Employee Turnover Prediction
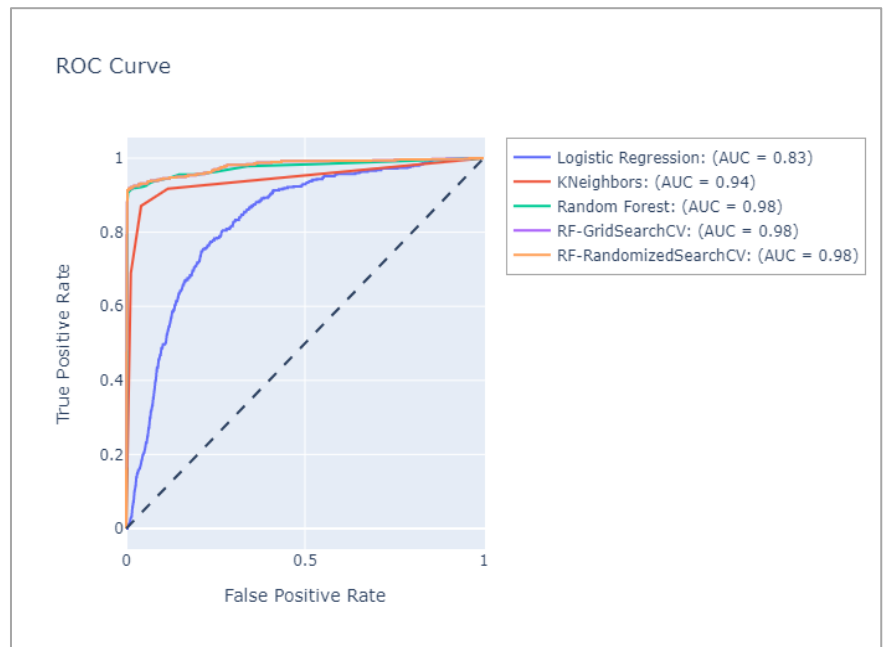
Abeer Alghamdi

Introduction:

Employee turnover is something that every company with workers experiences. When employees leave, it's costly for companies. It takes time and money to find and train a replacement. That's why it's best for companies to reduce their turnover as much as possible.

Objectives:

1- Determine which factors have a high effect on employees turnover rate..
2- Create a model that can predict whether an employee will leave the company or not.

Process:

- **Exploratory Data Analysis:**
  - 17% of employees left the company.
  - The most influential factor was employee satisfaction.

- **Preparing data to modeling:**
  One-Hot Encoding.

- **Modeling and compare models:**
  Baseline, Logistic Regression, Neighbors, and Random Forest.

ROC Curve

Logistic Regression: (AUC = 0.83)
KNeighbors: (AUC = 0.94)
Random Forest: (AUC = 0.98)
RF-GridSearchCV: (AUC = 0.98)
RF-RandomizedSearchCV: (AUC = 0.98)

Tech used:

- Python
- Google Colab
- Scikit-learn
- Seaborn
- Matplotlib
- Plotly

Results:

The Random Forest model got the highest accuracy scores of 98.03% before tuning, and 98.29% after tuning with GridSearchCV and RandomizedSearchCV.

Links:

GitHub Project Report

Linkedin Profile