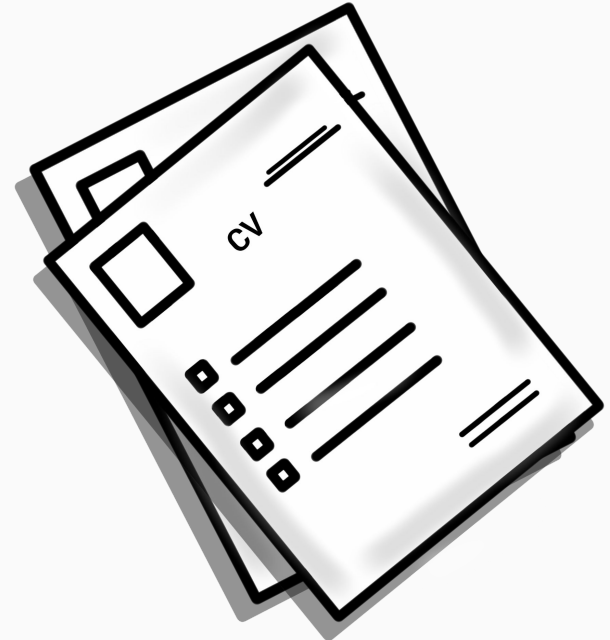


# Employee Attrition

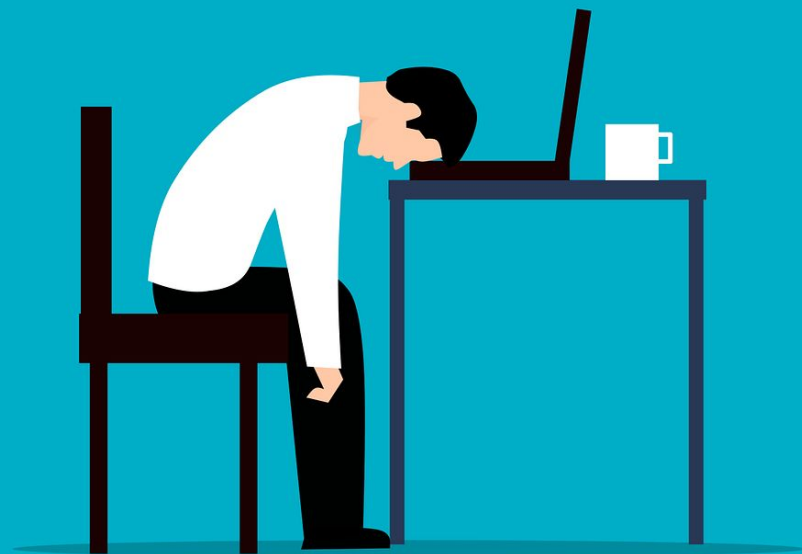
Natural Chan

# Background

- “The Great Resignation”
- Worker apathy & Anti-Work Mentality
- Who is likely to leave their job?
- Identify strong indicators
- Classify based on demographics & employment information



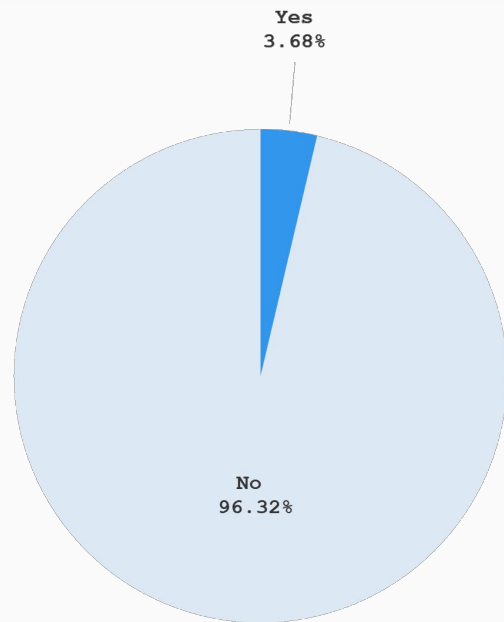
# Data



- U.S. Bureau of Labor Statistics
- National Longitudinal Surveys (**NLS97**)
- Job Search Label

# Modeling

- **Class Imbalance**
  - Oversample positive label
  - **SMOTE** with **5-Fold CV**
- 
- **Create pipelines** for each model
  - F-2 Score



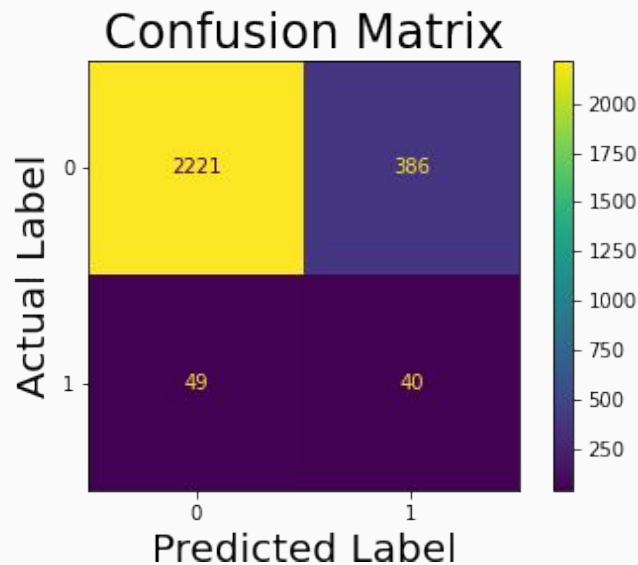
# Base Model(s)

	Accuracy	Precision	Recall	F1	<b>F2</b>
Decision Tree	.9230	.1508	<b>.1995</b>	.1686	<b>.1848</b>
Random Forest	.9591	.2233	.0309	.0533	.0371
XGBoost	.9587	.3506	.0764	.1240	.0902

# Hyperparameters

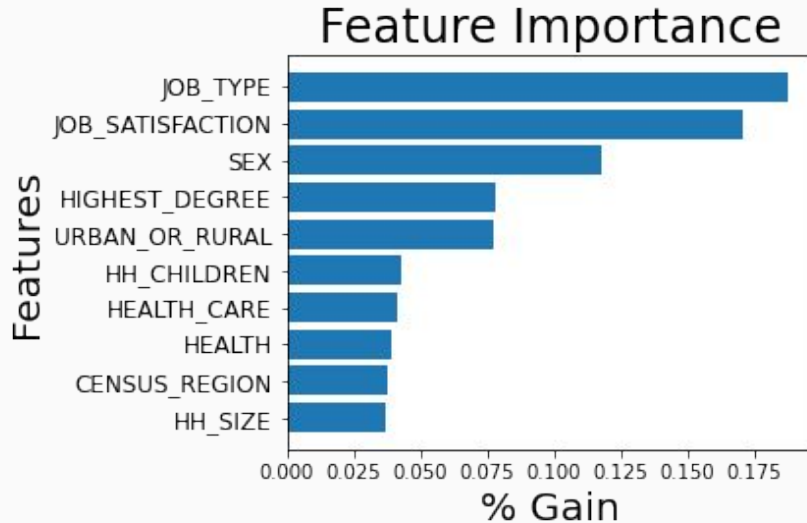
	Accuracy	Precision	Recall	F1	<b>F2</b>
Decision Tree	.8750	.1076	.3820	.1679	.2530
Random Forest	.7637	.0873	<b>.6517</b>	.1541	.2843
XGBoost	.8587	.1164	<b>.4944</b>	.1887	<b>.2997</b>

# Results



- **XGBoost**
- Hyperparameters:
  - colsample\_bylevel 0.9729
  - colsample\_bynode 0.4627
  - colsample\_bytree 0.8789
  - eta 0.0556
  - gamma 8.4169
  - max\_delta\_step 54
  - max\_depth 5
  - min\_child\_weight 9

# Results



- **Job Type** (Standard, Self-employed, non-traditional, military)
- **Job Satisfaction**
- **Household Information** is a weaker indicator



# Conclusion/Future Work

- Identify **employee satisfaction** via surveys/feedback
- **Limitations of Data**
- Specify career fields
- Increase sample size by analyzing other years



# Appendix

PUBID	JS_PROB	JOB_SEARCH
4111.0	0.842040	1
3800.0	0.829341	1
1515.0	0.827314	1
1504.0	0.822710	0
7944.0	0.818371	1
1408.0	0.808698	0
3418.0	0.796898	0
4709.0	0.796657	0
3856.0	0.793794	0
3006.0	0.791338	0

Top 10 Most Confident Predictions

Only 4/10 were actually searching for a job

Top 3 were true positive

All 3 responded: “Think it is OK” for job satisfaction at a Non-Traditional Job