Team 30

(HOW EARLY)

CAN WE PREDICT BANK FAILURE?

Team 30

MINYOUNG RHO

XINYI XU

NAOMI CANNELL

DESSISLAVA VELTCHEVA

ELISA IMPARA

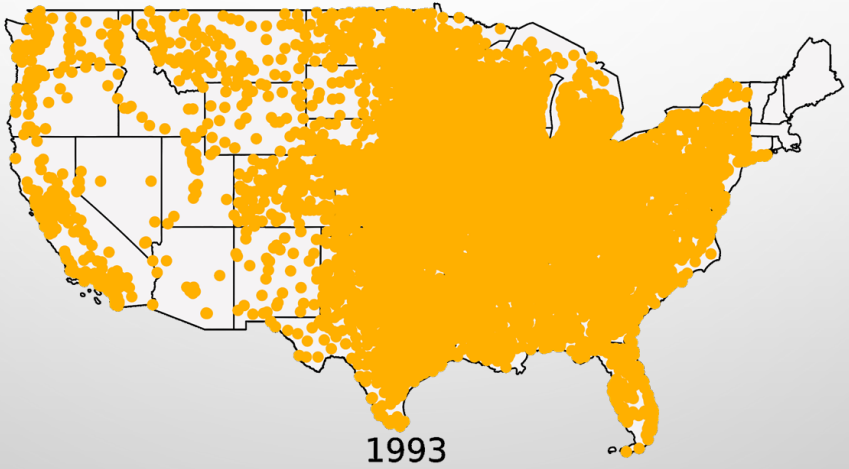
ESIN ONAL

AVNI MALHOTRA

MOTIVATION

- Banks are closely linked to the health of an economy

US bank decline from 1993 to 2021



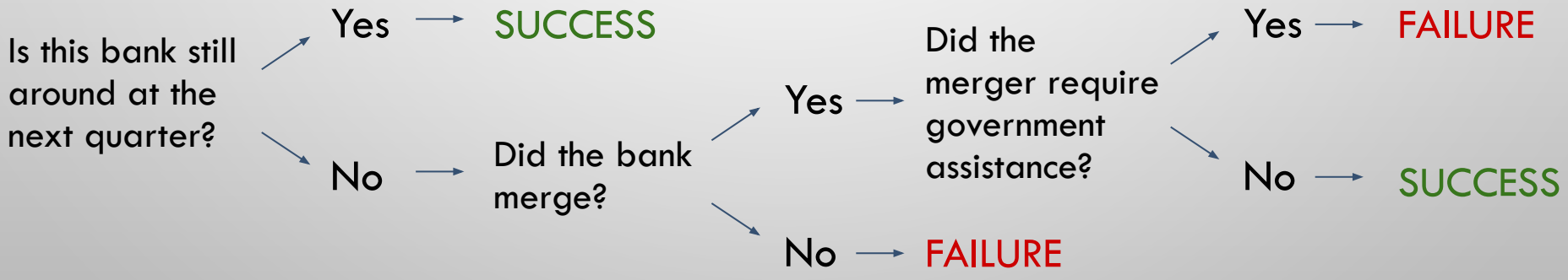
- Can we predict whether a bank will likely to fail?
- How early can we predict failure?

DATA: y

- SUCCESS vs. FAILURE at each quarter

Data from **ALL** (1993 - 2015):

- Mergers and acquisitions (M&A) of banks across the US
- Quarterly call reports submitted by all US commercial banks (assets, liabilities, etc.)



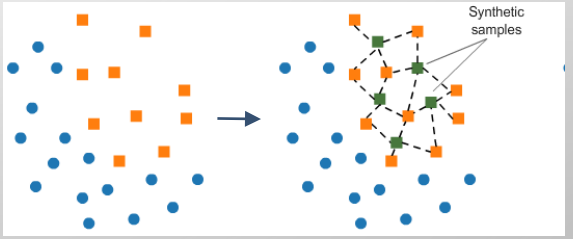
DATA: X

Financial Condition of bank from **Quarterly call reports submitted by all US commercial banks (assets, liabilities, etc)**

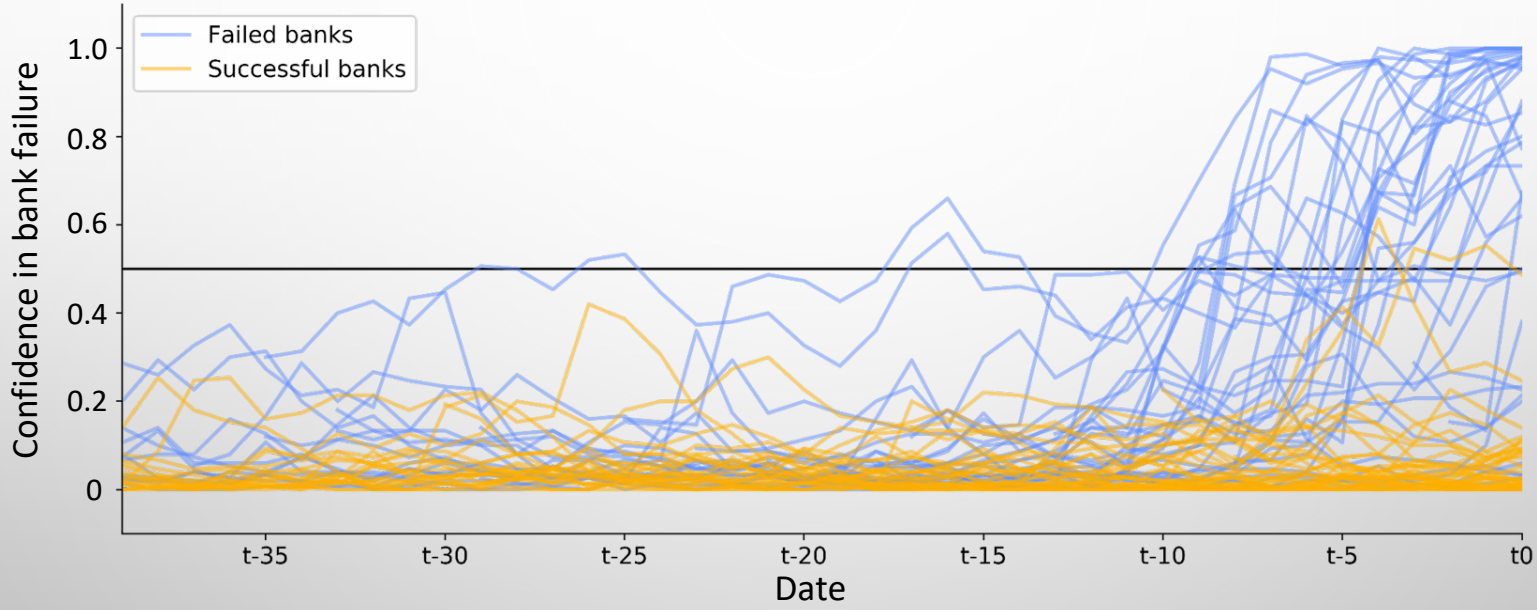
- Use within-bank ratios to avoid having to normalise or control for inflation.
- For each bank at each timepoint we use variables from the previous 5 quarters as predictors.

Bank ID	**Date**	**ROA *t0***	**ROA *t-1***	**...**	**NCO *t0***	**NCO B *t-1 ...***
56255	**06/1993**	-0.7	0.4		0.08	0.55
56255	**09/1993**	0.6	-0.7		0.13	0.08
83943	**09/1993**	4.8	0.3		0.74	0.27
83943	**12/1993**	1	4.8		0.24	0.74

CHALLENGES

Challenge	Solution
Conventional success metrics could not be used because failed banks are “successful” until the point of fail	Visualize the likelihood of failure at each time period
Traditional train-and-test split on our dataset would contain data on bank & quarter level, but we want this to be on a bank level	Test: selected 40 successful and 40 failed banks over their whole lifetime Train: rest of the data points
Imbalance on classifier	SMOTE oversampling method to produce balanced data 

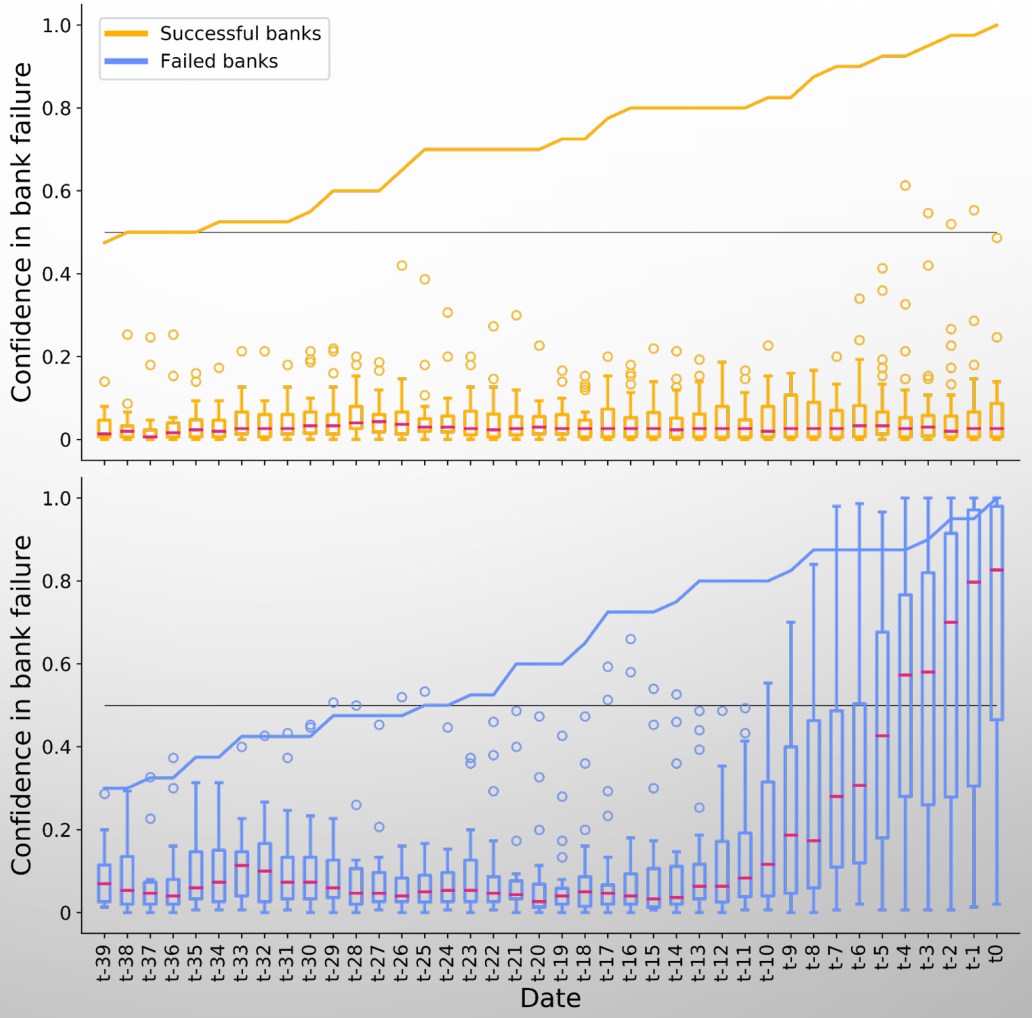
RANDOM FOREST PREDICTION OF BANK FAILURE



- Training set: 60,000 samples (1:1 ratio of fails and successes)
- Validation set: 3102 samples (lifespan of 80 banks)

PREDICTION POSSIBLE!

- Prediction of failure **IS POSSIBLE**
- On average we can predict failure 5 quarters in advance



CONCLUSION & FUTURE WORK

- Predicting bank failure:
  - 5 quarters prior to failure date
  - Challenges → complex data

- Future work:
  - *HOW CAN WE IMPROVE OUR RESULTS?*
    Adding more features, including:
    - Growth/decline (differencing the lagged variables) of a variable
    - Averaging the variable over time

  - *CAN WE EXPLAIN WHAT FACTORS LEAD TO FAILURE?*
    - Use of other classifiers (e.g. logistic regression)
    - Clustering techniques to understand traits of successful and unsuccessful banks
    - Interpret the results from random forest

MINYOUNG RHO

XINYI XU

NAOMI CANNELL

Thank You!

DESSISLAVA VELTCHEVA

ELISA IMPARA

ESIN ONAL

AVNI MALHOTRA