



Data Dynasty

FINAL PROJECT RAKAMIN BATCH 33

Employee Attrition

Prediction

Source : Dataset

Code : Picture





As Data Scientist consultants, our responsibility involves analyzing the factors that contribute to employee attrition within a company. We will create a machine learning model capable of predicting the likelihood of an employee leaving the company, and provide business recommendations based on the results of the predictions and analysis.

Meet the Teams



Citta Mudita
as
*Project
Leader*



Ikhlasul Amala
as
Data Analyst



Afif Rahman
as
*Data
Scientist*



Farrell
Wahyudi
as
*Data
Scientist*



M. Galuh
Saputra
as
*Business
Analyst*



Adi Nur
Rokhim
as
*Business
Analyst*

About the Client



SlothCompany

Human resources are considered as an important aspect of an organization, and voluntary **employee attrition has been identified as a key issue**. **Sloth Company** one of the Biggest FMCG Company in Indonesia is concerned about this topic.

Based on data have been collected by their HR Department, **Sloth Company** want to identify whether an **employee is likely to attrite** so they can increase their ability to intervene on time and possibly provide a remedy to the situation to prevent attrition.

Sloth Company asked **Data Dynasty** to help them Predict and Identify which aspects drive the **employee to attrite**.

What is The Problem ?

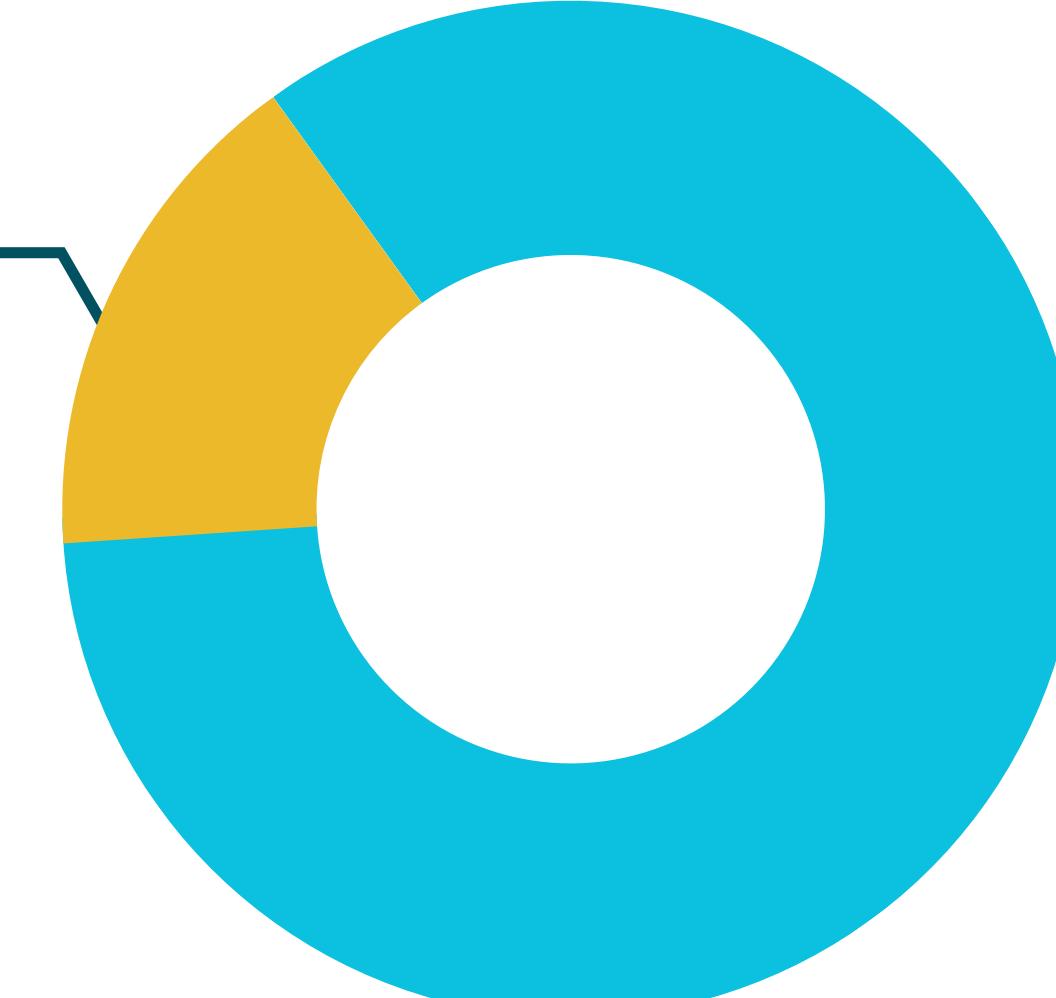


It's a High rate of Attrition

Attrite
237

Not Attrite
1233

Employee Attrition Rate



What is a good attrition rate for an organization?

A "good" attrition rate will depend on the company size, niche and the number of employees, among other factors. But most organizations should aim for a **10% attrition rate to keep the company running smoothly.**

Source : Certified Human Resource Management Professional

What is The Problem ?



Employee attrition refers to the natural reduction in the employees in an organization due to many unavoidable factors. Employee attrition results in a massive loss for an organization. The Society for Human Resource Management (SHRM) determines that USD 4129 is the average cost-per-hire for a new employee. According to (Raza, Munir, Almutairi, Younas, & Fareed, 2022)



What is The Problem ?

Background

Sloth Company want to know which aspects drive the employee to attrite.

Goals

Decrease the Attrition Rate from the target feature from 16,1% to at least 10%.

Objectives

Create a model to predict potential employee attrition and identify the aspects of employee attrition.

Business Metrics

Attrition Rate (%)

Exploratory Data Analysis

-  Dataset Information
-  Preliminary Insight



Datasets



The Dataset consists of **1470 rows** and **35 columns**



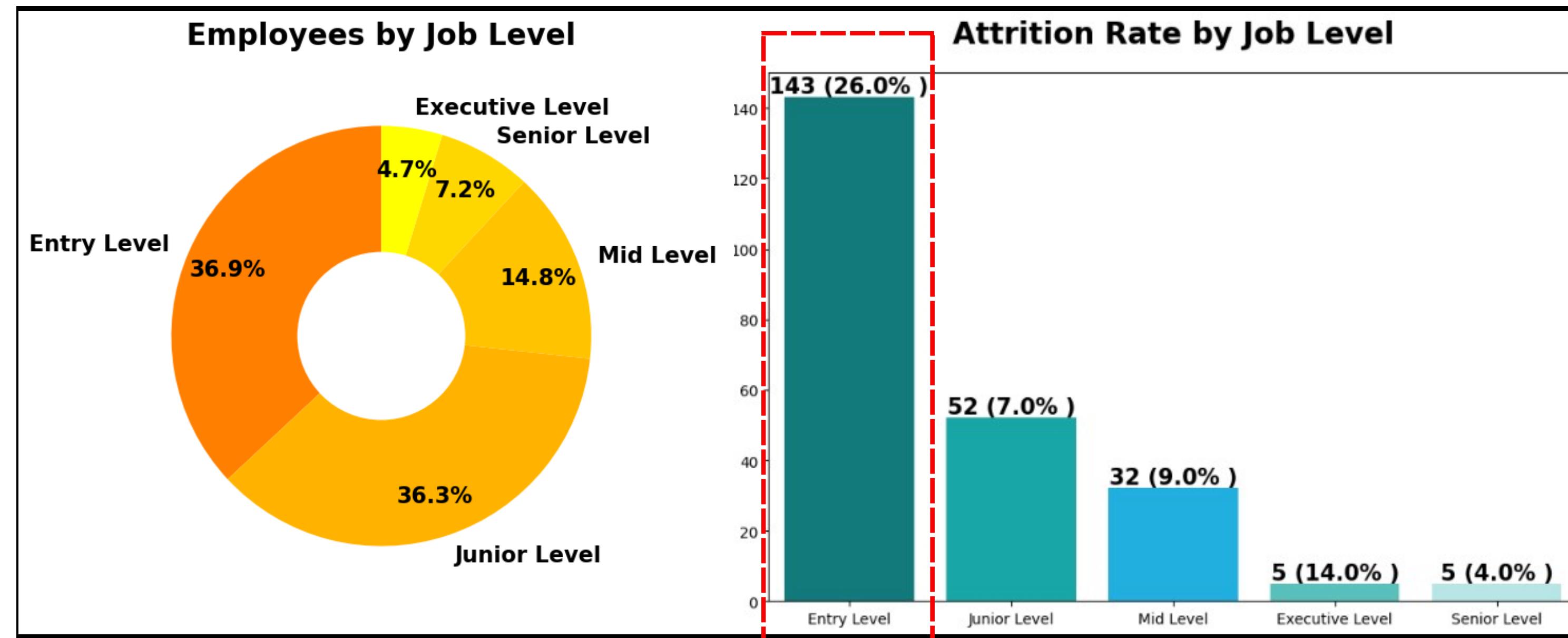
It has **No Missing Value** and **No Duplicated**



Target variable is the **Attrition** column

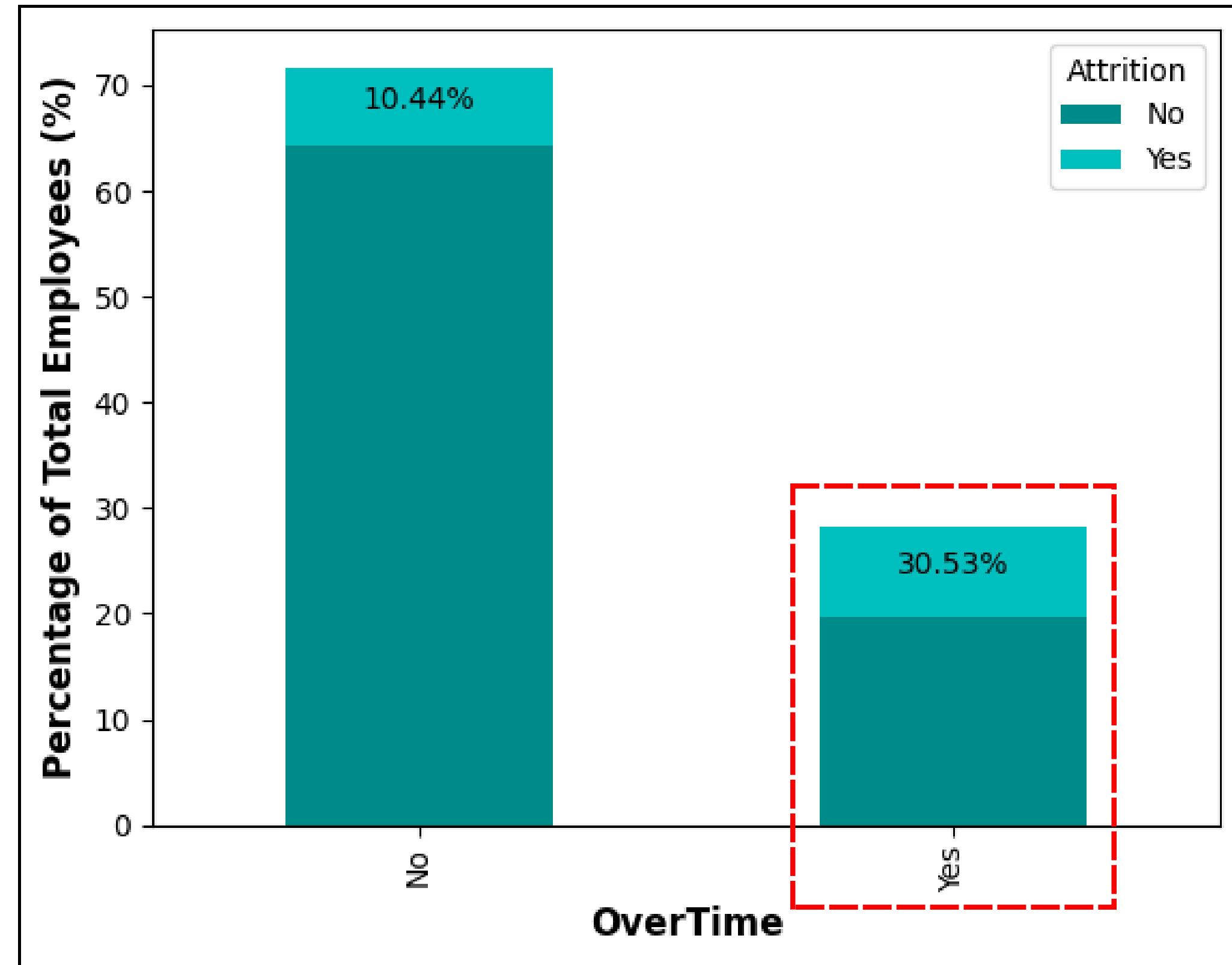
Attrition by Job Level

Majority of employees in the organization are at Entry Level or Junior Level. The Highest Attrition is at the Entry Level.



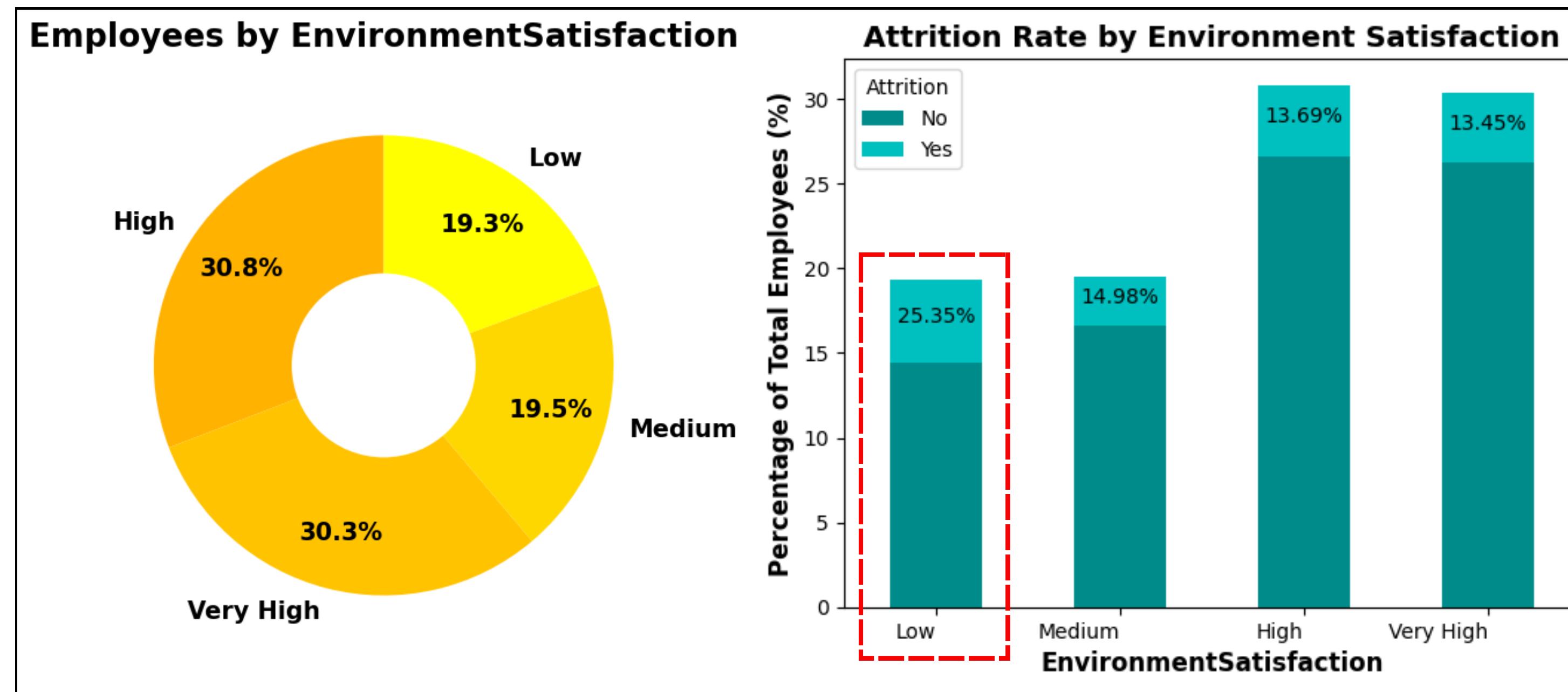
Attrition by Over Time

Employees who **overtime** are more likely to **attrition** compared to employees who do not overtime



Attrition by Environment Satisfaction

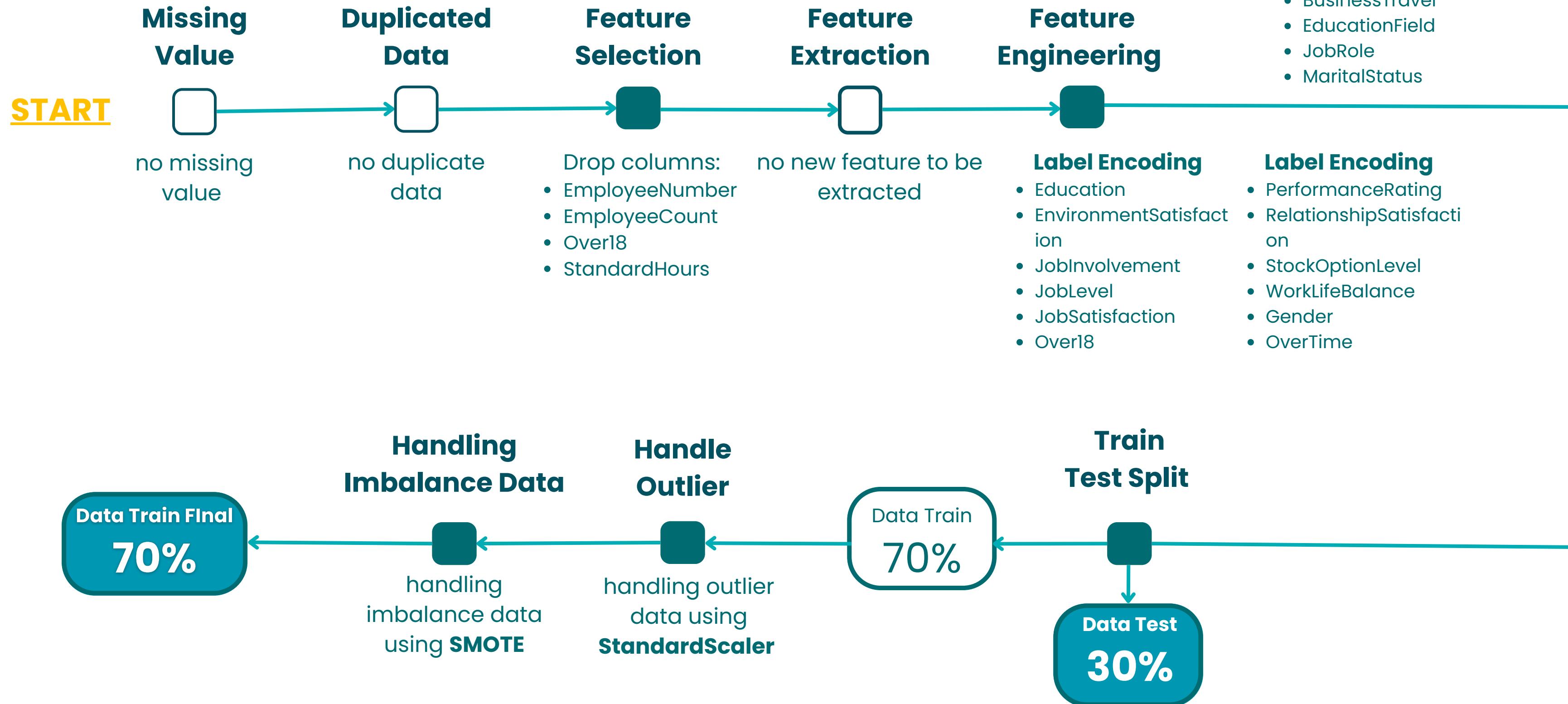
Majority of employees rate organizational environment satisfaction **High** & **Very High**. However, there are **still very high levels of attrition** in this environment.



Data Preprocessing

- Feature Engineering
- Handling Imbalance Data

Data Preprocessing Flow



Machine Learning Modelling

-  Model Evaluation
-  Feature Importance

Model Evaluation

Models tested:

- Logistic Regression
- **XGBoost**
- Random Forest
- AdaBoost

After Hyperparameter tuning, the model with the highest Recall and fairly high accuracy is XGBOOST

Model Comparison

MODELS	ACCURACY (TRAIN)	ACCURACY (TEST)	RECALL (TRAIN)	RECALL (TEST)	AUC (TRAIN)	AUC (TEST)
LOGISTIC REGRESSION	0.83	0.83	0.61	0.59	0.82	0.81
XGBOOST	0.85	0.81	0.78	0.75	0.89	0.82
RANDOM FOREST	0.70	0.75	0.63	0.56	0.73	0.72
ADABOOST	0.86	0.82	0.64	0.54	0.86	0.81

Model Evaluation

Models tested:

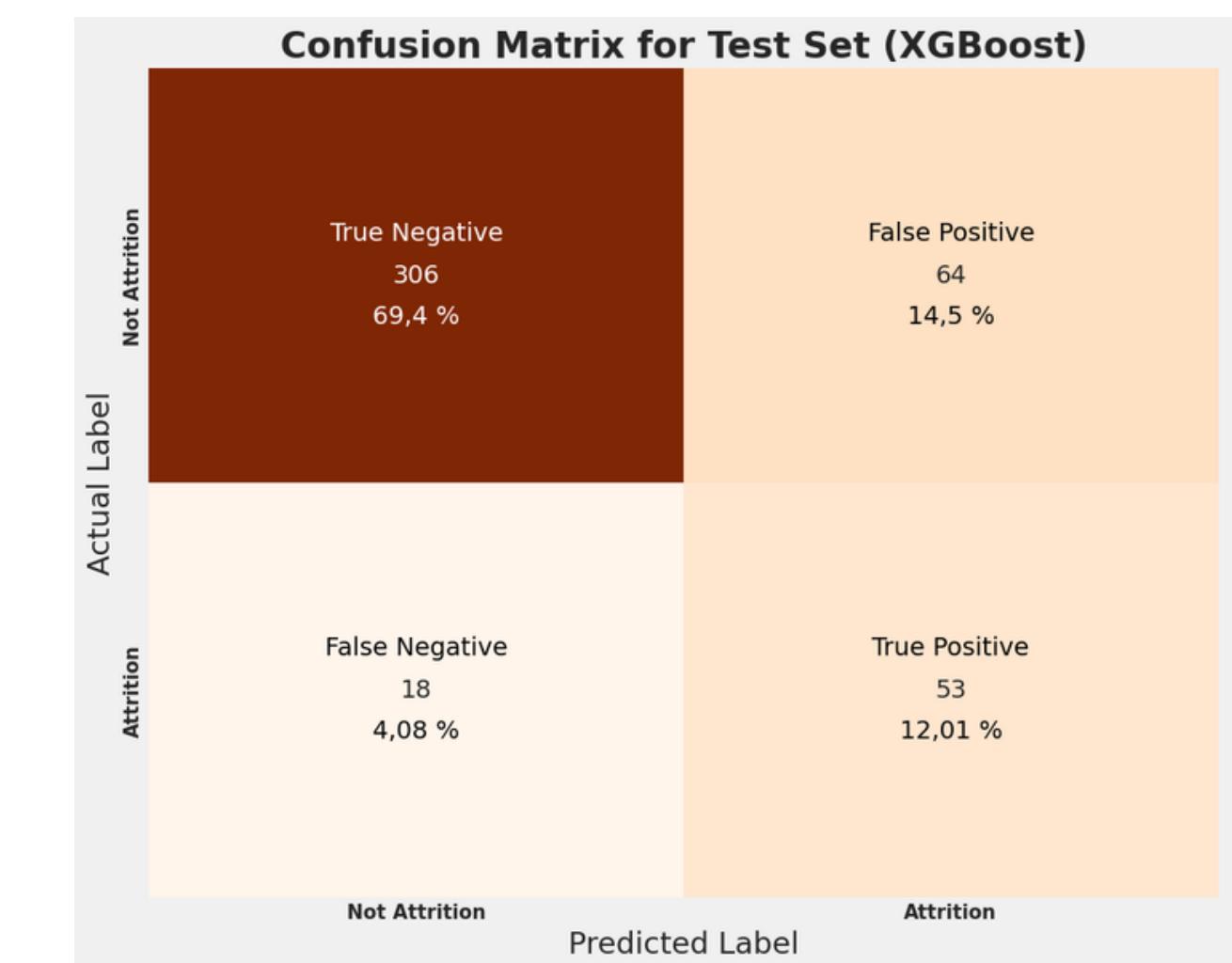
- Logistic Regression
- AdaBoost
- Random Forest
- **XGBoost**

XGBoost

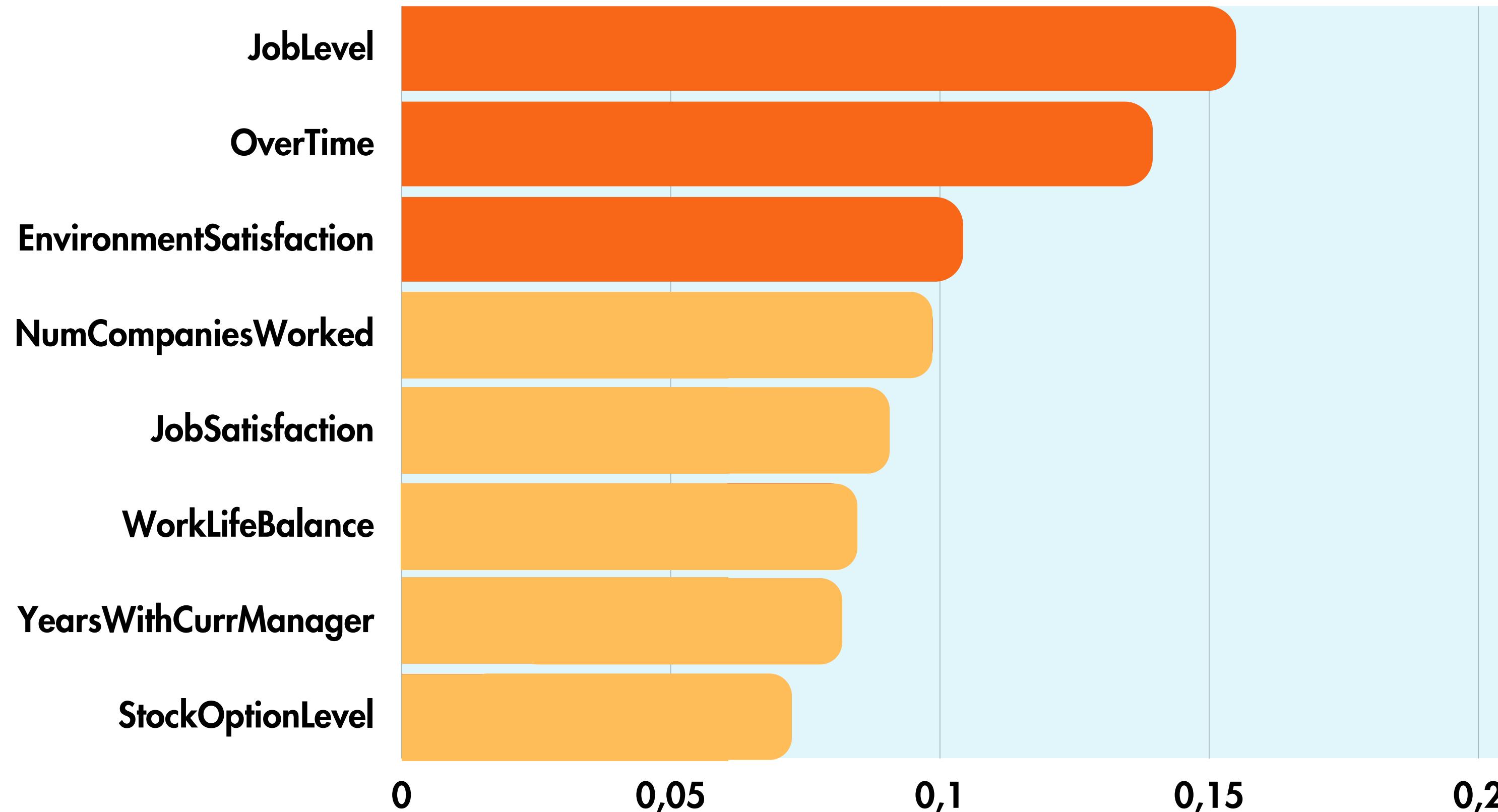
- SMOTE
- Hyperparameter Tuning

Recall measures the ratio between the number of employees that are predicted to leave who actually leave the company, and the number of employees that are predicted to stay but actually end up leaving the company.

Maximizing recall means minimizing the number of employees falsely predicted to stay.



Feature Importance



Business Recommendation



Recommendation



Business Impact

Recommendation

Recommendation HR

Based on the **Feature Importance** analysis conducted after **modeling**, it is known that the top 3 most influential features.



Recommendation HR

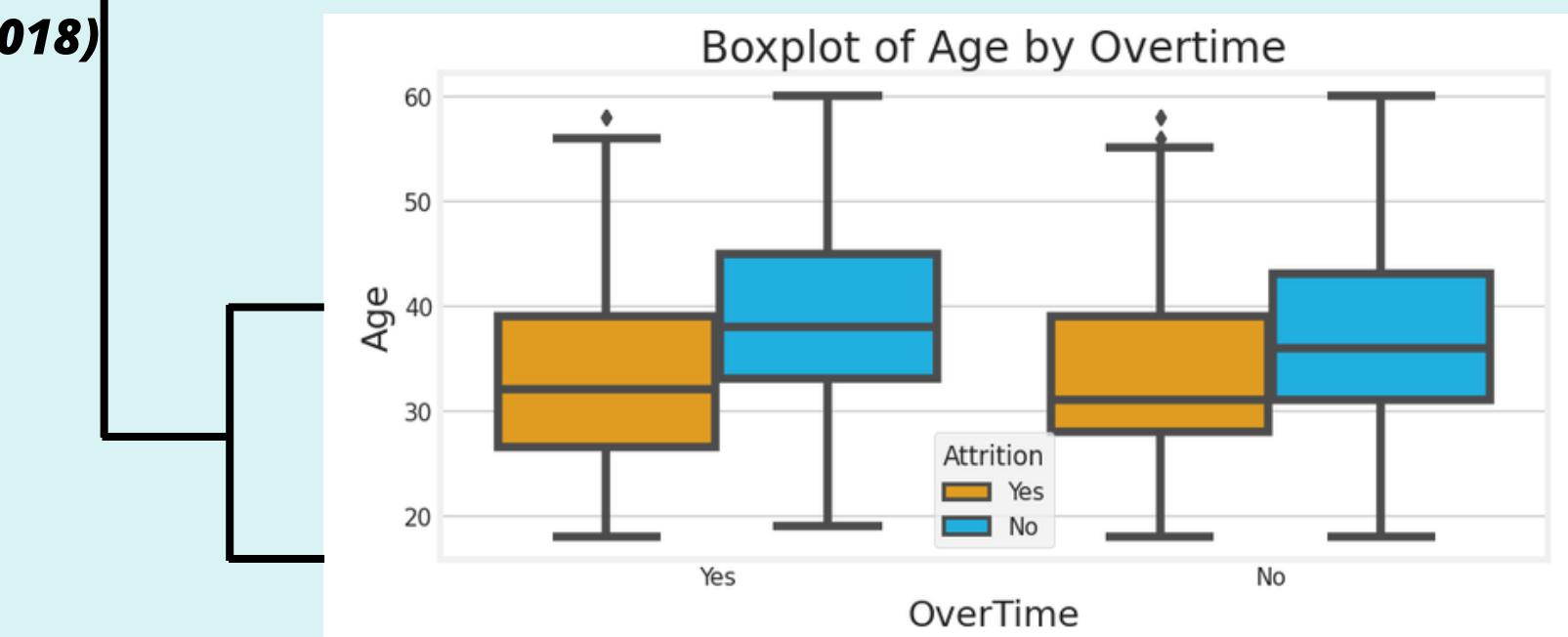
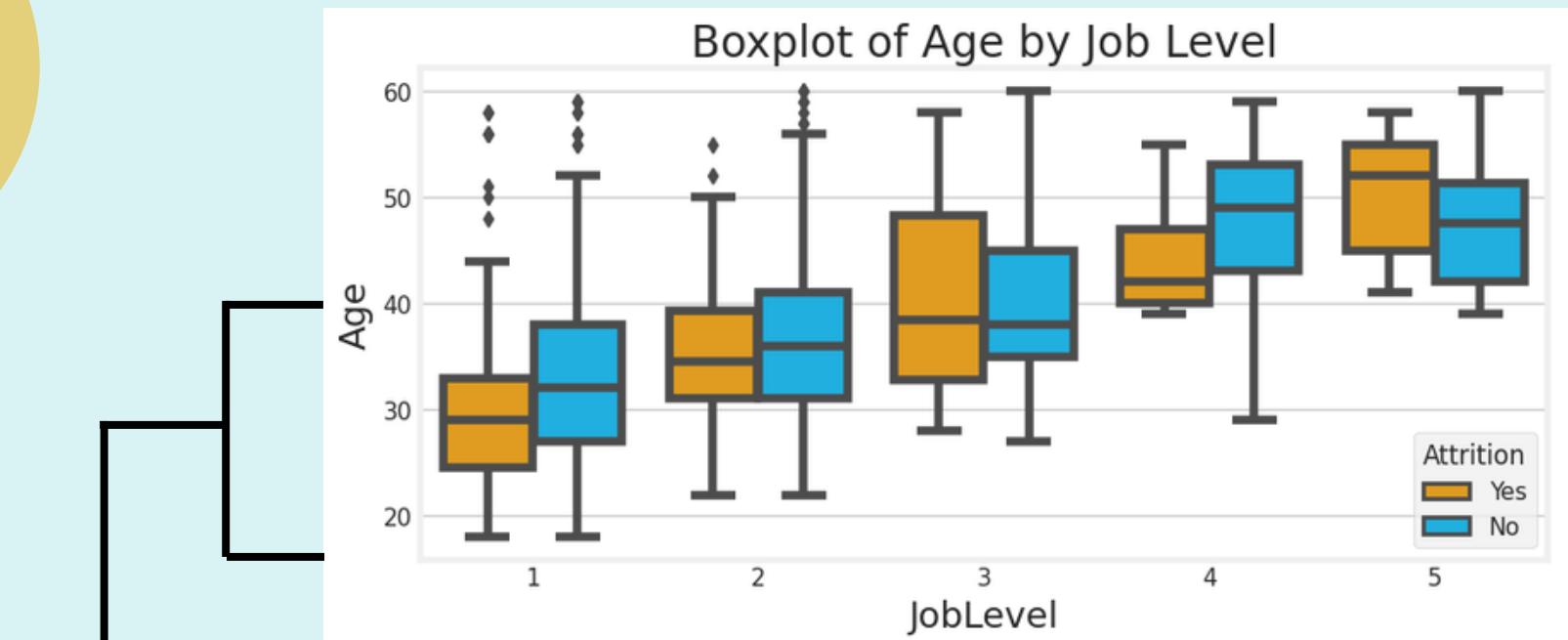


Provide opportunities for advancement and Improve work efficiency

Career Development

“Career development has a positive and significant effect on employee retention”
–Yadewani, D. (2021)

Productive age
-Pranata, H. (2018)



Recommendation HR



Improve environment facility

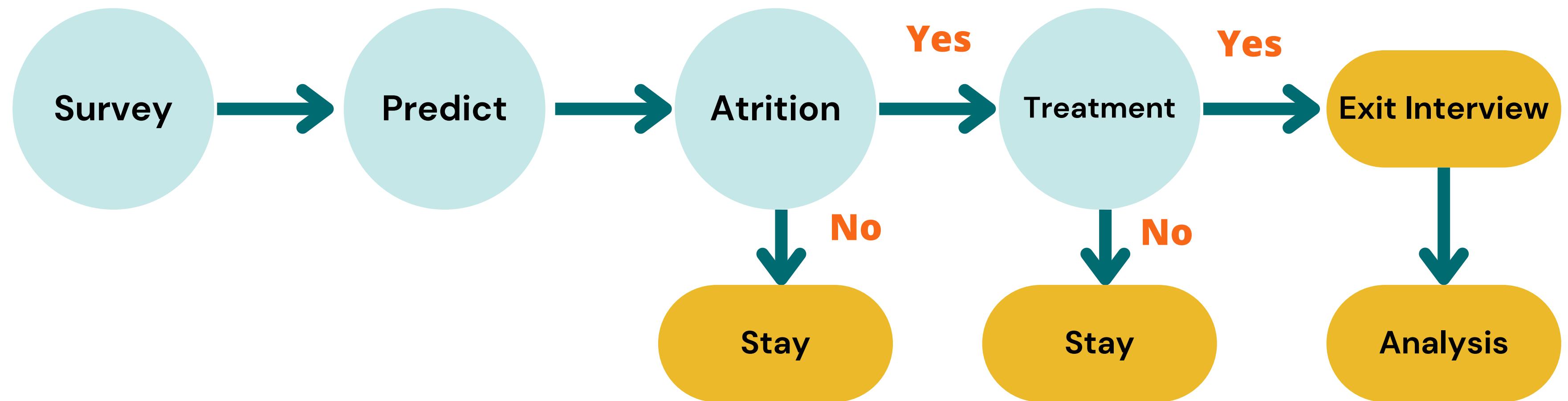
Add and/or improve the quality of environmental facilities

"A positive work environment can help employees feel more motivated and productive, thereby reducing the likelihood of them leaving the organization."

-Nam Choi, J (2019)



Recommendation Software Development



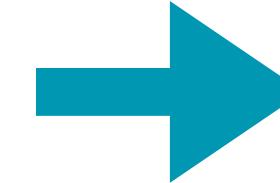


Business Impact

Attrition Rate

True Positive (TP) = 53
False Negative (FN) = 18
Total = 441

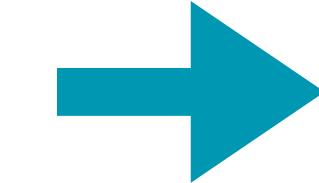
237



Assumption 50%

ATTRITION RATE = (0.5(TP)+FN)/TOTAL

Before Model



147

After Model



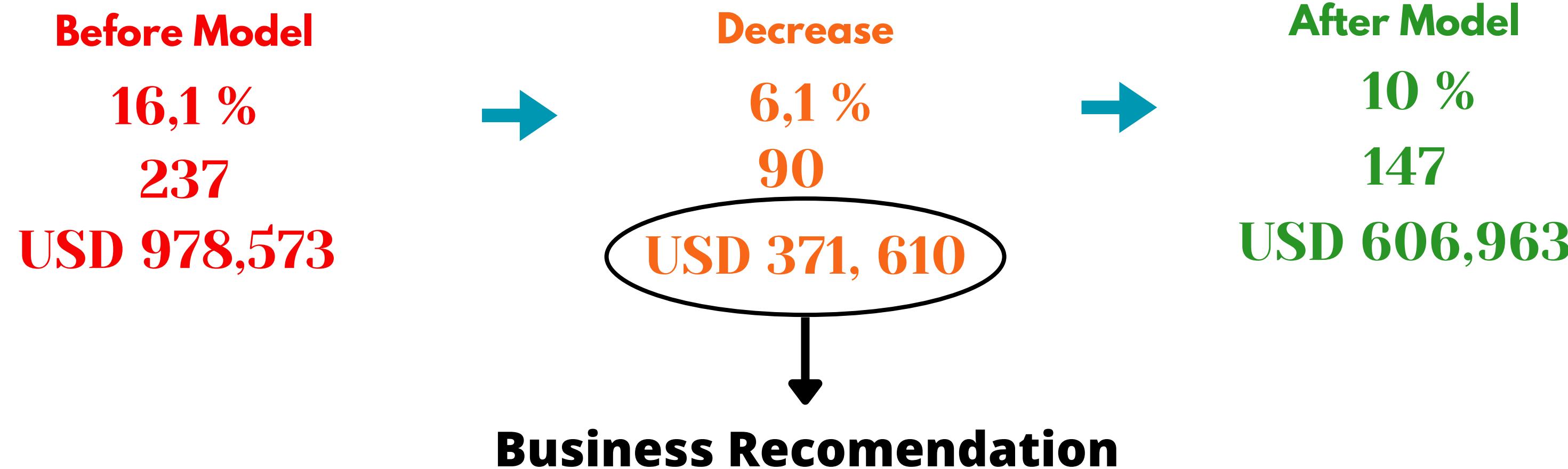
90

A DECREASE OF 38% ▼



Business Impact

Recruitment Cost

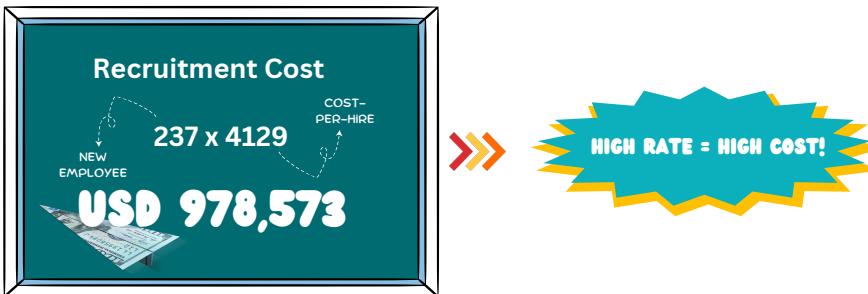
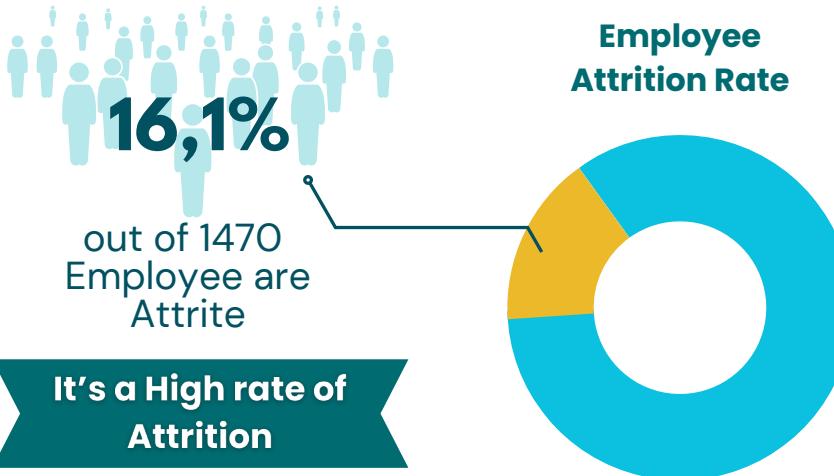




Employee Attrition

Prediction

Business Problem



■ Goals

Decrease the Attrition Rate from the target feature from 16,1% to at least 10%.

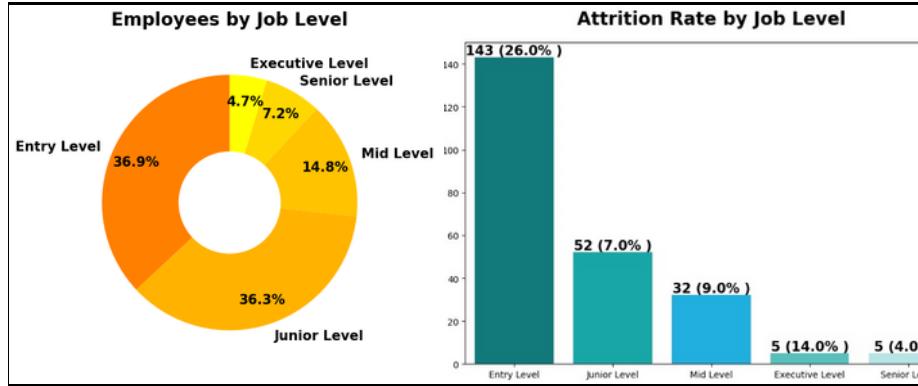
■ Objective

Create a model to predict potential employee attrition and identify the aspects of employee attrition.

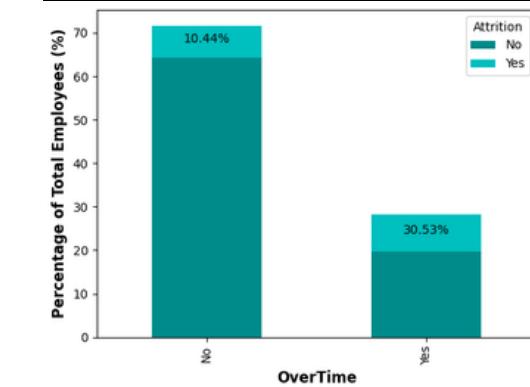
■ Business Metric

Attrition Rate (%)

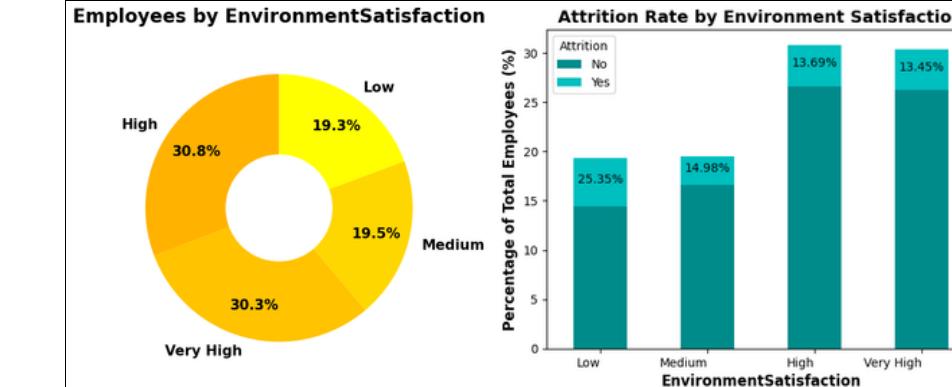
Attrition by Job Level



Attrition by Over Time



Attrition by Environment Satisfaction



Modelling

■ XGBoost

- SMOTE
- Hyperparameter Tuning



Recommendation

■ Recommendation HR

- Provide opportunities for advancement and Improve work efficiency with program Career Development
- Improve environment facility. Add and/or improve the quality of environmental facilities

■ Recommendation Software Development

Attrition Rate

Decrease

6,1 %

Saving Cost

USD 371, 610

Thank You