Predicting Startup Failure

By Natalya Doris

Business Problem

With the economy at a possible turning point, ABC Ventures hopes we can answer the following questions:

- Given a set of data points describing funding levels, location, market & timing, is it possible to **predict** whether a startup will succeed* or fail, with emphasis on identifying possible **failures**?

- What **factors** are most important in predicting failure?

*Success is defined as a company that has been acquired, as labeled in the data

Data

- Source: Crunchbase / [Kaggle](#)

- Over 50,000 observations; focused on subset of 6,200

- Features include funding metrics, location, timing & market category

Objectives

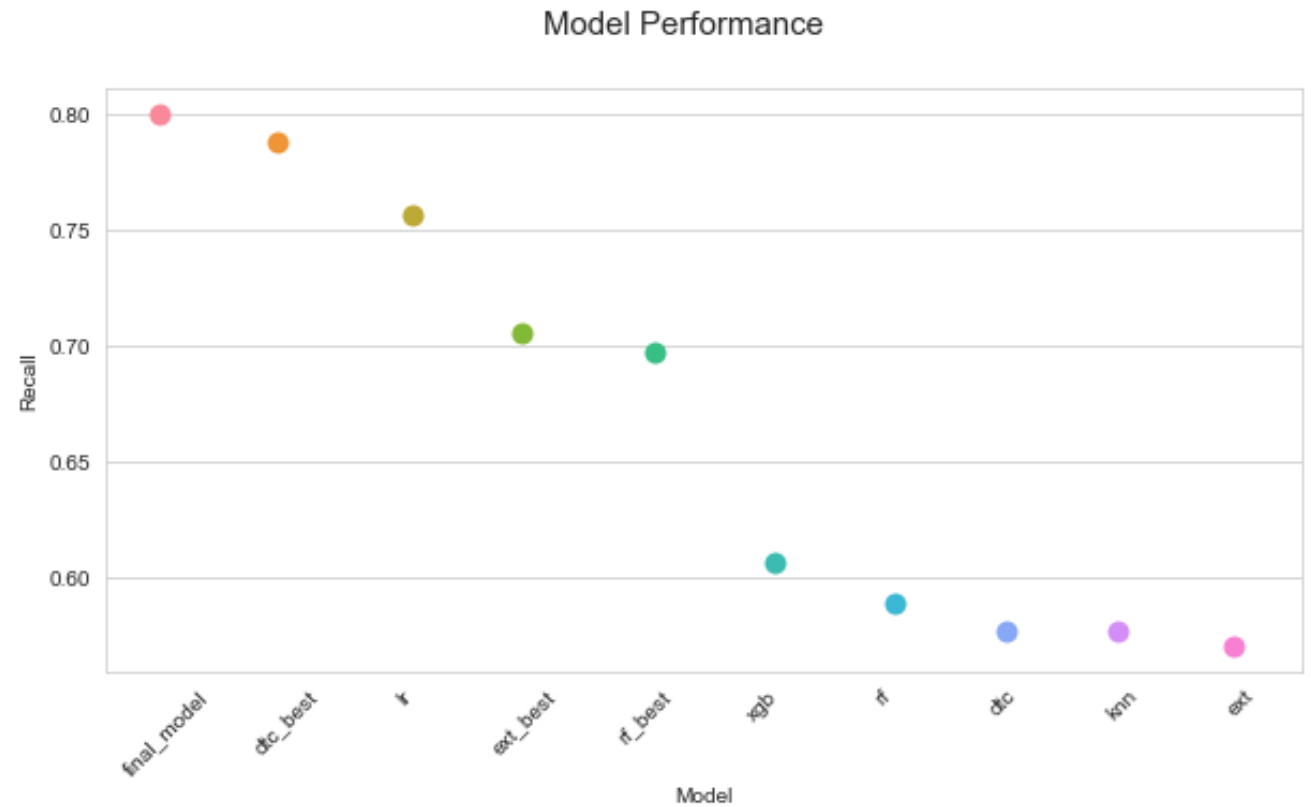Targeting **status** (company acquired or closed), we would like to:

1. Correctly identify the companies that **failed**: recall

2. Keeping in mind **false positives**: ROC AUC score

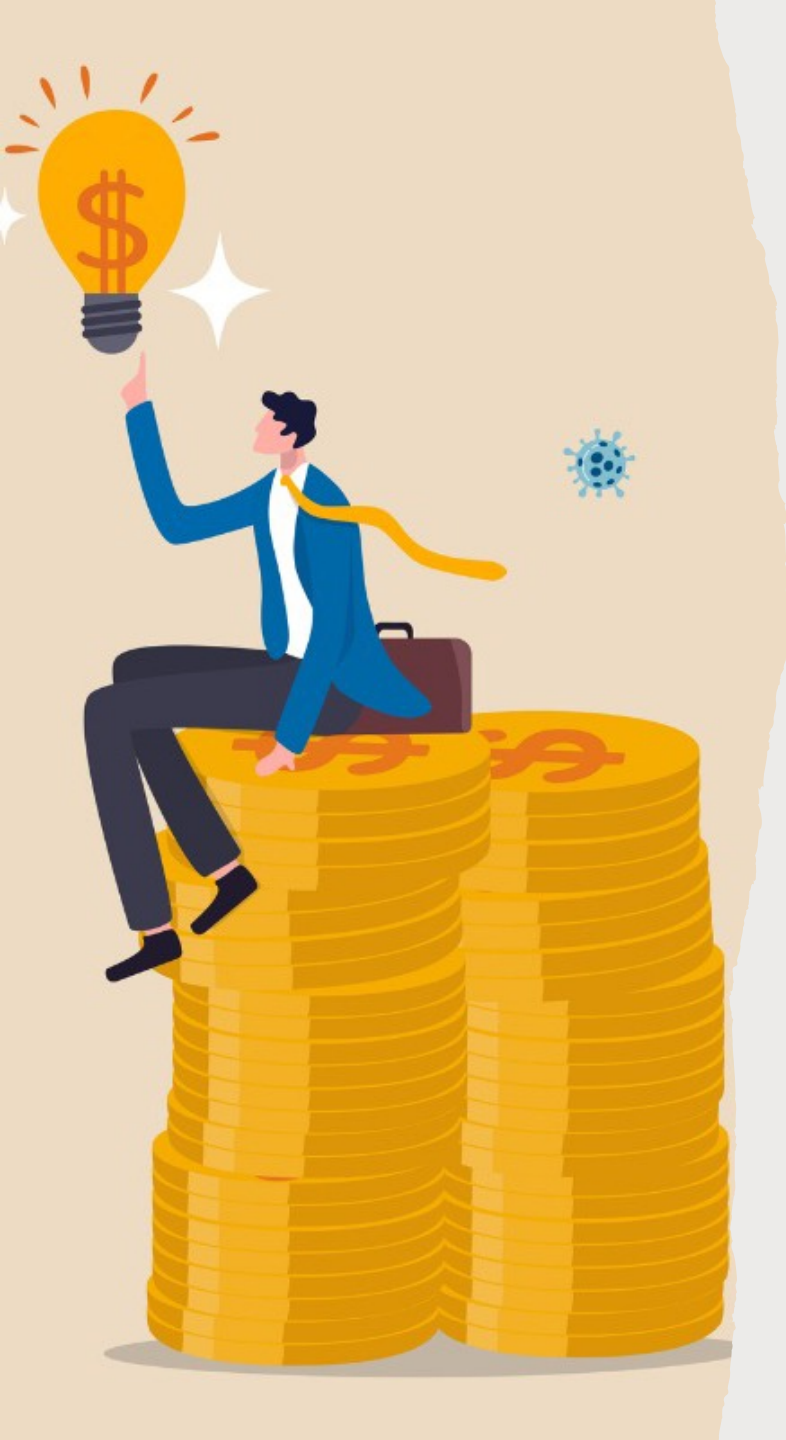3. Produce an **accurate** model

Model Selection

Final Model: XG Boost

80% Recall
- Correctly identified 80% of the companies that failed

73% ROC AUC
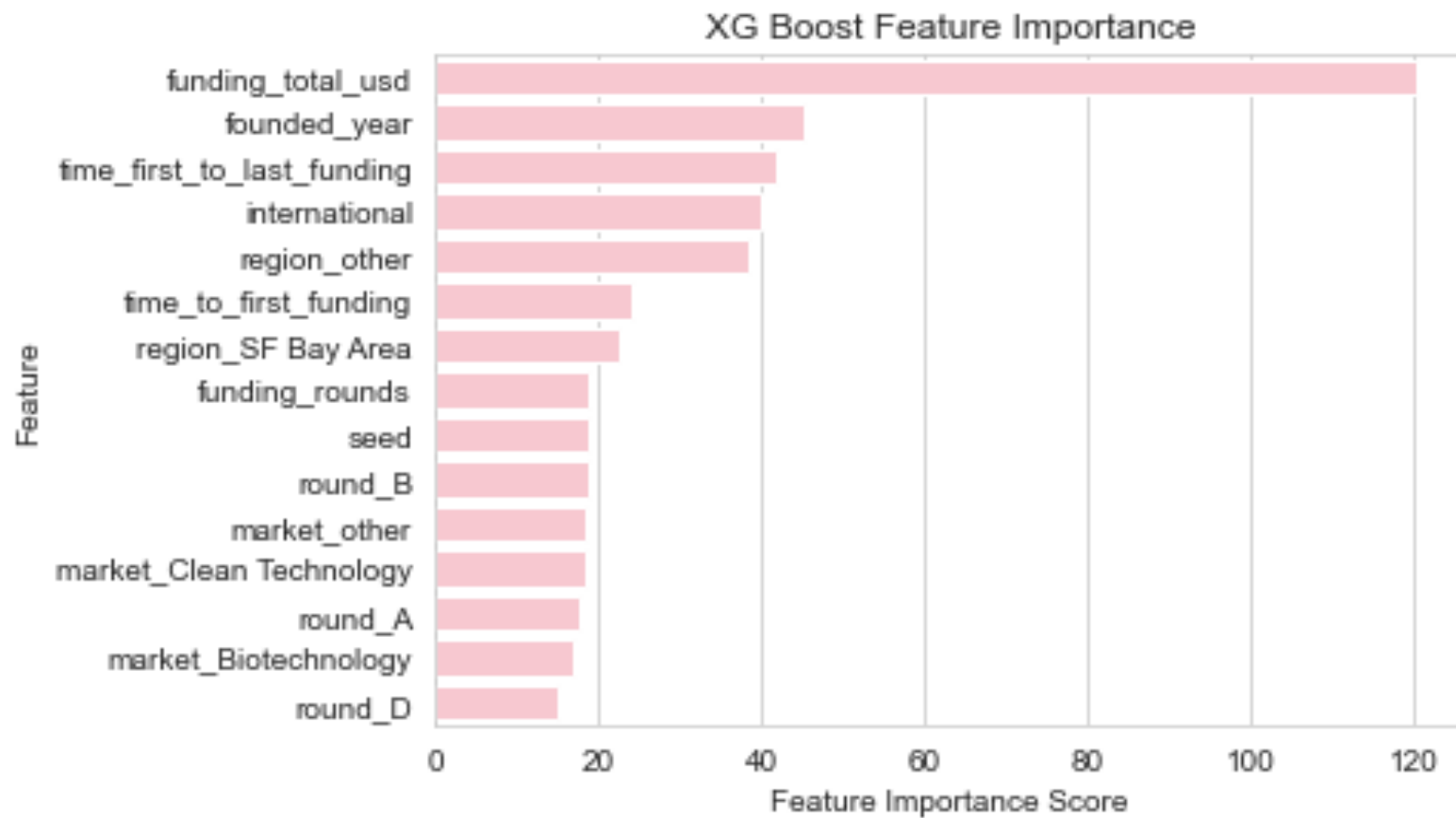- A measure of how good the model is at distinguishing between success and failure

72% Accuracy
- Achieved 72% Accuracy

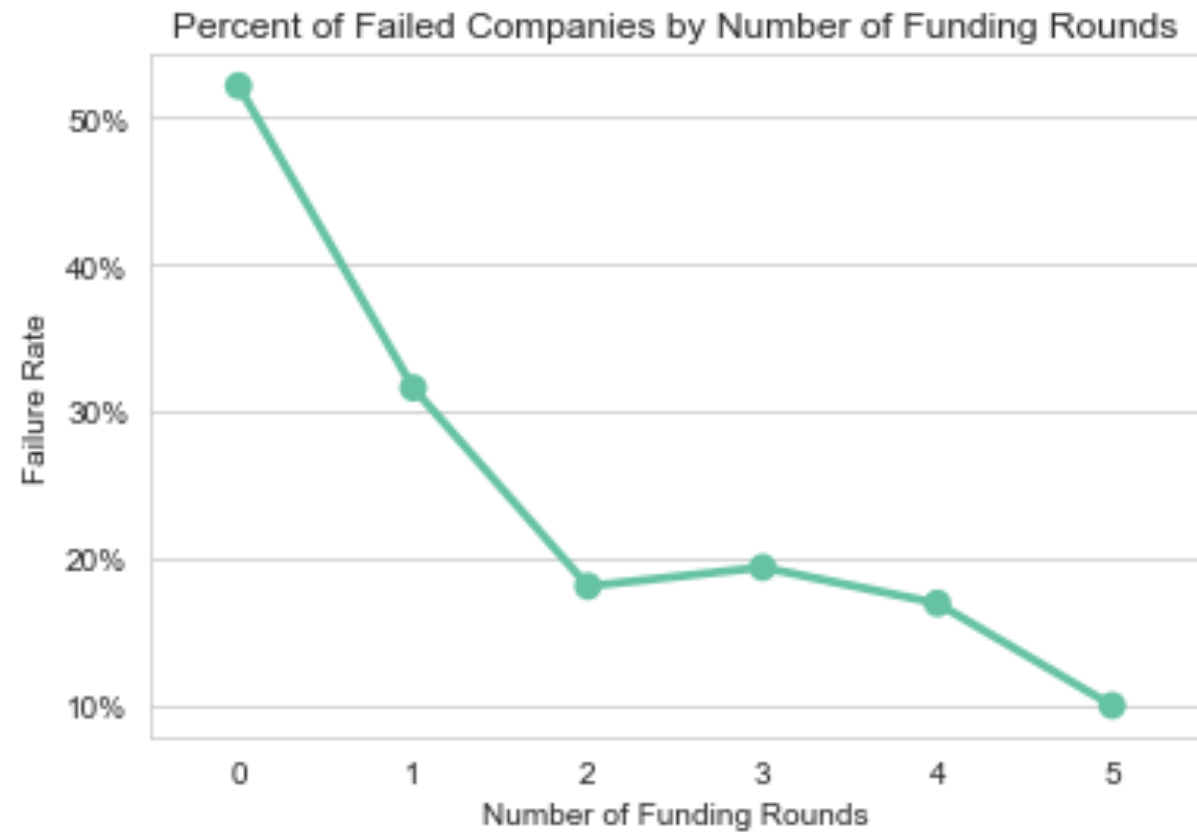# Final Model: XG Boost

XG Boost Feature Importance

Conclusion

- The final classification model can correctly identify **80%** of the companies that failed

- The most influential features were total **funding**, **year** founded, **time** passed from first funding round to last, and whether a company was **U.S. base**
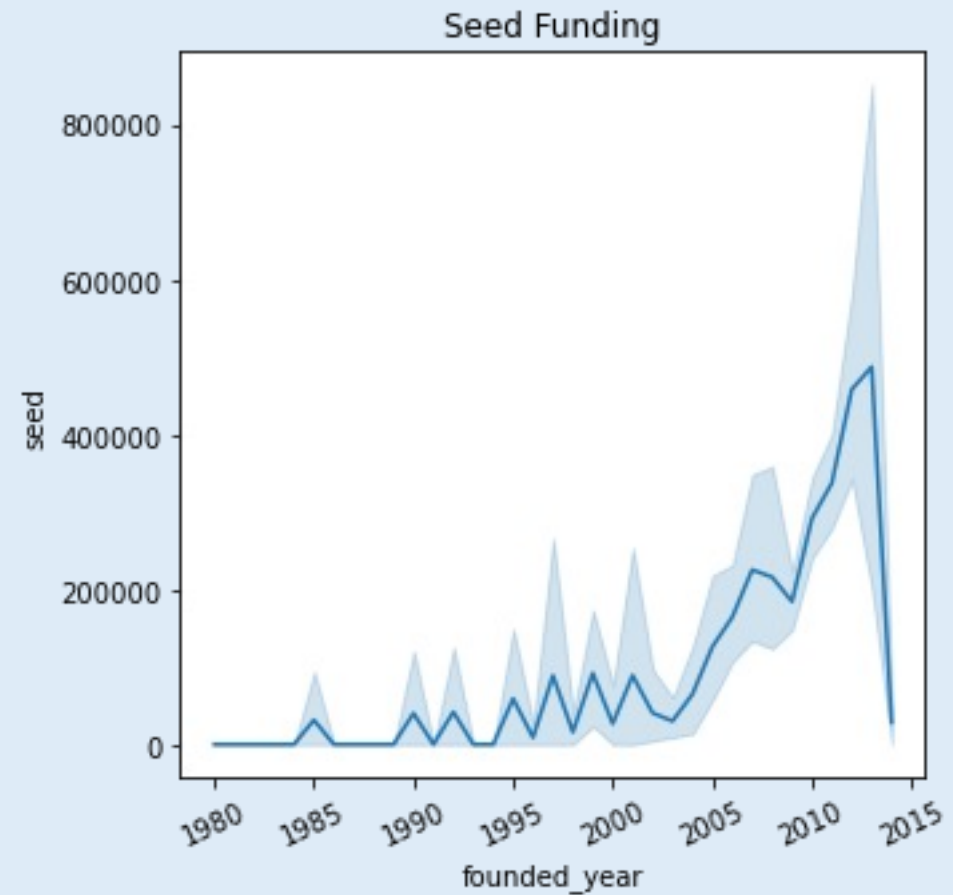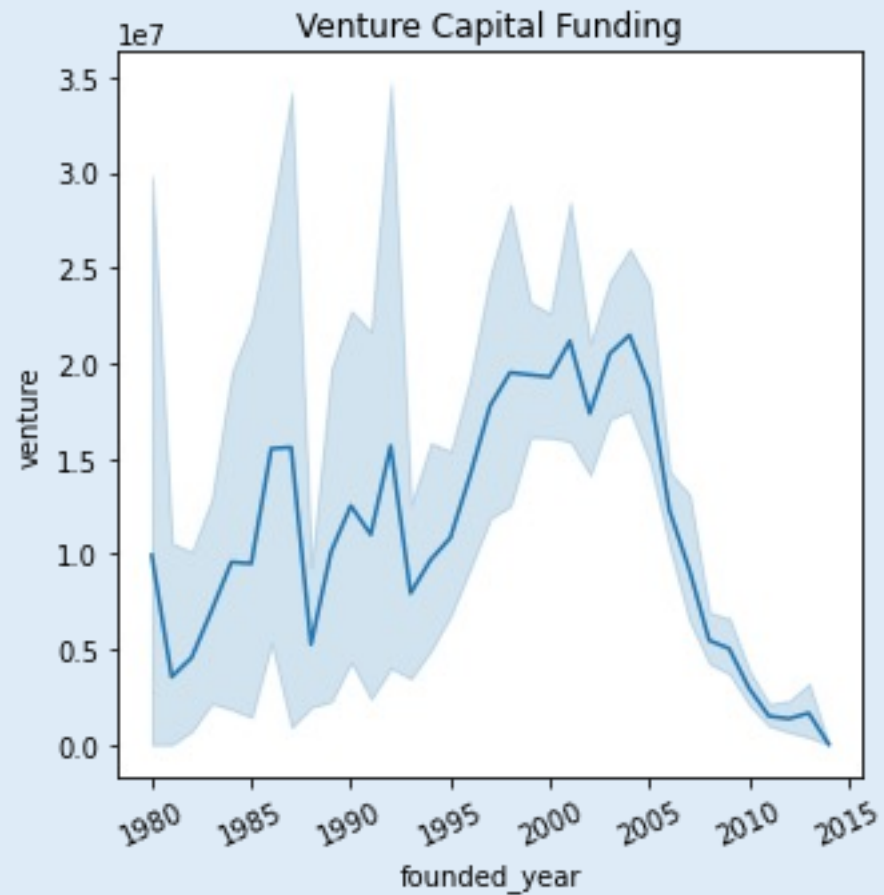
Funding First

Percent of Failed Companies by Number of Funding Rounds

Timing Matters

Location, Location, Location

Next Steps

- *Source more recent data*

- *Use economic indicators to improve prediction*

- *Separate U.S. and international companies, creating separate models for each category*

If any questions, you can reach me at ntdoris2@gmail.com. Thank you!