Predicting Attrition Rates of Employees

Will Byrd
Flatiron School 2024
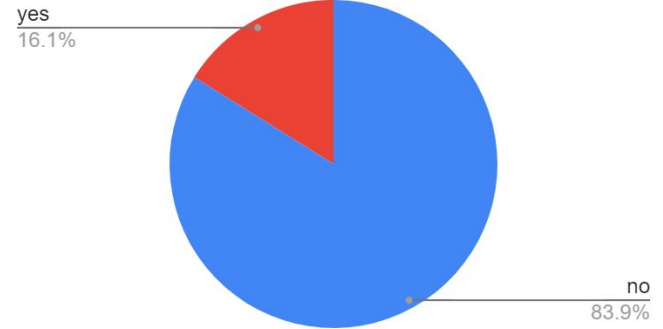
Business Overview

Count of Employees by Age and Attrition
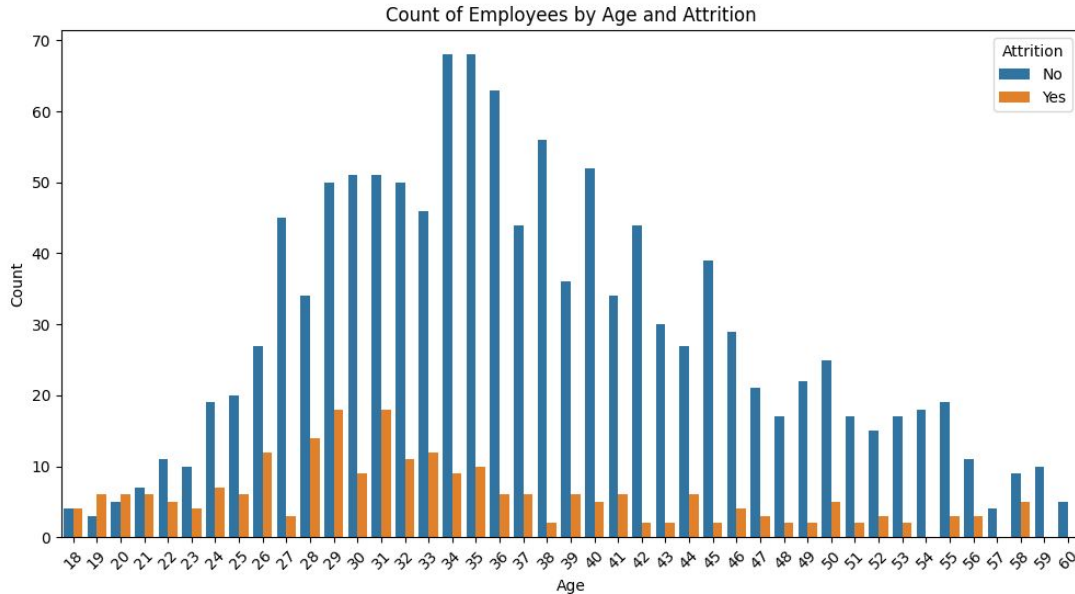
Most of 'Attrition' is happening to employees under 40
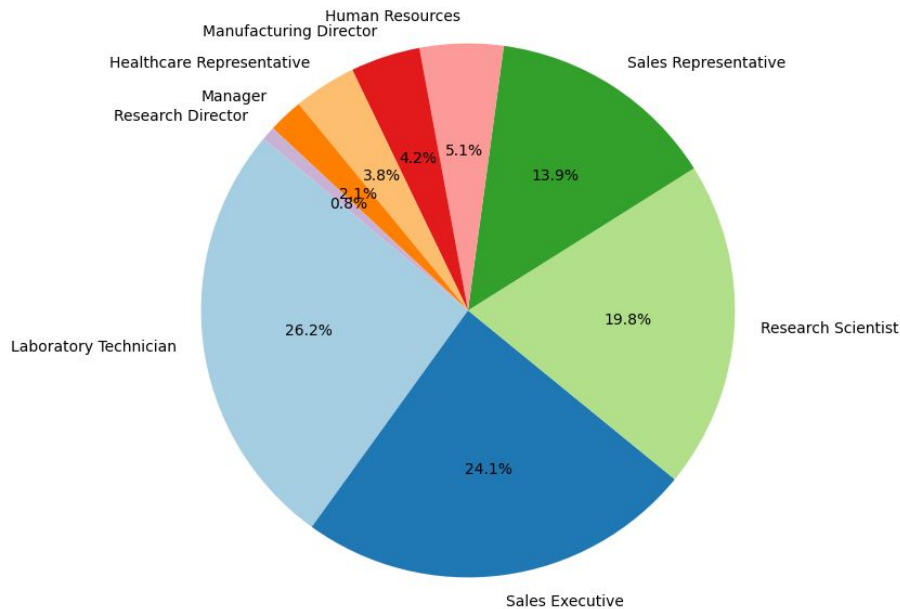
Attrition

yes
16.1%

no
83.9%

- **Predicting Attrition**
- Resource planning
- Cost of *false negative*
  - An employee leaving unexpectedly
  - Cost to hire/train new employees
  - No Counteractive measures for retention
- Cost of *false positive*
  - Employee expected to leave, but stays
  - Poor culture persists

Data Overview

Job Role Distribution of Terminated Employees

- IBM HR Dataset
  - **1470 Employees**
    - Sales
    - Research & Development
    - Human Resources

- Objectives
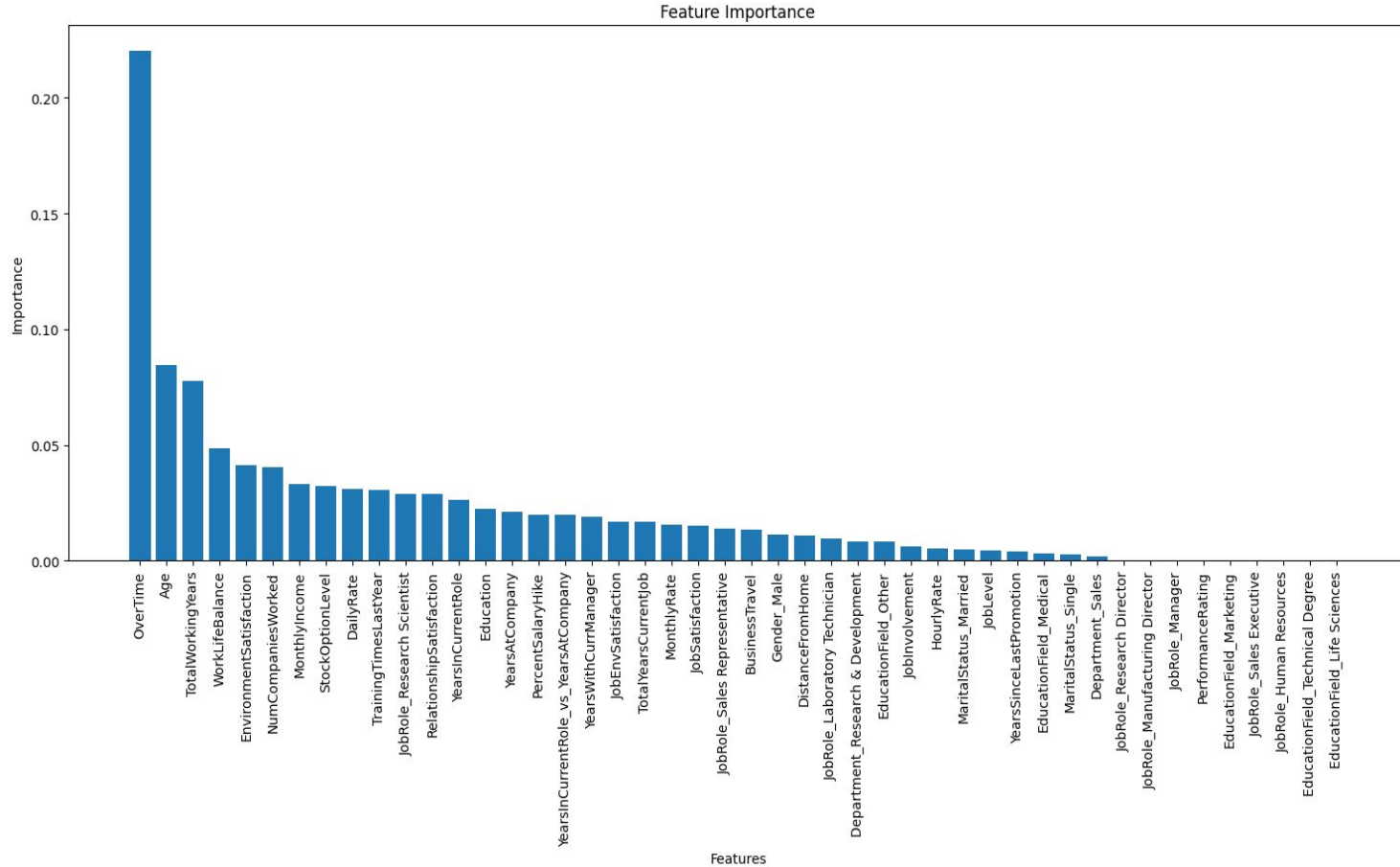 - Predict retention of employees
  - Optimize Recall

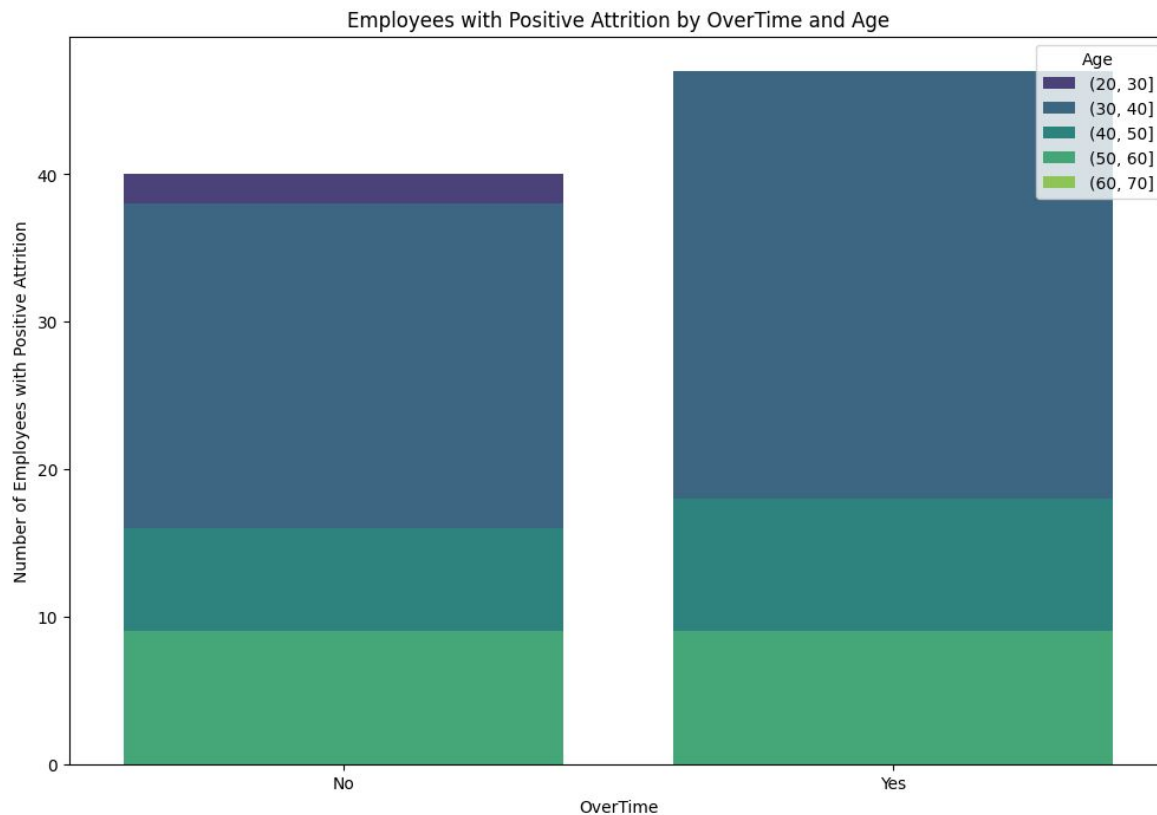*** Notice Sales, Research Scientist, and Laboratory Technician make up roughly 80% of terminated employees*

Data Analysis

*Our most important features in predicting 'Attrition' are OverTime, Age, TotalWorkingYears, and WorkLifeBalance*

Feature Importance

Data Analysis Cont.

Employees with Positive Attrition by OverTime and Age

Employees aged 30-40 have the highest rates of attrition (and frequency in dataset)

Measuring

Receiver Operating Characteristic (ROC) Curve

- **True negative-87%**
  - Model correctly predicted an employee will stay with the company
  - *83% of our dataset*
- **True Positive- 51%**
  - Model correctly predicted an employee will leave the company
  - *16% of our dataset*
- Generally, AUC equal to 50% is no better than guessing
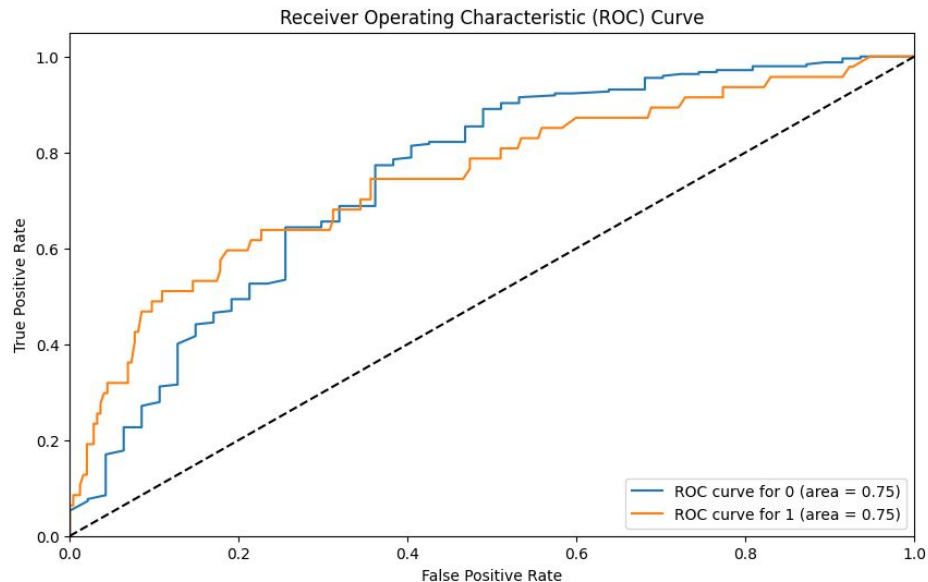
Results

- **20 models**
- *6 iterations* of each (baseline, balanced, fine-tuned)
  - Decision Tree
  - Random Forest
  - K-Nearest neighbors
- **2 Voting models**

|	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Baseline	0.758503	0.300000	0.382979	0.336449	0.606469
Final Model	0.812925	0.428571	0.510638	0.466019	0.690542

Recommendations

- Create a culture that excites and motivates
- Appeal to younger employees=sustainable growth
  - Increased wages
  - Better work life balance
    - remote/hybrid
    - Counter measures for employees likely to leave

Limitations in this data include:

- Small overall size
- Imbalanced
- Missing context on 'Attrition' reason
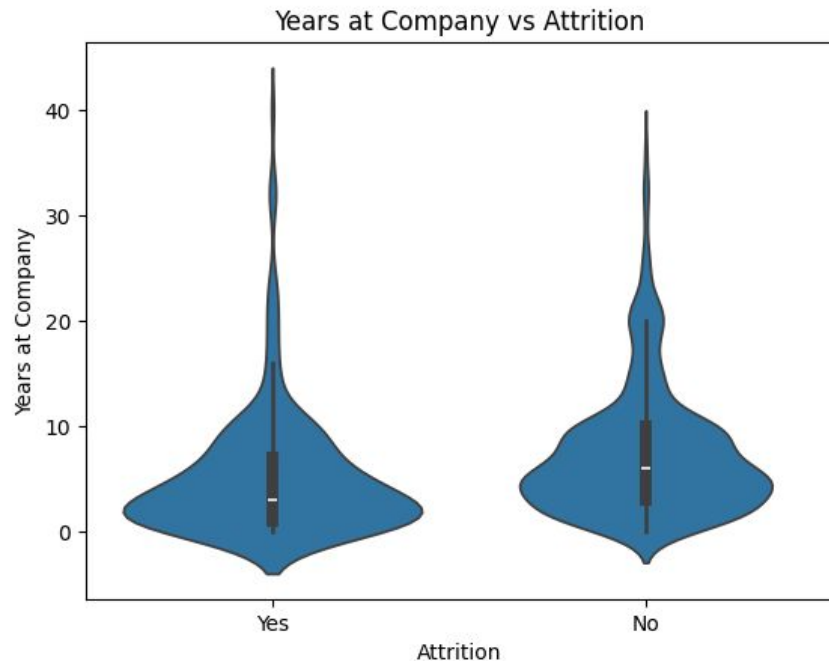
Employee with 40 years of tenure is probably retiring, but we can't be for sure with this dataset

Additional Questions?

Thank you!

By: Will Byrd

byrdwcrawford@gmail.com

The Flatiron School