edirisinghe, S. D., & Manuel, S. (2015). Factors affecting sales employee turnover in hotel & travel industry: A literature review. *Journal of Humanities and Social Science, 24*, 35–54.

El-Rayes, N., Fang, M., Smith, M., & Taylor, S. M. (2020). Predicting employee attrition using tree-based models. *International Journal of Organizational Analysis, 28*(6), 1273–1291.

Fallucchi, F., Coladangelo, M., Giuliano, R., & Luca, E. W. D. (2020). Predicting employee attrition using machine learning techniques. *Computers, 9*(4), 86. https://doi.org/10.3390/computers9040086

Frian, A., & Mulyani, F. (2018). Millennials employee turnover intention in Indonesia. *Innovative Issues and Approaches in Social Sciences, 11*, 1855-0541.

Gao, X., Wen, J., & Zhang, C. (2019). An improved random forest algorithm for predicting employee turnover. *Mathematical Problems in Engineering, 2019*, 1–12. https://doi.org/10.1155/2019/4140707

Gaur, L., Singh, G., Solanki, A., Jhanji, N. Z., Bhatia, U., Sharma, S., Verma, S., Kavita, Petrović, N., Fazal Ijaz, M., & Kim, W. (2021). Disposition of youth in predicting sustainable development goals using the neuro-fuzzy and random forest algorithms. *Human-Centric Computing and Information Sciences, 12*, 1–19. http://repository.usp.ac.fj/id/eprint/12807.

Gelbard, R., Ramon-Gonen, R., Carmeli, A., Bittmann, R. M., & Talyansky, R. (2018). Sentiment analysis in organisational work: Towards an ontology of people analytics. *Expert Systems, 35*, 12289.

Ghazi, A., Elsayed, S. I., & Khedr, A. E. (2021). A proposed model for predicting employee turnover of information technology specialists using data mining techniques. *International Journal of Electrical and Computer Engineering Systems, 12*, 113–121. https://doi.org/10.32985/ijeces.12.2.6

Gilpin, L. H., Bau, D., Yuan, B. Z., Bajwa, A., Specter, M., & Kagal, L. (2018). Explaining explanations: An overview of interpretability of machine learning. In *2018 IEEE 5th International Conference on data science and advanced analytics* (pp. 80–89). IEEE.

Gladence, L. M., Karthi, M., & Anu, V. M. (2015). A statistical comparison of logistic regression and different Bayes classification methods for machine learning. *ARPN Journal of Engineering and Applied Sciences, 10*, 5947–5953.

Grund, C., & Titz, K. (2022). Affective commitment through further training: The roles of firm provision and employee participation. *Review of Managerial Science, 16*, 1195–1226. https://doi.org/10.1007/s11846-021-00460-1

Grzenda, W., & Buczynski, M. K. (2015). Estimation of employee turnover with competing risks models. *Folia Oeconomica Stetinensia, 15*, 53–65.

Guha, S., & Chakrabarti, S. (2014). Employee turnover: A study on information technology sector. *Journal of Business & Management, 20*, 71–87.

Harl, M., Weinzierl, S., Stierle, M., & Matzner, M. (2020). Explainable predictive business process monitoring using gated graph neural networks. *Journal of Decision Systems, 29*(sup1), 312–327.

Jain, D. (2017). *Doctoral dissertation.* National College of Ireland.

Jain, R., & Nayyar, A. (2018). Predicting employee attrition using xgboost machine learning approach. In *2018 International Conference on System Modeling & Advancement in Research Trends (SMART)* (pp. 113–120).

Jayanthi, R., Goyal, D. P., & Ahson, S. I. (2008). Data mining techniques for better decisions in human resource management systems. *International Journal of Business Information Systems, 3*, 464–481.

Ju, B., & Li, J. (2019). Exploring the impact of training, job tenure, and education-job and skills-job matches on employee turnover intention. *European Journal of Training and Development, 43*, 214–231. https://doi.org/10.1108/EJTD-05-2018-0045

Kabungaidze, T., Mahlatshana, N., & Ngirande, H. (2013). The impact of job satisfaction and some demographic variables on employee turnover intentions. *International Journal of Business Administration, 4*, 53–65.

Kaushalya, R. K. N., & Perera, G. D. N. (2018). The impact of work life balance on employee turnover intention in selected private banks in Colombo District, Sri Lanka. *Human Resource Management Journal, 6*(1). https://doi.org/10.31357/hrmj.v6i1.3582

Kazemitabar, J., Amini, A., Bloniarz, A., & Talwalkar, A. S. (2017). Variable importance using decision trees. In *Advances in neural information processing systems, 30* pp. 426–435). NIPS.

Khera, S. N., & Divya. (2019). Predictive modelling of employee turnover in Indian IT industry using machine learning techniques. *Vision, 23*(1), 12–21. https://doi.org/10.1177/0972262918821221

Kovalev, M. S., Utkin, L. V., & Kasimov, E. M. (2020). SurvLIME: A method for explaining machine learning survival models. *Knowledge-Based Systems, 203*, Article 106164.

Langer, M., & König, C. J. (2023). Introducing a multi-stakeholder perspective on opacity, transparency and strategies to reduce opacity in algorithm-based human resource management. *Human Resource Management Review, 33*(1), Article 100881.

Lee, J. S., Back, K. J., & Chan, E. S. (2015). Quality of work life and job satisfaction among frontline hotel employees: A self-determination and need satisfaction theory approach. *International Journal of Contemporary Hospitality Management, 27*, 768–789. https://doi.org/10.1108/IJCHM-11-2013-0530

Liborius, P., & Kiewitz, C. (2022). When leader humility meets follower competitiveness: Relationships with follower affective trust, intended and voluntary turnover. *Journal of Vocational Behavior, 35*, Article 103719.

Linardatos, P., Papastefanopoulos, V., & Kotsiantis, S. (2020). Explainable ai: A review of machine learning interpretability methods. *Entropy, 23*(1), 18.

Loyola-Gonzalez, O. (2019). Black-box vs. white-box: Understanding their advantages and weaknesses from a practical point of view. *IEEE access : practical innovations, open solutions, 7*, 154096–154113.

Lu, C. H. (2021). The impact of artificial intelligence on economic growth and welfare. *Journal of Macroeconomics, 69*, Article 103342. https://doi.org/10.1016/j.jmacro.2021.103342

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S-I. (2020). From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence, 2*(1), 56–67. https://doi.org/10.1038/s42256-019-0138-9

Madariaga, R., Oller, R., & Martori, J. C. (2018). Discrete choice and survival models in employee turnover analysis. *Employee Relations, 40*, 381–395. https://doi.org/10.1108/ER-03-2017-0058

Mahbub, M. K., Biswas, M., Gaur, L., Alenezi, F., & Santosh, K. C. (2022). Deep features to detect pulmonary abnormalities in chest X-rays due to infectious DiseaseX: Covid-19, pneumonia, and tuberculosis. *Information Sciences, 592*, 389–401.

Marín Díaz, G., Galán, J. J., & Carrasco, R. A. (2022). XAI for churn prediction in B2B models: A use case in an enterprise software company. *Mathematics, 10*(20), 3896.

Mate, Y., Potdar, A., & Priya, R. L. (2022). Ensemble methods with bidirectional feature elimination for prediction and analysis of employee attrition rate during COVID-19 pandemic. In *Artificial Intelligence and Technologies* (pp. 89–101). https://doi.org/10.1007/978-981-16-6448-9_11

McCartney, S., & Fu, N. (2022). Bridging the gap: Why, how and when HR analytics can impact organisational performance. *Management Decision, 60*, 25–47. https://doi.org/10.1108/MD-12-2020-1581

Mhatre, A., Mahalingam, A., Narayanan, M., Nair, A., & Jaju, S. (2020). Predicting employee attrition along with identifying high risk employees using Big data and Machine Learning. In *2nd International Conference on Advances in Computing, Communication Control and Networking* (pp. 269–276).

Mölders, S., Brosi, P., Spörrle, M., & Welpe, I. M. (2019). The effect of top management trustworthiness on turnover intentions via negative emotions: The moderating role of gender. *Journal of Business Ethics, 156*, 957–969. https://doi.org/10.1007/s10551-017-3600-9

Nagadevara, V., Srinivasan, V., & Valk, R. (2008). Establishing a link between employee turnover and withdrawal behaviours: Application of data mining techniques. *Research and Practice in Human Resource Management, 16*, 81–99.

Neumeyer, X., Santos, S. C., & Morris, M. H. (2020). Overcoming barriers to technology adoption when fostering entrepreneurship among the poor: The role of technology and digital literacy. *IEEE Transactions on Engineering Management, 68*, 1605–1618.

Oberoi, S., Kumar, S., Sharma, R. K., & Gaur, L. (2022). Determinants of artificial intelligence systems and its impact on the performance of accounting firms. *Machine learning, advances in computing, renewable energy and communication* (pp. 411–427). Springer. https://doi.org/10.1007/978-981-16-2354-7_38

Oosthuizen, R. M., Coetzee, M., & Munro, Z. (2016). Work-life balance, job satisfaction and turnover intention amongst information technology employees. *Southern African Business Review, 20*, 446–467.

Oxborough C., Birchall, A., Cameron, E., Townsend, A., Rao, A., and Westermann, C. (2018). Explainable AI: Driving business value through greater understanding. Retrieved from PWC: Https://www.pwc.co.uk/audit-assurance/assets/explainable-ai.pdf.

Ozcan, S., Sakar, C. O., & Suloglu, M. (2020). Human resources mining for examination of R&D progress and requirements. *IEEE Transactions on Engineering Management, 68*, 1372–1387.

Park, I., Yoon, B., Kim, S., & Seol, H. (2019). Technological opportunities discovery for safety through topic modeling and opinion mining in the fourth industrial revolution: The case of artificial intelligence. *IEEE Transactions on Engineering Management, 68*, 1504–1519.

Prasman, V. J., & Ariyanto, E. (2020). Analyse the effect of work satisfaction and work environment on employee turnover intention at ecommerce industry. *IOSR Journal of Humanities and Social Science, 25*, 58–65.

Punnoose, R., & Ajit, P. (2016). Prediction of employee turnover in organizations using machine learning algorithms. *International Journal of Advanced Research in Artificial Intelligence, 5*(9). https://doi.org/10.14569/IJARAI.2016.050904

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). Why should I trust you? Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1135–1144). https://doi.org/10.1145/2939672.2939778

Ribeiro, M. T., Singh, S., & Guestrin, C. (2018). Anchors: High-precision model-agnostic explanations. In *Thirty-Second AAAI Conference on Artificial Intelligence* (pp. 1527–1535). AAAI.

Ribes, E., Touahri, K., & Perthame, B. (2017). Employee turnover prediction and retention policies design: A case study. https://doi.org/10.48550/arXiv.1707.01377.

Rybnicek, R., Bergner, S., & Gutschelhofer, A. (2019). How individual needs influence motivation effects: A neuroscientific study on McClelland's need theory. *Review of Managerial Science, 13*, 443–482. https://doi.org/10.1007/s11846-017-0252-1

Saeed, S., Haroon, H. B., Naqvi, M., Jhanjhi, N. Z., Ahmad, M., & Gaur, L. (2022). A systematic mapping study of low-grade tumor of brain cancer and CSF fluid detecting approaches and parameters. *Approaches and applications of deep learning in virtual medical care* (pp. 236–259). IGI Global. https://www.igi-global.com/chapter/a-systematic-mapping-study-of-low-grade-tumor-of-brain-cancer-and-csf-fluid-detecting-approaches-and-parameters/298113.

Saradhi, V. V., & Palshikar, G. K. (2011). Employee churn prediction. *Expert Systems with Applications, 38*, 1999–2006. https://doi.org/10.1016/j.eswa.2010.07.134

Sekaran, K., & Shanmugam, S. (2022). Interpreting the factors of employee attrition using explainable AI. In *2022 International Conference on Decision Aid Sciences and Applications* (pp. 932–936). IEEE.

Setiawan, I., Suprihant, S., Nugraha, A. C., & Hutahaean, J. (2020). HR Analytics: Employee attrition analysis using logistic regression. In *, 830. IOP Conference Series: Materials Science and Engineering*, Article 032001.

Shafique, U., & Qaiser, H. (2014). A comparative study of data mining process models (KDD, CRISP-DM and SEMMA). *International Journal of Innovation and Scientific Research, 12*(1), 217–222.

Siong, Z. M. B., Mellor, D., Moore, K. A., & Firth, L. (2006). Predicting intention to quit in the call centre industry: Does the retail model fit? *Journal of Managerial Psychology, 21*, 231–243.

Sisodia, D. S., Vishwakarma, S., & Pujahari, A. (2017). Evaluation of machine learning models for employee churn prediction. In *2017 International Conference on Inventive Computing and Informatics* (pp. 1016–1020). IEEE.

Srivastava, D. K., & Tiwari, P. K. (2020). An analysis report to reduce the employee attrition within organisations. *Journal of Discrete Mathematical Sciences and Cryptography, 23*, 337–348. https://doi.org/10.1080/09720529.2020.1721874

Stamolampros, P., Korfiatis, N., Chalvatzis, K., & Buhalis, D. (2019). Job satisfaction and employee turnover determinants in high contact services: Insights from employees' Online reviews. *Tourism Management, 75*, 130–147. https://doi.org/10.1016/j.tourman.2019.04.030

Stiglic, G., Kocbek, P., Fijacko, N., Zitnik, M., Verbert, K., & Cilar, L. (2020). Interpretability of machine learning-based prediction models in healthcare. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10*(5), e1379.

Storey, D. J. (2016). The growth of small firms. *Understanding the small business sector* (pp. 112–160). London: Taylor and Francis. https://doi.org/10.4324/9781315544335

Tao, Z., Wu, C., & Zhao, S. (2021). Research on the prediction of employee turnover behavior and its interpretability. In *Proceedings of the 2021 5th International Conference on Electronic Information Technology and Computer Engineering* (pp. 760–767). https://doi.org/10.1145/3501409.3501547

Tarallo, E., Akabane, G. K., Shimabukuro, C. I., Mello, J., & Amancio, D. (2019). Machine learning in predicting demand for fast-moving consumer goods: An exploratory research. *IFAC-Papers OnLine, 52*, 737–742. https://doi.org/10.1016/j.ifacol.2019.11.203

Tharani, S. M., & Raj, S. V. (2020). Predicting employee turnover intention in IT&ITeS industry using machine learning algorithms. In *2020 Fourth International Conference onI-SMAC(IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)* (pp. 508–513).

Thompson, S. C., Holmgren, A. J., & Ford, E. W. (2022). Information system use antecedents of nursing employee turnover in a hospital setting. *Health Care Management Review, 47*(1), 78–85.

Tsoka, T., Ye, X., Chen, Y., Gong, D., & Xia, X. (2022). Explainable artificial intelligence for building energy performance certificate labelling classification. *Journal of Cleaner Production, 355*, Article 131626.

Vignesh, S. G., Sarojini, V., & Vetrivel, S. (2018). Employee attrition and Employee retention-challenges & suggestions. In *International Conference on Economic Transformation with Inclusive Growth* (pp. 1–15).

Walsh, K., & Taylor, M. (2007). Developing in-house careers and retaining management talent: What hospitality professionals want from their jobs. *Cornell Hotel, Restaurant and Administrative Quarterly, 48*, 163–210. https://doi.org/10.1177/0010880407300521

Wasilefsky, D., Caballero, W. N., Johnstone, C., Gaw, N., & Jenkins, P. R. (2024). Responsible machine learning for United States Air Force pilot candidate selection. *Decision Support Systems, 180*, Article 114198.

Yan, X., & Su, X. (2009). *Linear regression analysis-theory and computing*. Singapore: World Scientific Publishing Co.

Yang, S., & Islam, M. T. (2020). *Employee attrition analysis*. Cornell University. https://doi.org/10.48550/arXiv.2012.01286. arXiv:1806.10480v1 [stat.ML].

Yedida, R., Reddy, R., Vahi, R., Jana, R., Abhilash, G.V., & Kulkarni, D. (2018). Employee attrition prediction. arXiv:1806.10480 [stat.ML]. https://doi.org/10.48550/arXiv.1806.10480.

Yin, C., Li, D., & Salmador, M. P. (2021). Institutional change of compensation policy and its impact on CEO turnover and firm performance. *Review of Managerial Science, 16*, 2527–2552. https://doi.org/10.1007/s11846-021-00507-3

Yingjie, L., Deng, S., & Pan, T. (2019). Does usage of enterprise social media affect employee turnover? Empirical evidence from Chinese companies. *Internet Research, 29*, 970–992. https://doi.org/10.1108/INTR-03-2018-0140

Yuan, L., Gao, X., Zheng, Z., Edmonds, M., Wu, Y. N., Rossano, F., Lu, H., Zhu, X., & Zhu, S. C. (2022). In situ bidirectional human-robot value alignment. *Science Robotics, 7*(68), eabm4183. https://doi.org/10.1126/scirobotics.abm4183

Zhang, Y. (2016). A review of employee turnover influence factor and countermeasure. *Journal of Human Resource and Sustainability Studies, 4*, 85–91. https://doi.org/10.4236/jhrss.2016.42010

Zhang, C. A., Cho, S., & Vasarhelyi, M. (2022). Explainable artificial intelligence (xai) in auditing. *International Journal of Accounting Information Systems, 46*, Article 100572. https://doi.org/10.2139/ssrn.3981918