Predicting Attrition Using ML

Isaac Liem
Corrina Hanson
Lara Mechling

Attrition is a costly business

- Loss of productivity
- Loss of Skills
- Essential knowledge
- Overall chemistry & synergy
- Monetary Cost of replacement

People are the Business

In order to mitigate this loss, it is important to understand what causes employees to leave. For businesses to thrive, they must continually look at their bottom line. One way to improve upon this is to cut or mitigate employee attrition losses.

Top Dataset Features

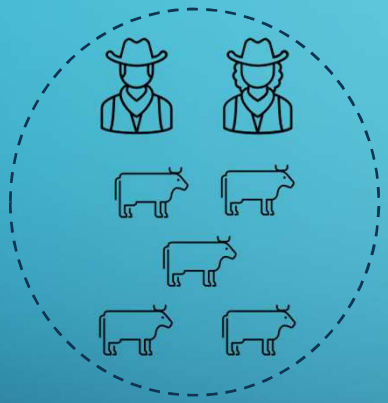
The features that had the greatest impact on attrition were:

- Years Since Last Promotion
- Over Time
- Number of Companies Previously Employed
- Distance from Home

There were an additional 30 features that are not listed

Method for Outcome

Data Wranglin'



- Cleaning
- EDA
- Feature Engineering

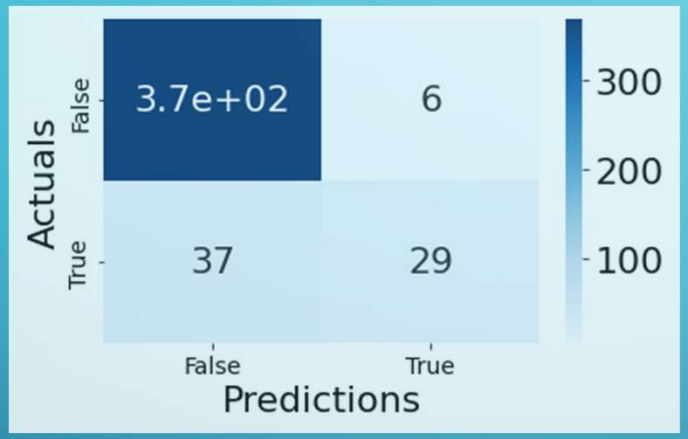
Data Modelling

- Sklearn
- Random Forest Classifier

Validation

- Confusion Matrix
- Classification Report

Model Metrics



	precision	recall	f1-score	support
0	0.91	0.98	0.94	375
1	0.83	0.44	0.57	66
accuracy			0.90	441
macro avg	0.87	0.71	0.76	441
weighted avg	0.90	0.90	0.89	441

Recall and Accuracy are great for 0, non-attrition, but recall drops off for 1, attritional employees.

References

Ismail, K. (2022). Forecasting Attrition With HR Data. Can HR Analytics Predict Attrition? (reworked.co)

Mcfeely, S., & Wigert, B. (2019). This Fixable Problem Costs U.S. Businesses $1 Trillion. This Fixable Problem Costs U.S. Businesses $1 Trillion (gallup.com)

Scikit-Learn Machine Learning in Python. scikit-learn. (n.d.). Retrieved November 13, 2022, from https://scikit-learn.org/stable/