Employee Performance and Retention Analysis - Final Report

Project Objective

The goal of this project was to analyze employee data to understand key factors influencing employee performance and attrition, and to build predictive models using machine learning and deep learning techniques.

Dataset Overview

The dataset 'Employee_data.csv' contains the following key features:

- Employee ID, Name

- Age, Department, Salary

- Years at Company

- Performance Score

- Attrition (Yes/No)

After removing duplicates and missing values, the dataset was ready for analysis and modeling.

Key Analyses and Findings

1. Exploratory Data Analysis (EDA)

- Mean, median, variance, and standard deviation were calculated for numerical features.

- Correlation analysis showed that Salary and performance score are positively correlated.

- Boxplots helped identify outliers.

2. Statistical and Probability Analysis

- Attrition Probability: Higher in departments with low average performance.

Employee Performance and Retention Analysis - Final Report

- Bayes' Theorem: Higher attrition among low performers.

- ANOVA test showed significant difference in performance scores across departments.

Predictive Modeling

3. Employee Attrition Classification (Random Forest)

- Accuracy: ~85%

- Balanced classification with good precision and recall.

4. Performance Score Regression (Linear Regression)

- R² Score: ~0.68

- Acceptable error and strong prediction trend.

Deep Learning Models

5. Neural Network - Performance Prediction

- Feedforward model with better fit than linear regression.

6. Neural Network - Attrition Classification

- Accuracy: ~87%

- Outperformed classical models in prediction confidence.

Insights and Recommendations

Insights:

- Low performance, short tenure, and specific departments (Sales, Support) are high-risk.

Employee Performance and Retention Analysis - Final Report

- Performance depends on salary, experience, and department.

Recommendations:

- Mentorship and training programs.

- Monthly feedback systems.

- Clear career paths for new employees.

- Employee engagement initiatives.

Visualizations Summary

- Line plots for performance over years.

- Bar charts for attrition by department.

- Scatter plots: Salary vs. Performance.

- Confusion matrices and regression visualizations.

Conclusion

The analysis revealed key factors behind employee performance and attrition. ML and DL models predicted these effectively. These insights support data-driven HR decisions to improve retention and growth.