World Journal of Advanced Research and Reviews

eISSN: 2581-9615 CODEN (USA): WJARAI

Cross Ref DOI: 10.30574/wjarr

Journal homepage: https://wjarr.com/

(RESEARCH ARTICLE)

Investigating employee attrition using machine learning techniques

Ida Godwin Ogah *

Department of Computer Science, Faculty of Applied Sciences, WSB University, Dąbrowa Górnicza, Poland.

World Journal of Advanced Research and Reviews, 2025, XX(XX), XXX–XXX

Publication history: Received on 03 April 2025; revised on 11 May 2025; accepted on 13 May 2025

Article DOI: xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx

Abstract

Introduction: This study investigates underlying issues that employees might not openly disclose in exit interviews by leveraging machine learning techniques to explore the factors causing employee turnover, offering insights beyond churn predictions and traditional exit interviews. The novelty of this research lies in the use of ML causal inference to draw conclusions.

Methods: The machine learning algorithm was trained on 10 features of the dataset with 14,999 records. The feature importance analysis and clustering highlighted the most influential factors in predicting attrition. Then, propensity score matching was used to estimate the causal effect of these features on attrition by comparing similar groups of employees who stayed and left.

Results: The model achieved an impressive accuracy of 95.25% and an F1-score of 96.0%, demonstrating the robustness of the algorithm. Further analysis, including clustering and causal inference using propensity score matching, revealed distinct patterns among departing employees, such as low, frustrated, and high performers.

Conclusion: By employing causal inference rather than merely prediction, this study offers a more objective understanding of the causes of attrition. The causal model in this research provided greater transparency into the decision-making process, allowing HR teams to visualize the factors driving attrition and make informed retention policies.

Keywords: Machine Learning; Data Science; Causal Inference; Data Analytics; Human Resources; Employee Retention

1. Introduction

Attrition, also referred to as employee turnover, presents a critical challenge for organizations globally. Employee attrition poses significant financial and operational challenges for many organizations worldwide, and traditional methods like exit interviews and machine learning predictions may not always reveal the true reasons for employees leaving (Eades, 2022). Moreover, some employees themselves may not even be fully conscious of all the factors contributing to their decision to quit the job. In exit interviews, quitting employees may also be hesitant to share their true reasons for leaving, especially if they involve negative feedback about the company or their managers. They might also fear repercussions or simply want to avoid conflict. However, machine learning models can go beyond predictions to identify patterns and relationships in data that might not be apparent to individuals. This study introduces a novel approach by leveraging ML causal inference to draw conclusions, potentially uncovering hidden issues that contribute to employee turnover. The objective of this study is to provide valuable insights for human resource (HR) professionals, enabling them to develop data-driven strategies that mitigate attrition and enhance employee retention policies.

* Corresponding author: Ida Godwin Ogah.

In 2021, the U.S. Bureau of Labor Statistics reported a record 4.5 million resignations in November alone, highlighting the urgency of addressing employee attrition and its causes (JOLTS, 2021). Traditional approaches, such as exit interviews, often fail to reveal the true reasons for employees' departure, as departing employees may withhold information. Thus, predicting attrition using machine learning (ML) models allows for objective insights based on employee data, bypassing subjective biases.

Machine learning's integration into human resource (HR) analytics is relatively new but rapidly growing. Research has shown that voluntary attrition is often linked to job dissatisfaction and unfulfilled career goals (Tae et al., 2008). As organizations strive for competitiveness, retaining employees and understanding the underlying factors leading to voluntary resignations have become imperative. This research aims to build on existing work by applying decision tree algorithms to predict attrition and identify the key variables affecting employee decisions to leave the host organization. The study contributes to the growing body of literature by focusing on practical applications in HR management and improving retention strategies through data-driven decision-making.

Key factors like employee satisfaction, evaluation scores, and time spent at the company are shown to strongly influence turnover (Pettman, 1973). By accurately identifying employees at risk of leaving, HR departments can proactively address these issues through tailored interventions, reducing turnover rates and minimizing operational disruptions.

This project explored predicting and understanding employee attrition using machine learning and statistical techniques, replacing traditional exit interviews. Decision trees identified key factors like satisfaction level, last evaluation, and time spent in the company as significant attrition drivers. Further analysis using clustering and propensity score matching helped to reveal patterns among employees who left, providing insights into potential causes, like low performance, frustration, or seeking better opportunities. This approach offers valuable data-driven insights into attrition, guiding interventions, and strategies for employee retention without relying on exit interviews.

This paper is divided into sections: Section 1 provides the background of the study. Section 2 presents a literature review of existing work in this domain. Section 3 details the methodological approach employed. Section 4 presents the results and discusses their implications. Finally, Section 5 offers conclusions and recommendations for future research and practice.

2. Literature review

The integration of machine learning into human resource (HR) analytics is a relatively new but rapidly growing field. Several studies have successfully demonstrated the use of the K-Nearest Neighbors (KNN) classifier in predicting employee attrition (Yedida et al., 2018), with a focus on model performance and prediction accuracy.

According to research published in the Harvard Business Review, traditional evidence-based approaches to identifying the causes and nature of attrition have limitations compared to machine learning algorithms (Klotz, 2019). HR professionals and managers often rely on exit interviews and feedback from colleagues who were closest to the departing employee to find the cause of attrition, however, research indicates that many employees do not disclose their true reasons for leaving during exit interviews (König et al., 2022). Richard et al. (2021) suggest using decision tree algorithms for classification and turnover prediction, but the research does not address causality.

By leveraging machine learning techniques such as decision tree feature importance models, a data-driven approach can be implemented to identify key factors influencing attrition and establish causative relationships. Machine learning offers a robust alternative to traditional qualitative and quantitative methods, particularly in handling complexity and confidentiality (Binoy et al., 2010). These algorithms can outperform conventional statistical approaches while mitigating biases introduced by employees withholding information. With decision trees and expert knowledge, organizations can effectively analyze employee turnover without relying solely on exit interviews.

2.1. Employee retention statistics and trends

Employee retention is a concern for almost every firm, according to statistics from around the world. Even the most successful CEOs have struggled to keep their top performers. Every manager should make every effort to keep people on board. When you view the numbers below, you'll realize the importance of this. As contained in Employeepedia (2017), the following statistics were gathered about attrition and retention in recent years:

- One-third of new hires quit their jobs within the first six months. This point is a crucial stage for every company as they need to intensify their retention process.

- 73% of organizations are constantly revamping their onboarding processes to improve employee retention.
- 45% of referred employees will leave after two years. Improving your referral program can help you keep the staff for longer.
- 78% of managers value employee retention. These are the companies that include employee retention in their budgets way before they hire.
- 33% of recruits knew whether they were going to stay for the long-term or short-term within their first week.
- 50% of remote workers are less likely to quit as they are more satisfied with their working conditions. They can work at their own pace in an environment in which they feel comfortable. Remote work is an added advantage to companies as they will not have to work hard to retain their employees.
- 35% of employees will look for new positions if they don't get a pay rise in the next 12 months. If your company provides no salary increments or bonuses, expect some possible retention problems. All employees, besides the newbies, expect higher salaries to match their productivity.
- 9.32% of employers expect their employees to job-hop. This will act as mental preparation for most employers who fail at retaining employees for an extended period. It is wise to keep such trends at the back of your mind, at the same time you find out the cause.
- 33% of supervisors and managers are looking for new opportunities. No one in the company is immune to leaving, including the senior leadership. The leadership will move if they feel undervalued by your organization. Guard your business so it does not get to this point.

2.2. The Dynamics of Employee Attrition

A researcher once said that "the positive and negative effects of employees coming and going are two sides of the same coin - employee turnover (Mayhew, 2019). A good understanding of the factors that influence an employee's decision to quit can help the organization better position itself. An employee's decision to quit can have either a positive or a negative impact on the company. Most managers make the mistake of believing that no employee is indispensable (Navlani, 2018). Of course, no employee is indispensable. However, they neglect the losses that the company suffered during this time. Employee turnover is more than just the annual percentage of employees who left and those who stayed. An employee who quits in a negative sense promotes a bad image for the company. In addition, there are the costs of hiring and training new employees, low production due to labor shortages, lost sales, and a bad reputation for the company.

While churn predictions have been widely used in attrition studies, research suggests they often fail to capture the true reasons behind employee departures. To address this limitation, our study applies causal inference, allowing for conclusions beyond predictions and self-reported data during exit interviews, which may be biased.

3. Methodology

This study employs causal inference techniques to analyse attrition patterns, enhancing employee churn predictions and eliminating the need for exit interviews, which are often unreliable.

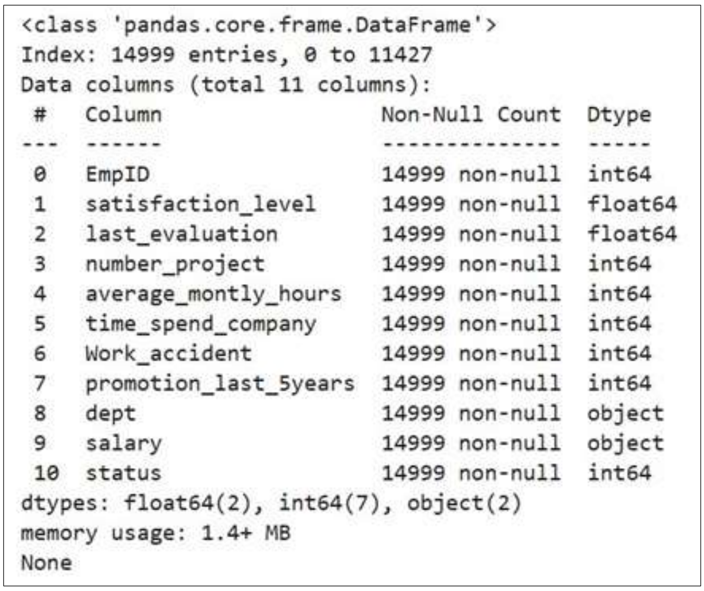
This project employed a mixed-method approach to analyse employee attrition using a combination of causal inferences and a machine learning algorithm. The algorithm was trained on 10 features of the dataset with 14,999 records. The feature importance analysis was used to identify the most influential factors in predicting attrition. To further understand the underlying causes, K-means clustering was applied to group employees who left based on their satisfaction level and last evaluation, revealing distinct clusters like low performers, frustrated performers, and high performers. Finally, propensity score matching was used to estimate the causal effect of these features on attrition by comparing similar groups of employees who stayed and left.

This research offers an innovative solution compared to traditional methods for understanding employee attrition, such as employee churn prediction, which offers no insight into the causes, and exit interviews, which often rely on self-reported data and may not fully capture the actual underlying reasons for employees’ departure.

3.1. Data Source and Transformation

The dataset, sourced from an HR dataset from Kaggle, comprises 14,999 employees, of which 24% had voluntarily left (Ramin, 2021). This dataset included features such as employee satisfaction, last evaluation scores, average monthly working hours, and salary levels, among others. The data preprocessing steps to ensure the model's effectiveness involve cleaning, transforming categorical variables using one-hot encoding, and scaling continuous variables, among others. Additionally, dummy variables were created to represent categorical data like department and salary categories.

These details are explained in the subsequent sections. Figure 1.0 below represents the overview of the historical HR dataset.



```
<class 'pandas.core.frame.DataFrame'>
Index: 14999 entries, 0 to 11427
Data columns (total 11 columns):
 #   Column                 Non-Null Count  Dtype
---  ------                 --------------  -----
 0   EmpID                  14999 non-null  int64
 1   satisfaction_level     14999 non-null  float64
 2   last_evaluation        14999 non-null  float64
 3   number_project         14999 non-null  int64
 4   average_montly_hours   14999 non-null  int64
 5   time_spend_company     14999 non-null  int64
 6   Work_accident          14999 non-null  int64
 7   promotion_last_5years  14999 non-null  int64
 8   dept                   14999 non-null  object
 9   salary                 14999 non-null  object
 10  status                 14999 non-null  int64
dtypes: float64(2), int64(7), object(2)
memory usage: 1.4+ MB
None
```

Figure 1 Summary of employee dataset used in this study. Source: Kaggle, 2021

To prevent data leakage, the dataset was meticulously split into training and testing sets before any data preprocessing or feature engineering steps were applied. This ensured that information from the testing set did not influence the model's training and that the evaluation results accurately reflected its performance on unseen data. Careful attention was paid to avoid including features that implicitly contained information about the target variable in the testing set, thereby further mitigating the risk of data leakage (Shachar et al., 2011).

3.2. Feature Engineering

3.2.1. Mean Imputation

In order to prevent data loss and improve model robustness, the missing values in the dataset are filled using the mean substitution. In statistics, mean imputation is a method where missing values for a particular variable are replaced with the mean of the observed values for that variable (Lin et al., 2020). This approach addresses individual missing values (not entire records) and is often used when only some components of a dataset are missing (Waljee et al., 2013).

$$X = \frac{\sum_{I} n = 1^{xi}}{n}$$

Were

- x′ is the imputed value,
- xi are the known values in the feature (column),
- n is the number of observed (non-missing) values in the column.

3.2.2. Dummy Trapping and One-Hot Encoding

To avoid dummy trapping or multicollinearity of categorical data, one category of the department variable was removed as a reference point. The dropped category becomes the baseline or reference level against which the other categories are compared. This improves stability and interpretation during model training. This ensures that the model can interpret the department information correctly without any issues. This transformation allows the model to utilize the department information for predictions. Multicollinearity is a statistical issue that arises when two or more independent variables in a decision tree model exhibit a high degree of correlation, signifying a strong linear relationship between the predictor variables (Chan et al., 2022).

One-hot encoding was used based on the idea of indicator variables in statistics, representing categorical data as binary vectors to avoid unintended ordinal relationships among categories (Powers, 2008).

$$\text{one} - \text{hot(Ci)} = \begin{cases} 1 \text{ if the value belongs to Ci} \\ 0 \text{ for others} \end{cases}$$

This is a transformation where each unique category (Ci) in the categorical feature becomes a new binary feature, represented as a vector. To regulate high cardinality due to the explosion in the number of dimensions associated with one-hot encoding, different strategies were deployed to reduce the impact on memory, computational efficiency, and model performance.

- Feature Importance Analysis: Before applying one-hot encoding, we analyze important categories using feature importance scores from models. This removes irrelevant or redundant categories, reducing dimensionality.
- Hybrid Encoding: Combine one-hot encoding for low-cardinality variables and target encoding or frequency-based thresholding for high-cardinality variables. This helps to balance the trade-off between interpretability and computational efficiency.
- Principal Component Analysis (PCA): Transform the high-dimensional binary data into a smaller set of principal components while retaining most variance. This helps to maintain relationships between categories in the reduced dimensions.
- Clustering Similar Categories: This method groups similar categories based on data-driven similarity metrics by K-means clustering.

3.2.3. Min-Max Scaling (Normalization):

The dataset was rescaled to a specific range between [0, 1] to ensure consistency in some features and enhance model performance. The formula for min-max scaling is:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)}$$

Where

- x is the original value,
- min(x) is the minimum value of the feature,
- max(x) is the maximum value of the feature,
- x' is the scaled value within the range [0, 1]

3.3. Decision Tree Model Training

The decision tree algorithm was selected due to its interpretability and ability to handle non-linear relationships. A decision tree's purpose is to divide the training set into homogeneous zones with only one iris species present based on the features provided, in this case, petal and sepal widths. The tree is created iteratively from the root to the last leaf.

The decision tree model implementation was done using the sci-kit-learn library in Python. The features and targets selected for model training were based on feature importance calculation by the decision tree classifier. The features with high importance scores were detected as major contributors to the predictive model. These features included satisfaction level, last evaluation, tenure in the company, work accident, number of projects, average monthly hours, department, and salary.

To prepare the data for machine learning, it is necessary to clearly define the target variable ("stayed_or_left") and the feature variables (all other columns). This separation is crucial for training and evaluating the model's performance in predicting employee attrition based on the provided features (Sharma, 2012). The decision tree model was trained on the prepared training data, which consisted of the selected features (features_train) and corresponding target variables (target_train), while the variables for testing (features_test) and (target_test) respectively.

- Target Variable: The "status" column was selected as the target variable because it's the variable the model is trying to predict.
- Feature Variables: The remaining columns (except "status") were selected as the features used to make predictions about the target variable. These features include employee satisfaction, salary, department, etc.

- Model training: By splitting the data into training and testing sets (target_train, target_test, features_train, features_test), the model was trained on one portion of the data and then evaluated its performance on a separate, unseen portion. This helps us understand how well the model is likely to generalize to new data.
- Fitting Process: The fit method of the DecisionTreeClassifier object was called, passing the training data as arguments. This process builds the decision tree structure by recursively splitting the data based on the selected features and the Gini impurity criterion (default).

3.4. Gini Impurity

This study uses Gini Impurity as the splitting criterion when building the decision tree model using DecisionTreeClassifier. This helps the algorithm to create a tree that effectively separates the data and makes accurate predictions. It reduces the likelihood of incorrect classification of a randomly chosen node and labels according to the distribution of labels in the node. The Gini impurity G for a node is calculated as:

$$G \ = 1 - \sum_{i=1}^{c} p^2 \, \mathrm{i}$$

Here, C represents the total number of classes, and Pi denotes the proportion of items in the node that belong to class i. The Gini impurity G ranges from 0 (indicating perfect purity, where all elements belong to a single class) to 0.5 (representing maximum impurity in binary classification).

The following assumptions were made while working with the decision tree algorithm.

- In the beginning, the whole training set is considered as the root.
- Feature values are preferred to be categorical.
- If the values are non-categorical, they are discretized or converted to dummy variables before building the model.
- Records are distributed recursively based on attribute values.

Order to place attributes as root or internal nodes of the tree is done by using a statistical approach.

3.5. Hyperparameter Tuning with GRIDSEARCHCV

Hyperparameter tuning is crucial for optimizing machine learning models and finding the best settings that yield the best performance. The study employs GridSearchCV from the sklearn.model_selection module for this purpose. The Cross-validation (CV) technique is used to evaluate the model's performance by splitting the dataset into training and validation sets multiple times. In k-fold cross-validation (CV), where k is the number of folds, the dataset was split into k equal parts or subsets. The model is trained k times, each time using k−1 folds for training and the remaining fold for validation. The average cross-validation score was calculated by

$$CV_{score} \ = \frac{1}{k} \sum_{i\,=\,1}^{k} score_i$$

The grid search systematically searches through a predefined set of hyper-parameters to find the best combination for a model. In this study, a five-fold grid search cross-validation was employed to fine-tune hyper-parameters, helping to prevent overfitting and enhance model generalizability. The cross-validation score was calculated as the average performance across the five folds, given by

$$CV_{score} \ = \frac{1}{5} \sum_{i\,=\,1}^{5} score_i$$

After the search is complete, the best_params_ attribute of the param_search object stores the combination of hyper-parameters that produce the highest performance on the validation data.

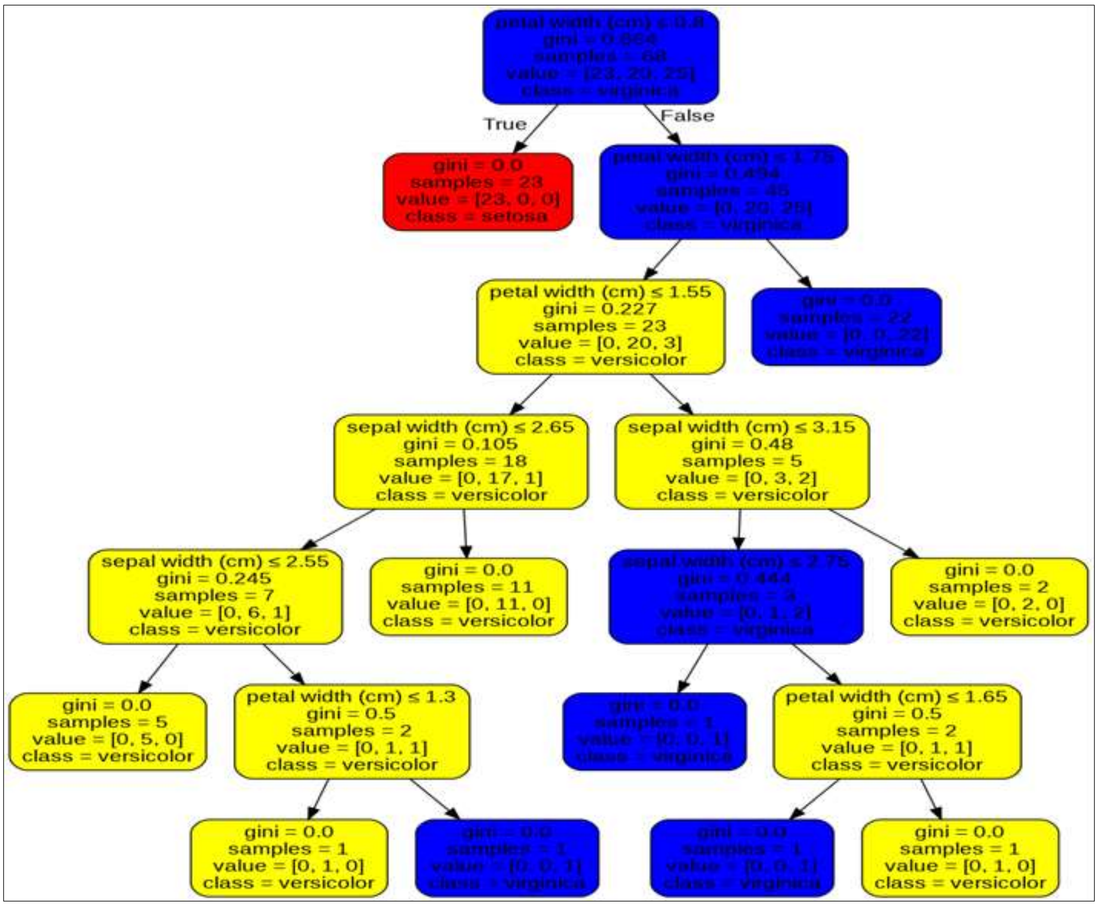


Figure 2 A typical decision tree split using the Gini impurity criterion

3.6. Model Generalization

The goal of machine learning is to build models that can generalize well to new, unseen data. The train-test split helps us assess the model's ability to generalize. Grid search uses cross-validation to evaluate each combination of hyper-parameters. For each set of hyper-parameters, the model was trained and validated using cross-validation. This ensures that the chosen hyper-parameters generalize well to unseen data.

By tuning the hyper-parameters using GridSearchCV, the study aims to improve the model's ability to generalize well to unseen data (avoid overfitting) and enhance its performance on the minority class (fixing imbalance). The resulting best parameters guide the creation of a final model that is expected to achieve better overall accuracy, recall, and ROC/AUC scores.

3.7. Overfitting Control and Hyper-parameter Tuning

Overfitting happens when a model learns the training data too well, including its noise and random fluctuations, and performs poorly on unseen or new data. The study addresses overfitting using the following techniques:

- Limiting Max Tree Depth: Setting max_depth in the DecisionTreeClassifier prevents the tree from growing too deep and becoming overly complex. This helps to generalize the model better to unseen or new data. The study experiments with different max_depth values (5 to 20) to find the optimal one.
- Limiting Min Sample Size on a Leaf: Setting min_samples_leaf ensures that a leaf node has a minimum number of samples. This prevents the tree from creating leaves that are too specific to individual training instances, thereby reducing overfitting. The study explores min_samples_leaf values ranging from 50 to 500 to identify the minimum sample size.
- Grid Search: The study uses GridSearchCV to systematically search for the best combination of max_depth and min_samples_leaf values. This helps to fine-tune the model and avoid overfitting. The best parameters were found to be max_depth of 6 and min_samples_leaf of 50.

3.8. Balancing the Class Imbalance

Class imbalance occurs when one class has significantly more instances than another. This can lead to a model that is biased towards the majority class. This study handles imbalance using the following techniques:

- Class Weighting: The study uses the balanced class_weight parameter in the DecisionTreeClassifier. This automatically adjusts the weights of the classes during training, giving more importance to the minority class. This helps the model to learn the patterns of the minority class better.
- Recall and ROC/AUC Scores: The study uses recall and ROC/AUC scores to evaluate the model's performance on the minority class. These metrics are more sensitive to class imbalance than accuracy. Focusing on these metrics ensures that the model is not biased towards the majority class.
- Comparing Balanced and Imbalanced Models: The study compares the performance of balanced and imbalanced models using recall and ROC/AUC scores. This helps to see the impact of class weighting on the model's performance. The results show that the balanced model performs better on the minority class, as expected.

3.9. Model Performance Evaluation using Confusion Matrix

The confusion matrix provides a detailed breakdown of the model's performance across true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). Using these values, we can interpret how the model performs for each class.

3.9.1. Interpretation

- True Positives (TP): Instances correctly predicted as positive.
- True Negatives (TN): Instances correctly predicted as negative.
- False Positives (FP): Instances incorrectly predicted as positive (Type I error).
- False Negatives (FN): Instances incorrectly predicted as negative (Type II error).

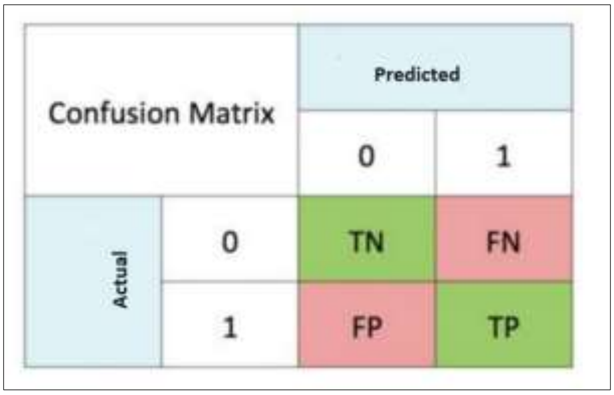
$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+ FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+ FN}}$$

$$\text{F1-Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision + Recall}}$$

In Figure 3.0 below, the confusion matrix used to evaluate predictions was illustrated, showing true positives and false positives clearly defined.



Confusion Matrix		Predicted	
		0	1
Actual	0	TN	FN
	1	FP	TP

Figure 3 Confusion matrix used to evaluate model predictions. Source: Author's work

From the figure above, our focus is on True-Positive (TP) is defined as the number of employees who actually left and were correctly labeled as left. False-Positive (FP) is defined as the number of employees who stayed but were wrongly labeled as left.

3.10. Methodology for Causal Inference

We use the propensity score to estimate the likelihood of an individual receiving the treatment based on their characteristics. The higher the propensity score, the more likely they are to be in the treatment group. By matching individuals with similar propensity scores, we aim to create comparable groups, reducing the influence of confounding factors and allowing for a more accurate estimation of the treatment's causal effect.

3.10.1. Mathematically

$P(X) = Pr(T = 1 \mid X)$

Where:

- P(X) = Propensity score, that is predicted probability of receiving treatment.
- Pr = Probability function
- T = Treatment status (1 = treated = left, 0 = control = stayed)
- X = Observed covariates (confounding factors)

4. Results and Discussion

This study successfully implements a machine learning model that uncovers hidden reasons for employee attrition. By analyzing employee data and identifying key predictors, the decision tree model achieved high accuracy in predicting factors leading to attrition. This suggests that the model can capture underlying issues that contribute to employees' decisions to leave, even when those factors are not explicitly stated by the departing employees. The findings of this study have important implications for HR professionals, enabling them to develop more targeted and effective retention strategies that address the root causes of attrition, rather than relying solely on potentially biased or incomplete information obtained through exit interviews.

The model demonstrated high performance and accuracy of 95.56%. The model successfully identified key factors such as satisfaction level, last evaluation, and time spent at the company as key factors influencing employees' decisions to quit their jobs.

4.1. Descriptive Statistics Summary

The dataset comprised 14,999 employees, with an attrition rate of 24%. Table 1.0 highlights the overall distribution of the employee dataset, while Figure 4.0 illustrates the proportion of employees who stayed with the company versus those who departed.

Table 1 The key features exhibited by the dataset

	count	mean	std	min	25%	50%	75%	max
satisfaction_level	14999	0.612834	0.248631	0.09	0.44	0.64	0.82	1
last_evaluation	14999	0.716102	0.171169	0.36	0.56	0.72	0.87	1
number_project	14999	3.803054	1.232592	2	3	4	5	7
average_montly_hours	14999	201.0503	49.9431	96	156	200	245	310
time_spend_company	14999	3.498233	1.460136	2	3	3	4	10
Work_accident	14999	0.14461	0.351719	0	0	0	0	1
promotion_last_5years	14999	0.021268	0.144281	0	0	0	0	1
stayed_or_left	14999	0.238083	0.425924	0	0	0	0	1

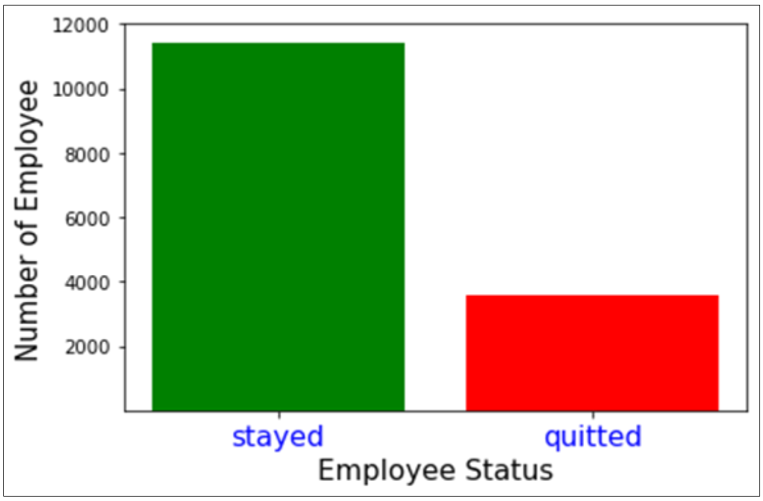


Figure 4 Distribution of employees. Source: Author’s work

- Satisfaction Level: The Average satisfaction score was noted with variation among employees.
- Last Evaluation: The scores were generally distributed across the performance spectrum.
- Time Spent at the Company: The average tenure was around 3-4 years, with some employees having longer or shorter durations.
- Other Features: The number of projects, average monthly hours, work accidents, and promotions showed variations across the dataset.

4.2. Correlation Analysis of the Variables

By analyzing the heatmap visual representation of the correlations among the variables in the employee dataset, we were able to uncover potential relationships and interactions between key factors. This not only provided valuable insights into the variables that might influence employee attrition but also highlighted dependencies (Gogtay, 2017). Understanding these correlations is crucial for effective feature selection, allowing us to refine the model and enhance its ability to predict employee churn with greater accuracy. This approach plays a critical role in optimizing the prediction model by focusing on the most impactful features (see Figure 5.0).

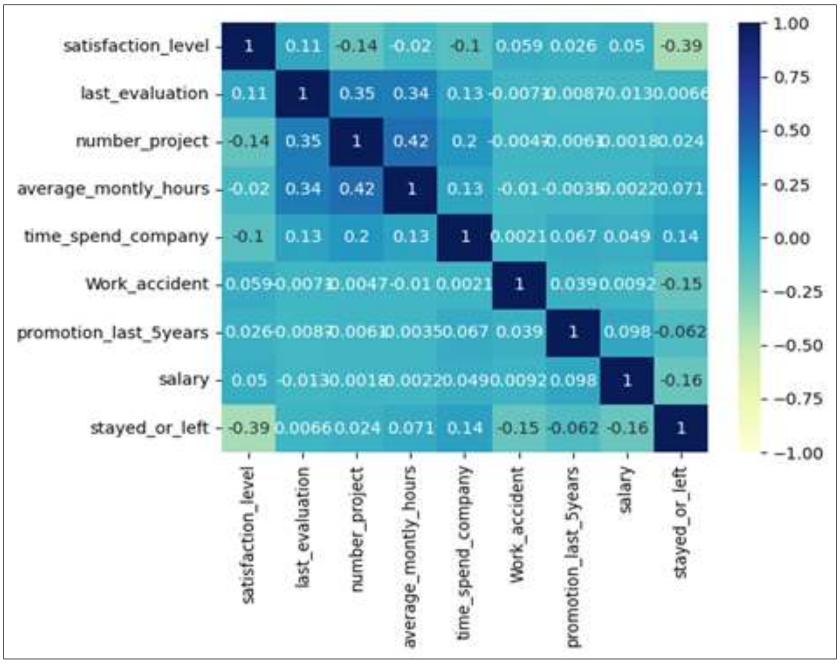


Figure 5 Correlation between the variables. Source: Author’s work

Generally speaking, the color bar on the right side gives a sense of how correlated the variables are. Dark blue represents a perfectly negative correlation. White shows that there is no correlation. Red shows a perfectly positive correlation. So, we basically look for the darkest colors to find potential relationships between variables.

- A strong negative correlation between the target variable and satisfaction_level means that as the satisfaction level decreases, the likelihood of an employee leaving increases.
- A moderate positive correlation between the target variable and time_spend_company suggests that employees who have spent more time with the company are slightly more likely to leave.
- A moderate negative correlation between the target variable and salary might indicate that employees with higher salaries are less likely to leave.
- A weak positive correlation between the target variable and last_evaluation might indicate that employees with better last evaluations could be slightly more likely to leave.
- A low correlation or no correlation between all other variables means they have minimal or no effect on employee leaving.

Subsequent sections will dive deeper for a better understanding of how these features affect attrition.

4.3. Evaluation of Decision Tree Model

To assess the model's robustness and performance, the study utilizes a confusion matrix and other metrics such as accuracy, recall, precision, roc/auc score, and F1-score to evaluate the model on both the training and test datasets.

The confusion matrix provides insights into the model's performance on the data it was trained on. It shows the number of true positives, true negatives, false positives, and false negatives, allowing us to assess the model's ability to correctly classify employees who stayed and those who departed within the training data. This matrix helps evaluate the model's learning and fitting to the training data (Figure 6.0).

The evaluation of the model using the test dataset also shows the model's generalization ability on unseen data (Figure 7.0). It reveals how well the model performs on data it has not encountered during training. By comparing the confusion matrices for the trained and test sets, we identify potential overfitting issues and adequately address them.

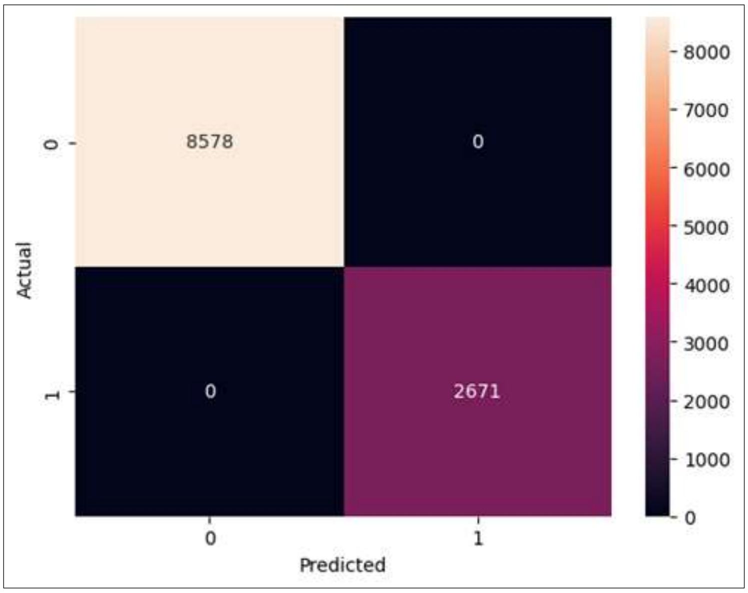


Figure 6 Evaluating the model performance on the training set

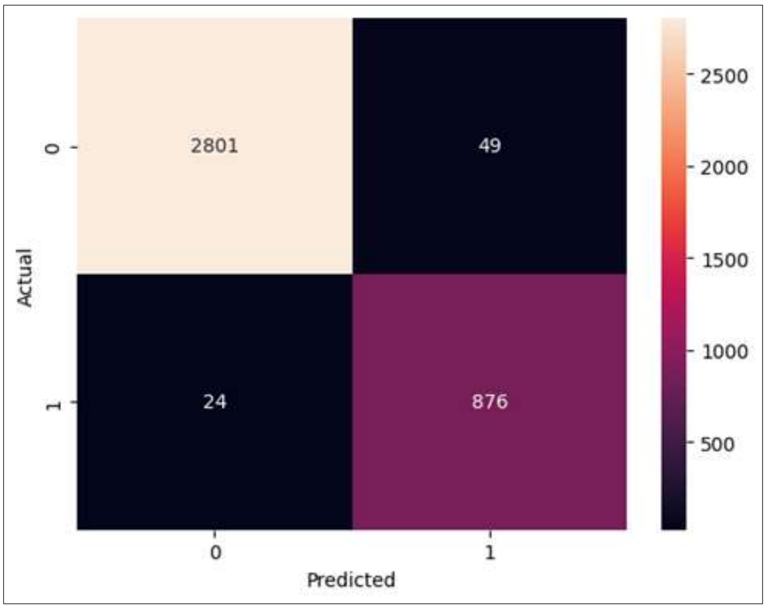


Figure 7 Evaluating the model performance on the test set (unseen data)

The F1-score provides a harmonic mean of precision and recall, balancing the trade-off between the two. It is useful because of the imbalance in the class distribution, particularly when both false positives and false negatives are important.

Table 2.0 showed that the decision tree algorithm achieved a remarkable accuracy of 95.25%, a recall score of 92.33%, a precision of 94.70%, an F1-score of 96.0%, and the Area Under the Receiver Operating Characteristic Curve (ROC/AUC) score of 94.24% further affirmed the model's effectiveness in predicting the likelihood of an employee staying or leaving.

Table 2 Model Performance Metrics

Metric	**Value**
Accuracy	95.25%
Recall	92.33%
Precision	94.70%
F1-score	96.00%
ROC/AUC score	94.25%

4.4. Model Feature Importance

Feature Importance scores help identify which features are contributing most to the decision-making within the model. The algorithm identified the most influential features for predicting employee turnover, ranked by their importance: satisfaction level (54.68%), time spent in the company (16.24%), last evaluation (8.73%), number of projects completed (13.36%), number of projects (10.43%), and average monthly hours (5.23%). These features emerged as the most critical predictors of employee turnover. The model was trained using 75% of the dataset, reserving the remaining 25% for testing. Table 3.0 below presents the features and their corresponding importance values, sorted by descending order of significance.

Table 3 Importance values of features used in the decision tree model

Feature	Importance	Percentage
satisfaction_level	0.546798	54.679789
time_spend_company	0.162368	16.236832
last_evaluation	0.133584	13.358394
number_project	0.104315	10.431482
average_montly_hours	0.052343	5.234260
salary_low	0.000592	0.059243
dept_marketing	0.000000	0.000000
dept_technical	0.000000	0.000000
dept_support	0.000000	0.000000
dept_sales	0.000000	0.000000
dept_product_mng	0.000000	0.000000
depths	0.000000	0.000000
dept_management	0.000000	0.000000
dept_accounting	0.000000	0.000000
dept_RandD	0.000000	0.000000
promotion_last_5years	0.000000	0.000000
Work_accident	0.000000	0.000000
salary_medium	0.000000	0.000000

One notable finding of this study is the zero-importance score assigned to the salary variable in the decision tree model. This suggests that salary, as represented in the dataset, did not have a significant impact on the model prediction. This result is somewhat unexpected, as salary is often considered a key factor in job satisfaction and retention. However, several factors could explain this observation. It is possible that other factors, such as satisfaction level, work-life balance, or career development opportunities, were more influential in driving attrition within this specific dataset. Additionally, the limited variation in salary or its correlation with other features could have contributed to its low importance score.

Despite the zero-importance score, it is important to acknowledge that salary could still play a role in attrition for some employees. Individual circumstances, career aspirations, and personal financial situations can influence the perceived importance of salary. While the model did not identify salary as a primary driver of attrition overall, it could be a significant factor for specific segments of the workforce. Further investigation is needed to explore this possibility. It is also worth considering that the way salary was represented in the dataset might have limited the model's ability to capture its full impact. Future research could explore more nuanced measures of compensation, such as total rewards packages or perceived salary fairness, to gain a deeper understanding of its relationship with attrition.

4.5. Key Insights from clustering analysis

Using K-means clustering to plot a graph of quitters against satisfaction and last evaluation, the algorithm identified three distinct clusters among employees who left the company, providing further insights into the characteristics of churned employees. The analysis revealed that employee satisfaction was the strongest determinant of attrition. Employees with lower satisfaction scores were more likely to leave, particularly when combined with low salaries and limited promotion opportunities. The second most important predictor was the last evaluation score, with employees receiving higher evaluations showing a tendency to seek new challenges outside the company. Table 4.0 and Figure 8.0 illustrate insights from the clustering analysis.

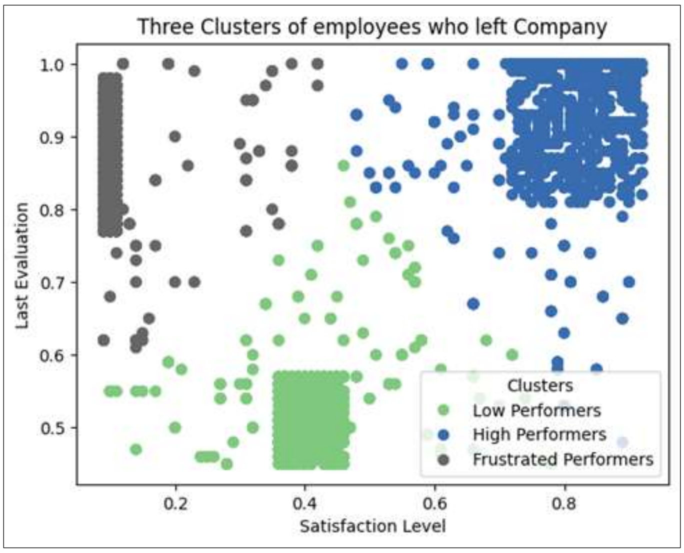


Figure 8 Clusters Analysis of departed employees

Table 4 Insights from KMeans Clustering Analysis

Cluster	**Condition**	**Insights**
Low Performers	Moderate Satisfaction, Low Evaluation	Employees are likely disengaged or mismatched in their roles
High Performers	High Satisfaction, High Evaluation	Employees who may seek growth opportunities elsewhere
Frustrated Performers	Low Satisfaction, High Evaluation	Employees who are performing well or overworked but feel unrecognized

The cluster analysis revealed a surprising trend with the seemingly satisfied employees leaving the company. That is, employees with high satisfaction and high evaluation (high-performers) also left the company. This highlights the limitations of relying solely on exit interviews, which may not capture the full picture. These employees, despite high evaluations and satisfaction levels, could be leaving due to hidden factors such as limited growth opportunities, low compensation, or a desire for a better work-life balance. This emphasizes the need for proactive measures like talent development programs, competitive benefits, and regular employee feedback to address possible attrition risks before they lead to quitting of treasured employees.

Combining feature importance with cluster analysis has enhanced our understanding of how features influence the prediction within different groups or segments, providing a more detailed and context-specific interpretation of the data. However, while this combination can uncover patterns and relationships, it doesn't establish full causality. To fully understand the causal attribution, we would need to integrate additional causal modeling techniques such as Propensity Score Matching.

4.6. Causal Inference using Propensity Score Matching

Propensity score matching helps to reduce bias and estimate causal effects, but it's important to remember that it is not a foolproof method. There might be other unobserved factors influencing both attrition and the outcomes. However, the results are specific to the dataset and the population it represents. They might not be generalizable to other companies or industries.

The results suggest that satisfaction level is a more important factor in employee attrition compared to the last evaluation score. Employees who left the company tended to have lower satisfaction levels, while the last evaluation scores were relatively similar between the two groups (Table 5.0).

Table 5 Estimated Causal Effect on Attrition

Feature	Effect on Attrition (Stayed vs. Left)	(Stayed, Left)	Insight
Satisfaction Level	Employees who stayed had a slightly higher average satisfaction level compared to those who left.	(0.468, 0.440)	This suggests that lower satisfaction levels might be causally linked to increased attrition, though the difference is relatively small.
Last Evaluation	Employees who stayed had a slightly higher average last evaluation score compared to those who left.	(0.725, 0.718)	This indicates a potential, but weak, causal relationship between lower evaluation scores and attrition.

The novelty of this research is the use of causal inference to draw conclusions without the need to conduct exit interviews, which, in most cases, are not honest representations of the key issues leading to attrition. The contributions of this study align with existing literature on attrition predictors. Similar studies, like those by Sainju (2021) and Memon (2021), emphasized the importance of satisfaction and performance evaluations in predicting turnover. However, the causal model in this research provided greater transparency into the decision-making process, allowing HR teams to visualize the factors driving attrition and make informed retention policies.

5. Conclusion and Recommendations

This study provides a more objective understanding of attrition factors by leveraging causal inference techniques, a decision tree model, and clustering analysis. The results revealed that satisfaction level, last evaluation, and company tenure were key factors influencing employee attrition. Propensity score matching provided evidence for the causal effect of lower satisfaction levels on increased attrition.

While exit interviews can provide valuable insights, they often rely on employees' willingness to share their true reasons for leaving. Employees may be hesitant to provide negative feedback or may not be fully aware of all the factors influencing their decision. The data-driven approach employed in this study offers a more objective way to understand attrition by identifying patterns and relationships in employee data. This finding underscores the importance of addressing employee satisfaction as a crucial driver of retention. Organizations should invest in initiatives to improve employee satisfaction, such as enhanced communication, work-life balance programs, and recognition and reward systems.

The practical implications of this study are significant for companies that struggle with high employee turnover. HR departments can utilize this model to prioritize interventions for at-risk employees, particularly those in the frustrated performers' cluster. Retention strategies such as revising promotion policies, providing more challenging projects, and offering better recognition programs could potentially retain high performers.

While the study identified important factors associated with attrition, it relied primarily on internal company data, potentially neglecting crucial external factors that might influence employee decisions. External factors, such as industry trends, economic conditions, and competitive job opportunities, could significantly impact attrition rates but were not incorporated into this analysis. This limitation hinders a comprehensive understanding of employee attrition and limits the generalizability of the findings to broader contexts.

Compliance with ethical standards

Acknowledgments

I would like to express my sincere gratitude to Prof. Adam Domanski for his valuable guidance, insightful feedback, and continuous support throughout this research. His expertise and encouragement have been instrumental in shaping this work.

Disclosure of conflict of interest

I declare that there are no financial, personal, or professional conflicts of interest that could have influenced the work reported in this paper.

Statement of ethical approval

This research utilizes publicly available anonymized datasets for model development, ensuring compliance with ethical standards and research integrity. As such, it does not require special approval or licensing.

Declaration of Interest Statement

There is no conflict of interest related to this research. There is no competing financial, professional, or personal interests that could have influenced the conduct or findings of this research. This study was carried out independently, with no external funding or conflicts of interest.

Funding sources

This research did not receive any specific grant from any funding agencies in the public, commercial, or not-for-profit sectors.

Statement of Contribution

This research was conducted independently, with all conceptualization, data collection, analysis, and interpretation carried out solely by me. My supervisor offered guidance and advice as needed, ensuring academic rigor and full autonomy in executing the study.

References

[1] Eades, C. (2022). Using exit interviews to enhance police employee retention and hiring. Saint Leo University.

[2] U.S. Bureau of Labor Statistics. (2021). Employee turnover statistics. Job Openings and Labor Turnover Survey (JOLTS). Retrieved from https://www.bls.gov/jlt.

[3] Tae, H. L. et al., (2008). Understanding Voluntary Turnover: Path-Specific Job Satisfaction Effects and The Importance of Unsolicited Job Offers. Retrieved from https://doi.org/10.5465/amr.2008.33665124

[4] Yedida, R. (2018). Employee attrition prediction using machine learning. Journal of Human Resources Analytics, 13(2), 124-135.

[5] Klotz, A. C., and Bolino, M. C. (2019). Do you really know why employees leave your company? Retrieved from Harvard Business Review. https://hbr.org/2019/07/do-you-really-know-why-employees-leave-your-company.

[6] Mayhew Ruth (2019), Employee Turnover Definitions and Calculations

[7] Employeepedia (2017). Statistics On Employee Retention.

[8] https://www.employeepedia.com/manage/retention/994-employee-retention-theories

[9] Richard, J., Shreyas, U., and Sanket, J. (2021). Employee attrition using machine learning and depression analysis. IEEE.

[10] Pettman, B. O. (1973). Some factors influencing labour turnover: A review of research literature. Industrial Relations Journal, 4(3), 43-61.

[11] Navlani A. (2018). Decision tree classification in Python tutorial. https://www.datacamp.com/community/tutorials/decision-tree-classification-python.

[12] Binoy B, et al., (2010). A Genetic Algorithm Optimized Decision Tree-SVM-based Stock Market Trend Prediction System

[13] Guru, V., and Suresh, K. (2018). Employee attrition and employee retention:

[14] Challenges and suggestions. ResearchGate. Retrieved from https://www.researchgate.net/publication/322896996.

[15] Scikit-learn Decision Trees User Guide. Retrieved from https://scikit-learn.org/stable/modules/tree.html

[16] Richer Valentin (2019). Understanding Decision Trees.

[17] Konig, C. J., Richter, M., and Isak, I. (2022). Exit interviews as a tool to reduce parting employees' complaints about their former employer and to ensure residual commitment. Management research review, 45(3), 381-397.

[18] Ramin Huseyn (2021) HR Analytics Data Set. Kaggle. https://www.kaggle.com/datasets/raminhuseyn/hr-analytics-data-set

[19] Shachar K., Saharon R., Claudia P. (2011). Leakage in data mining: Formulation, detection, and avoidance.

[20] Chan, J. Y. L., Leow, S. M. H., Bea, K. T., Cheng, W. K., Phoong, S. W., Hong, Z. W., and Chen, Y. L. (2022). Mitigating the multicollinearity problem and its machine learning approach: a review. Mathematics, 10(8), 1283.

[21] Lin, W. C., and Tsai, C. F. (2020). Missing value imputation: a review and analysis of the literature (2006–2017). Artificial Intelligence Review, 53, 1487-1509.

[22] Waljee, A. K., Mukherjee, A., Singal, A. G., Zhang, Y., Warren, J., Balis, U., ... and Higgins, P. D. (2013). Comparison of imputation methods for missing laboratory data in medicine. BMJ open, 3(8), e002847.

[23] Powers, D., and Xie, Y. (2008). Statistical methods for categorical data analysis. Emerald Group Publishing.

[24] Gogtay, N. J., and Thatte, U. M. (2017). Principles of correlation analysis. Journal of the Association of Physicians of India, 65(3), 78-81.

[25] Sharma[1], P., and Bhatia, A. P. R. (2012). Implementation of decision tree algorithm to analysis the performance.

[26] Sainju, B., Hartwell, C., and Edwards, J. (2021). Job satisfaction and employee turnover determinants in Fortune 50 companies: Insights from employee reviews from Indeed. com. Decision Support Systems, 148, 113582.

[27] Memon, M. A., Salleh, R., Mirza, M. Z., Cheah, J. H., Ting, H., Ahmad, M. S., and Tariq, A. (2021). Satisfaction matters: the relationships between HRM practices, work engagement, and turnover intention. International Journal of Manpower, 42(1), 21-50.