

**Deloitte.**



**WPI**

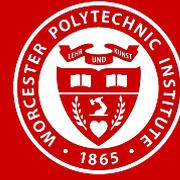
# Forecasting Bank License Revocation

Jacob Bortell

Jack Harding

Mike Giancola

Parmenion Patias



## Acknowledgements

- ◉ Dr. Jaroslav Bologov
- ◉ Dr. Alexey Kozionov

- ◉ Dr. Oleg Pavlov
- ◉ Dr. Thomas Balistriero

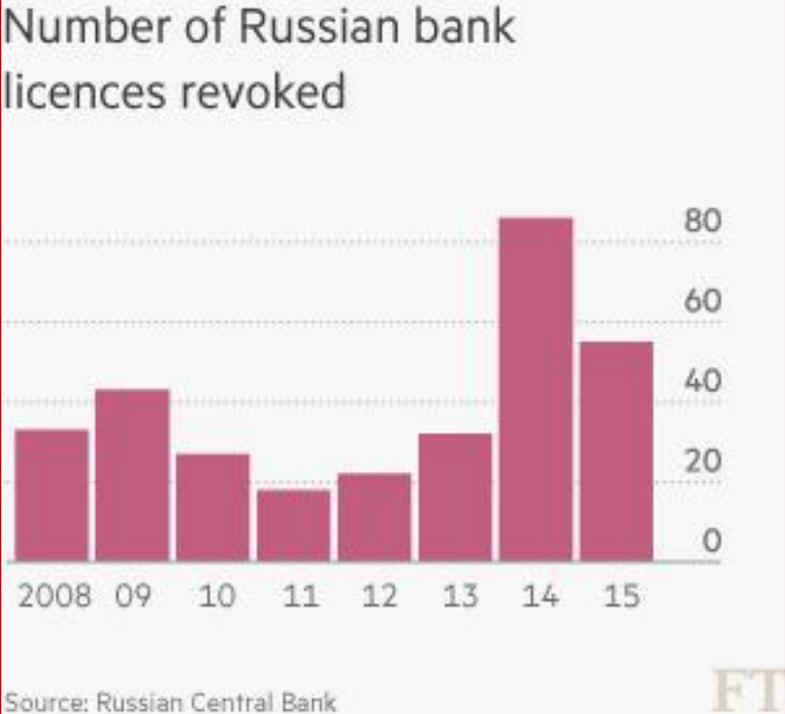
## Outline

1. Background
2. Methodology
3. Results & Analysis

## The Problem

- Bank licenses by Central Bank of Russia
- Difficult to forecast license revocation
- Affects general public

- 2013 CBR Reforms



- 2013 CBR Reforms
- 2014 Oil Crisis



- 2013 CBR Reforms
- 2014 Oil Crisis
- Russian Wages



## The Goal

*Forecast the likelihood of license revocation*

## Central Bank of Russia (CBR)

- Banking Standards
  - N1 - Capital Adequacy Ratio
  - N2 - Instant Liquidity Ratio
  - N3 - Short-Term Liquidity Ratio

## Dataset Structure

Lic Num	Date	N1	N2	N3	---	Quarters
127	2015-10-1	...	...	...	...	3
2288	2014-3-1	...	...	...	...	> 2 years
564	2015-7-1	...	...	...	...	Active

## Building Models

- Build models which
  - Analyze financial data
  - Learn from data to make predictions
- Two Types:
  - Statistical: Logistic Regression
  - Machine Learning: Random Forest

- Split data
  - Learn from  $\frac{2}{3}$  of data

Lic Num	Date	N1	N2	N3	---
127	<del>2016-1-1</del>	13	0.4	5	...

Quarters
3

2288	<del>2014-3-1</del>	11	0.5	2	...
------	---------------------	----	-----	---	-----

> 2 years
-----------

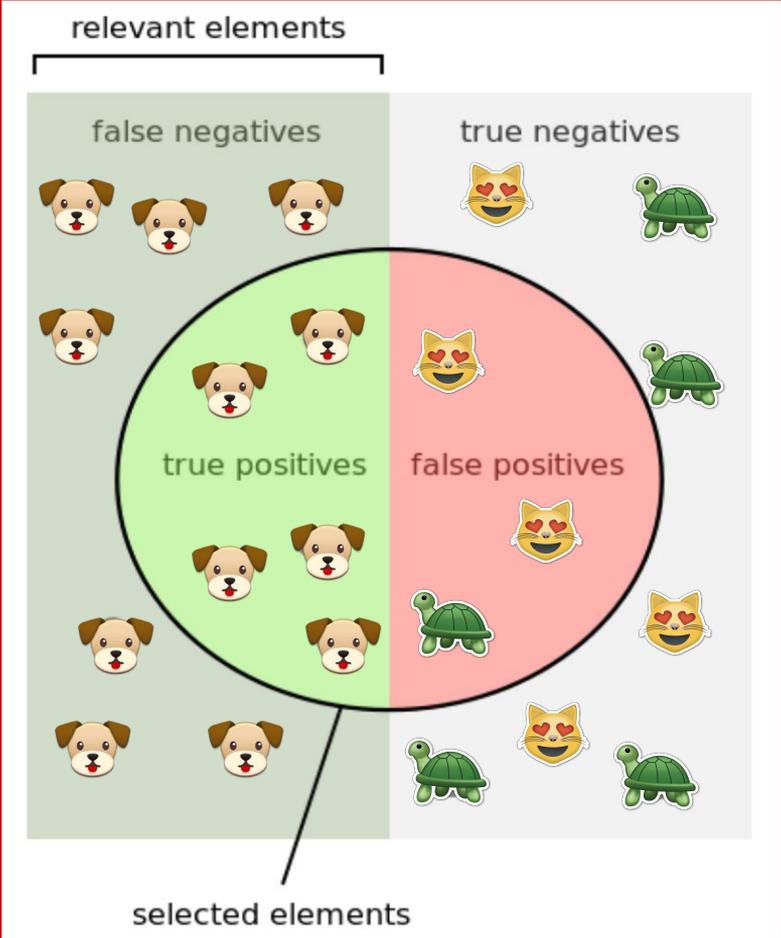
- Predict from  $\frac{1}{3}$  of data

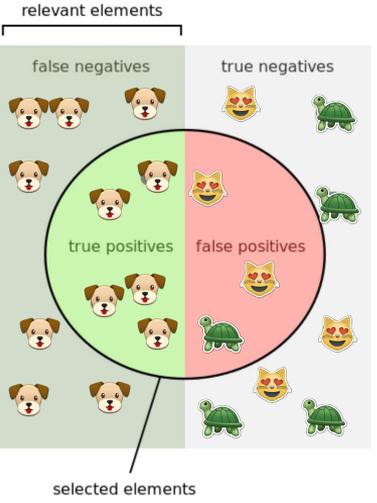
564	<del>2015-7-1</del>	14	0.6	4	...
-----	---------------------	----	-----	---	-----

?
---

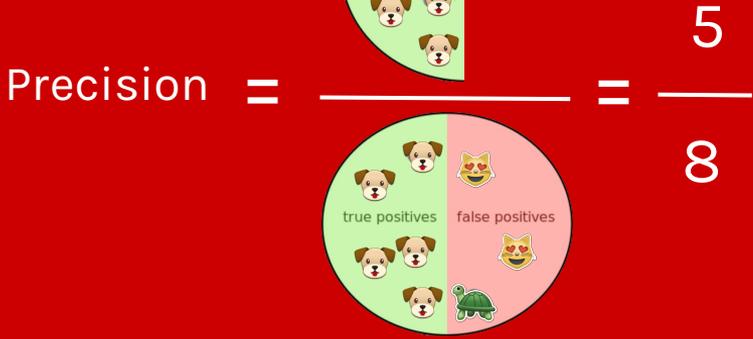
- Compare predictions to known values

# Performance Metrics

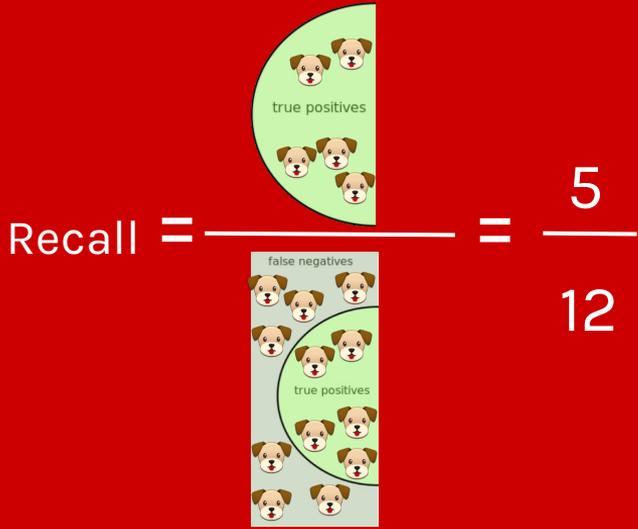




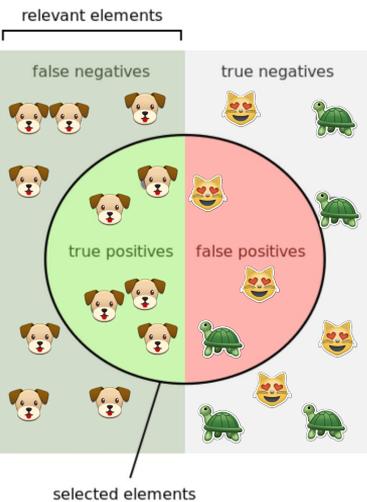
Precision =  $\frac{5}{8}$



Recall =  $\frac{5}{12}$



# Performance Metrics



$$F1 = \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} * 2$$

$$F1 = \frac{5/8 * 5/12}{5/8 + 5/12} * 2$$

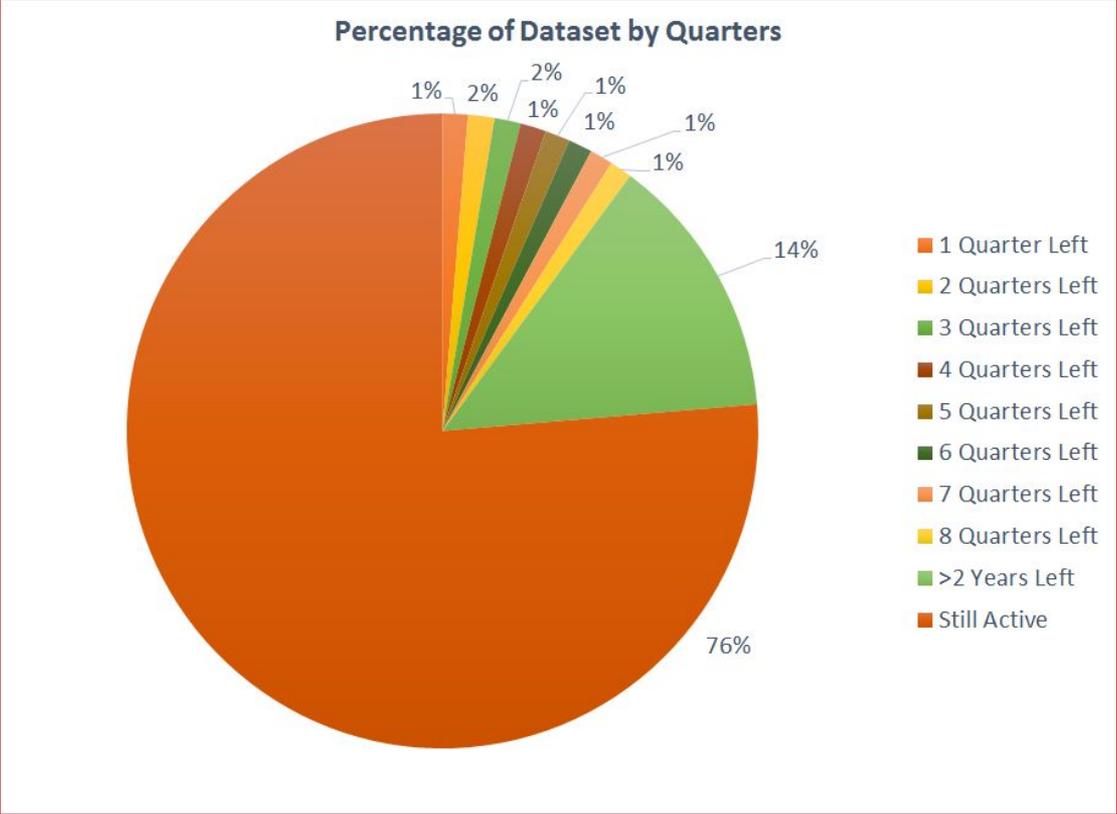
$$F1 = 1/2 = 50\%$$

# Performance Metrics

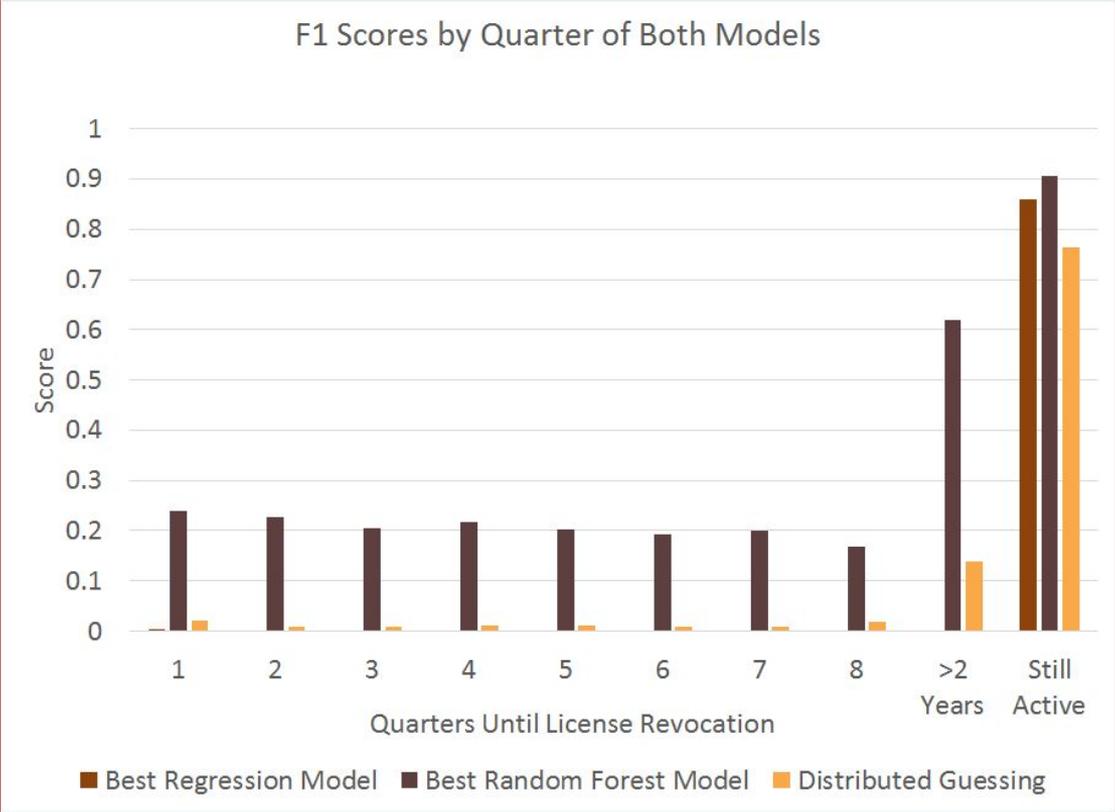
## Results

- Dataset Composition
- Performance Based on F1
- Significance of Features
- Recommendations

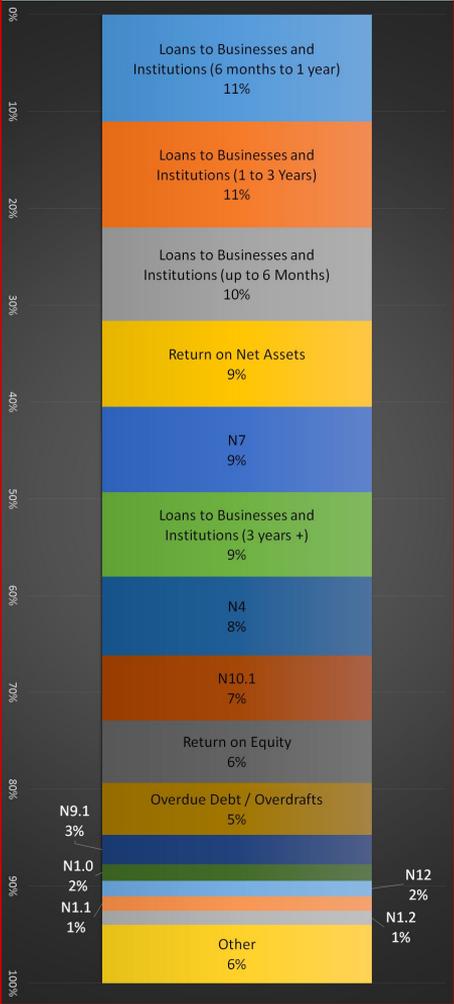
- “Still Active” banks majority of dataset
- First 8 quarters only 10% of dataset



- Regression outperformed by guessing
- Random Forest outperforms Regression



# Significant Factors





Loans to Businesses and  
Institutions (6 months to 1 year)  
11%

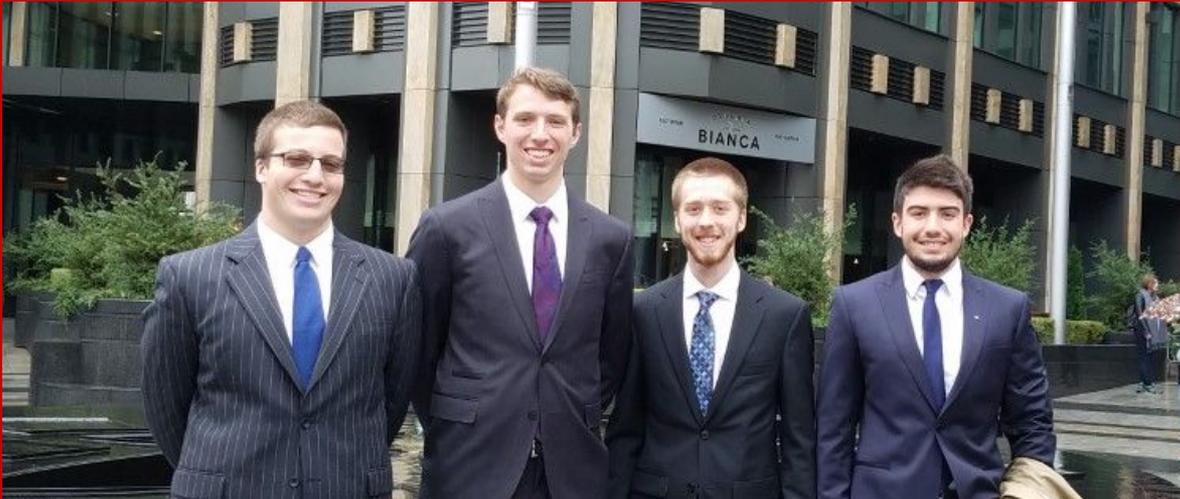
Loans to Businesses and  
Institutions (1 to 3 Years)  
11%

Loans to Businesses and  
Institutions (up to 6 Months)  
10%

# Significant Factors

## Recommendations

- Use Random Forest Model
- Research Clustering
- More Features
  - Risk Related
  - Socio-political
- Fewer Classes / More Granular Data
- Benefits
  - Banks
  - Investors
  - General Public



Mike  
Giancola

Jack  
Harding

Jacob  
Bortell

Parmenion  
Patias

**Deloitte  
Team**