**VISVESVARAYA TECHNOLOGICAL UNIVERSITY**

Jnana Sangama, Belgaum- 590 018



SYNOPSIS ON

**“Employee Attrition Prediction”**

**ARTIFICIAL INTELLIGENCE AND DATA SCIENCE**

During the Academic year 2023-2024  
Submitted By

**Student Name: Rahul babu B U**

**USN:1DB21AD038**

Under the Guidance of

**Prof. Subhashini B N**

Assistant professor

Dept. of AI&DS

**

**Department of Artificial Intelligence and Data Science**

**Don Bosco Institute of Technology**

**Mysore Road, Kumbalgodu, Bengaluru-560074**

**Contents**

* **Introduction**
* **Problem Statement & Motivation**
* **Literature Survey**
* **Aim & Objectives**
* **Proposed system**
* **Methodology**
* **Requirement Specification**
* **Implementation Details**
* **Innovativeness and Usefulness**
* **Drawbacks**
* **Result**
* **References**

**Synopsis**

**Project Title: “Employee Attrition Prediction”**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Batch no.** | | | | |
| **Sl.no** | **USN** | **Name** | **E-mail id** | **Contact number** |
| 1. | 1DB21AD038 | Rahul babu B U | uarbrb@gmail.com | 8197384971 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sl.no** | **Guide Name** | **Designation** | **E-mail id** | **Contact number** |
| 1. | Prof. Subhashini B N | Assistant professor | Subhashini.bn@dbit.com | 9901029705 |

**Introduction:**

Employee attrition, or turnover, is a significant issue for organizations as it leads to the loss of skilled workers and incurs high recruitment and training costs. Predicting attrition can help organizations identify employees at risk of leaving and implement retention strategies.

Employee attrition, or turnover, is a critical issue faced by organizations worldwide, leading to substantial costs and operational disruptions. With the advent of machine learning, predicting which employees are at risk of leaving has become a viable solution to mitigate these challenges. By analyzing patterns in historical employee data, machine learning models can provide insights into the factors influencing attrition. This project aims to leverage these advanced techniques to develop a predictive model for employee attrition. The goal is to enable organizations to proactively implement retention strategies, thereby maintaining a stable and productive workforce.

**Problem Statement & Motivation:**

**Problem Statement:** Employee attrition negatively impacts organizations by increasing costs and disrupting operations. Predicting which employees are likely to leave can help mitigate these effects.

**Motivation:** Understanding the factors leading to employee attrition and predicting it can help organizations improve retention strategies, thus saving costs and maintaining productivity.

**Literature Survey:**

A review of existing literature shows that various machine learning models have been used for predicting employee attrition, including logistic regression, decision trees, random forests, and neural networks. Studies highlight the importance of features such as job satisfaction, work environment, and personal factors in predicting attrition.

**Title:** "Employee Attrition Prediction using Machine Learning Models: A Review"

* **Authors:** Various
* **Journal:** ResearchGate
* **Year:** 2022
* **Overview:** This study reviews machine learning models used for predicting employee attrition, highlighting techniques such as Random Forest, XGBoost, and Artificial Neural Networks. It underscores the effectiveness of these models in handling complex HR datasets and identifying significant attrition factors.

 **Title:** "Predicting Employee Attrition Using Machine Learning Techniques"

* **Authors:** A Raza et al.
* **Journal:** MDPI
* **Year:** 2022
* **Overview:** The research provides a systematic approach to using machine learning for employee attrition prediction. It emphasizes the importance of data preprocessing and feature selection to enhance model accuracy. Techniques such as Decision Trees, SVM, and ensemble models are discussed.

 **Title:** "Employee Attrition Prediction Using Machine Learning Models: A Review Paper"

* **Platform:** SlideShare
* **Year:** 2024
* **Overview:** This review paper explores various machine learning algorithms, including Random Forest and Neural Networks, for predicting attrition. It also discusses the challenges of data imbalance and the need for more interpretable models.

 **Title:** "Employee Attrition Prediction Using Machine Learning Algorithms"

* **Authors:** Various
* **Journal:** ResearchGate
* **Year:** 2024
* **Overview:** The study employs Random Forest, XGBoost, and Neural Networks on Kaggle's IBM HR dataset, demonstrating their effectiveness in predicting attrition. It highlights the potential of using feature importance scores to identify key attrition factors.

 **Title:** "Predicting Employee Attrition"

* **Authors:** R van Dam
* **Journal:** Tilburg University
* **Year:** 2021
* **Overview:** This study explores the impact of employee performance and business performance on attrition rates, emphasizing the need for integrating performance metrics into predictive models.

**Objective/Aim of the Project:**

**Aim:** To develop a machine learning model that accurately predicts employee attrition.

**Objectives:**

* To identify key factors influencing employee attrition.
* To preprocess and analyze employee data.
* To build and evaluate different machine learning models.
* To select the best-performing model for predicting attrition.

**Proposed system:**

The proposed system involves collecting employee data, preprocessing it, and using machine learning techniques to build a predictive model. The system will use historical data to learn patterns and make predictions on future attrition.

**Methodology**

* **Data Collection:** Gather data from sources such as HR databases.
* **Data Preprocessing:** Clean the data, handle missing values, encode categorical variables,
* and normalize numerical features.
* **Feature Selection:** Identify the most relevant features for predicting attrition.
* **Model Building:** Train various machine learning models (e.g., logistic regression,
* decision trees, random forests, neural networks).
* **Model Evaluation:** Evaluate models using metrics such as accuracy, precision, recall, and
* F1-score.
* **Model Selection:** Select the best-performing model based on evaluation metrics.

**Requirement Specification:**

* **Software:** Python, Jupyter Notebook, libraries (pandas, numpy, scikit-learn, matplotlib,seaborn).
* **Hardware:** Standard computer with sufficient RAM and processing power.
* **Data:** Historical employee data with features such as demographics, job role, satisfaction level, and previous attrition status.

**Implementation Details:**

* **Data Preprocessing:** Implement scripts to clean and preprocess the data.
* **Model Training:** Train and validate multiple machine learning models.
* **Evaluation:** Use cross-validation to assess model performance.
* **Deployment:** Deploy the best model in a user-friendly interface for HR managers to use.

**Innovativeness and Usefulness:**

**Innovativeness:** The project utilizes advanced machine learning techniques to predict employee attrition, which is not commonly used in traditional HR practices.

* **Advanced Data Analysis**: Machine learning models can process large volumes of data to identify patterns and predict future employee behavior. This ability to analyze complex datasets significantly advances over traditional methods that rely on simple statistical analysis or subjective assessments.
* **Predictive Accuracy**: The use of algorithms such as Random Forest, Gradient Boosting, and Extra Trees enhances the predictive accuracy of attrition models, allowing organizations to foresee potential resignations and address underlying issues proactively.
* **Customization and Adaptability**: Machine learning models can be customized to reflect different organizations' unique characteristics and needs. They can adapt to changes in the workforce and continuously improve their predictions as new data becomes available.
* **Comprehensive Insights**: By analyzing factors such as monthly income, age, overtime, and distance from home, these models provide a thorough understanding of the factors contributing to employee attrition, enabling more effective human resource strategies.

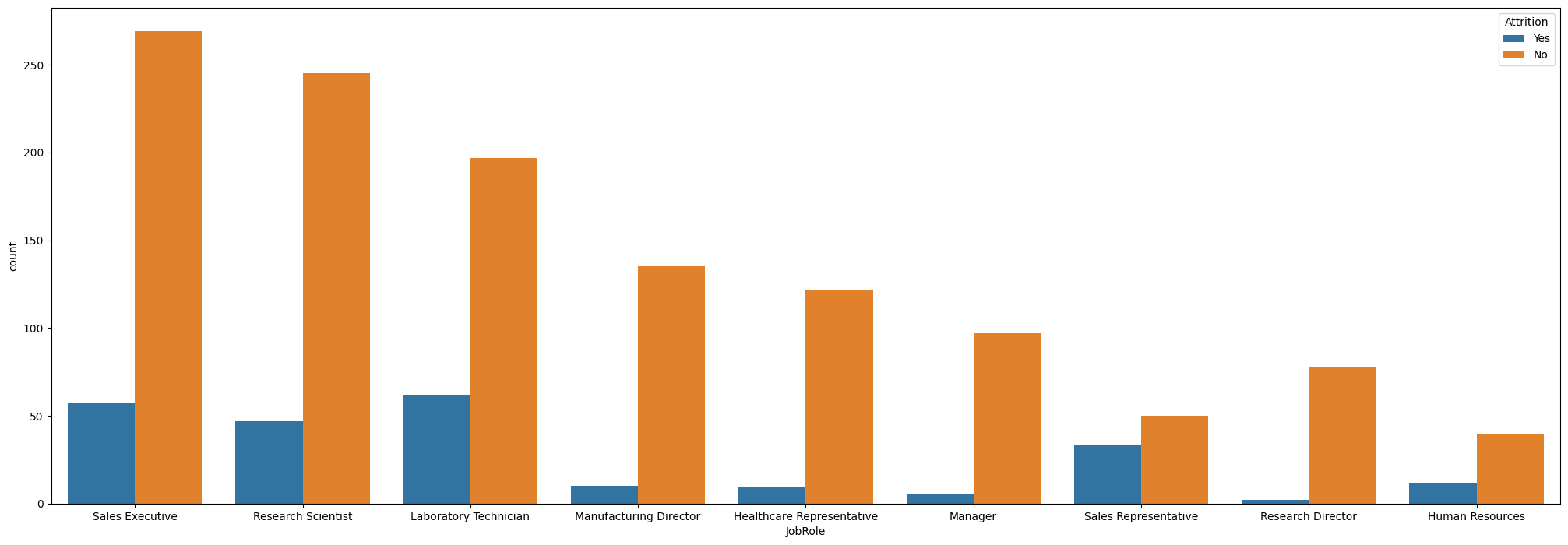
**Usefulness:** The predictive model can help organizations proactively address attrition, leading to better retention strategies and reduced costs.

* **Proactive Retention Strategies**: By predicting which employees are likely to leave, organizations can implement targeted retention strategies to address potential issues before they result in turnover.
* **Cost Reduction**: High employee turnover is costly due to recruitment, training, and lost productivity. Predictive analytics help reduce these costs by identifying at-risk employees early and enabling timely interventions.
* **Improved Workforce Planning**: Attrition predictions allow companies to plan for workforce changes effectively, ensuring that replacements are ready and minimizing disruptions in operations.
* **Enhanced Employee Engagement**: Understanding the factors leading to employee dissatisfaction helps organizations improve job satisfaction and engagement, reducing the likelihood of attrition.
* **Data-Driven Decision Making**: Machine learning models provide insights into the factors influencing attrition, helping HR teams make informed decisions and optimize human resource policies.

**Drawbacks:**

* **Data Quality and Availability**: Effective prediction models require high-quality, comprehensive datasets. Incomplete or biased data can lead to inaccurate predictions, making it difficult for organizations to trust the insights generated.
* **Privacy Concerns**: Collecting and analyzing employee data for attrition prediction can raise privacy issues. Employees may feel uncomfortable with the level of monitoring and analysis being conducted on their personal and professional data.
* **Complexity and Cost**: Implementing predictive analytics involves significant investment in technology and expertise. Smaller organizations may find it challenging to justify the cost and complexity associated with developing and maintaining such systems.
* **Overreliance on Models**: Organizations might over-rely on prediction models and overlook qualitative factors that contribute to employee satisfaction and retention. This can lead to decisions that fail to consider the broader context of employee experiences and workplace culture.
* **Resistance from Employees**: Employees might resist the use of predictive models due to fear of being unfairly categorized as potential leavers, which can affect morale and trust within the organization.

**Result:**



**References:**

* [researchgate.net - employee attrition prediction using machine learning ...](https://www.researchgate.net/publication/364322002_EMPLOYEE_ATTRITION_PREDICTION_USING_MACHINE_LEARNING_ALGORITHMS)
* [researchgate.net - Employee Attrition Prediction using Machine Learning ...](https://www.researchgate.net/publication/379860312_Employee_Attrition_Prediction_using_Machine_Learning_Models_A_Review_Paper)
* [mdpi.com - Predicting Employee Attrition Using Machine Learning ...](https://www.mdpi.com/2076-3417/12/13/6424)
* [slideshare.net - EMPLOYEE ATTRITION PREDICTION USING MACHINE ...](https://www.slideshare.net/slideshow/employee-attrition-prediction-using-machine-learning-models-a-review-paper/267182623)
* [ieeexplore.ieee.org - Predicting Employee Attrition using Machine Learning](https://ieeexplore.ieee.org/document/8605976)
* [link.springer.com - Explaining and predicting employees' attrition: a machine ...](https://link.springer.com/article/10.1007/s42452-020-2519-4)