Clinical Trials Prediction

By Jeffery Rosario

Problem Statement

Is the clinical trial successful?

Goal

To predict whether the clinical trial is success or unsuccessful

Business Value

- Tool for healthcare researchers to determine whether a clinical trial will succeed
- Help determine changes needed for a successful predicted outcome

Methodology

- 5395 Clinical Trials - clinicaltrials.gov

  - minimum/maximum age, intervention type, intervention name

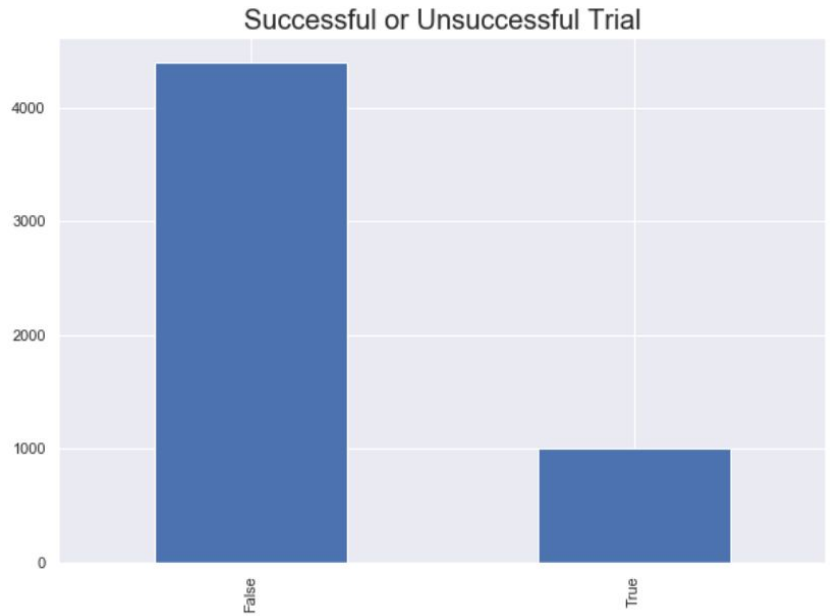
- ~30000 PMID reference papers
- Logistic Regression
- Cross Validation

Data Imbalance

-Imbalance of target variable

- 1:4 ratio

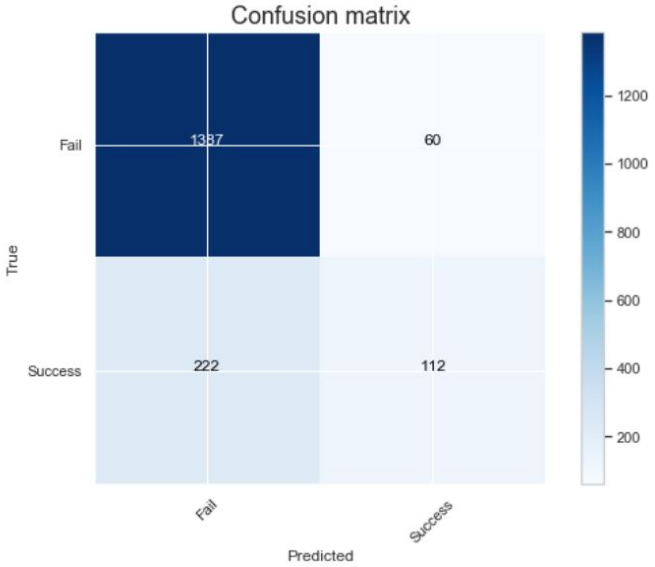
- Account for imbalance using SMOTE

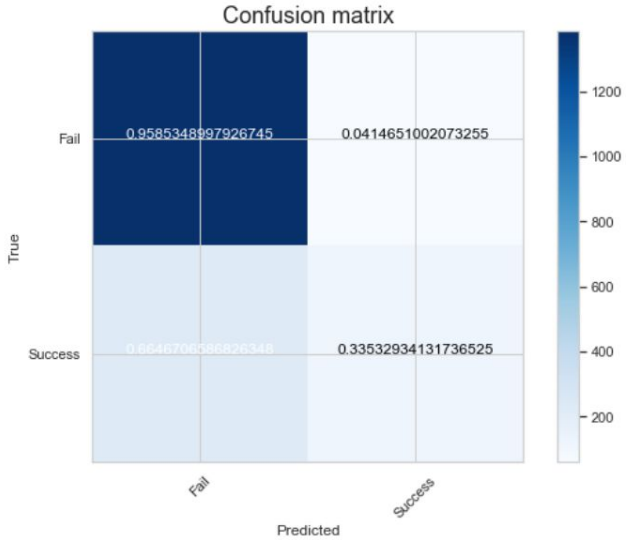


Classification Report

	Precision	Recall	F1-Score	Support
Success - w/out Text	0.55	0.25	0.34	334
Success	0.66(+0.11)	0.36(+0.11)	0.47(+0.13)	334
Failure - w/out Text	0.85	0.95	0.90	1447
Failure	0.87(+0.02)	0.96(+0.01)	0.91(+0.01)	1447

Confusion Matrix





Findings

- Including text vectorizers modestly improved overall model accuracy and precision
- Machine learning models can predict clinical trials outcomes

Recommendations

- Use model to predict outcomes of a clinical trial
- Be wary of unseen interventions when using the model as a predictor
-

Future Work

- Bigger Dataset, institutional access to data
- Search a bigger breadth of hyperparameters for machine learning models

Thank You!