

The Role Partnerships between Community Banks and Fintech Companies Played in PPP Loan Distribution

Gio DeFrank, Junghee Mun, Kristina Posner



Raymond A. Mason
School of Business
WILLIAM & MARY

Table of Contents

01

**Hypothesis
Overview**

02

**Related Literature
and Research**

03

Data Sources

04

**Data Collection and
Cleaning Methodology**

05

**Model & Methodology:
Impact of Fintech
Partnerships on Distribution
of Loans to Minorities**

06

**Model & Methodology:
Impact of Fintech
Partnerships on Distribution
of Loans to Low and
Moderate Income areas**

07

**Further Discussion &
Conclusion**



01 Hypothesis Overview



Our Hypothesis:

Community Banks who partnered with financial technology (Fintech) firms were ultimately more successful in distributing PPP loans to small businesses with racial minorities and underbanked populations.



**Related
Literature and
Research**

02

Current Literature Overview



Community Banks

- Community Banks played an outsized role distributing PPP loans to small businesses during the pandemic

(Community banks play outsized role, 2020)
(Erel & Liebersohn, 2020)



Fintechs

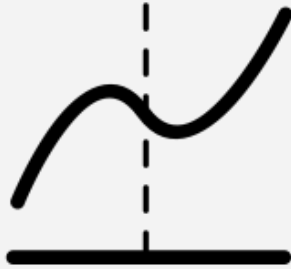
- Fintechs are more likely to reach minority and underbanked populations than conventional banks



Partnerships

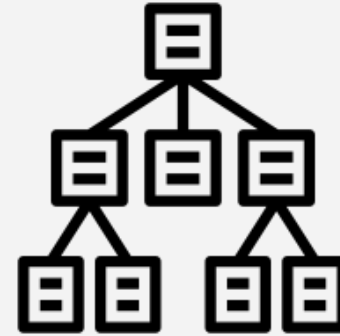
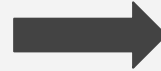
- Anecdotal evidence suggests Community Bank and Fintech partnerships played a large role in distributing PPP loans to disadvantaged populations during the pandemic

Model Summary and Justification



Logistic Regression

- Implemented logistic regression to evaluate the impact a Fintech Partnership had on the odds of a PPP loan being distributed to minority & LMI labeled businesses
- Our data violated half of the required assumptions of logistic regression



Random Forest

- Implemented random forest model to validate/test the results of the logistic regression model
- Random forest models do not have assumptions
- Random forest is an intuitive and easy-to-understand model
- Achieved accuracy rate of **67%** for predicting minority owned businesses and **70%** accuracy for predicting businesses located in LMI areas



03

Data Sources

Data Sources



PPP Loan Data

- Approximately 8 million rows
- Provided loan level information
- Formed the core of our analysis



FDIC Community Bank Designation Data

- Requirements that define Community Banks
- Additional information about each bank's financials



FDIC Yearly "How America Banks" Survey

- FDIC data about America's unbanked population
- Percent unbanked vs percent banked
- Confidence intervals for provided statistics



**Data
Collection
and Cleaning
Methodology**

04

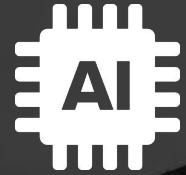
Our Data-Cleaning/Collection Methodology



Web scraping

alteryx

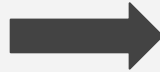
Data Cleaning
Workflow



Classification
Neural Network

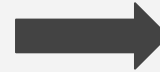
Step 1

- Used Python script to web-scrape Google for articles that mention Fintech & Community Bank partnerships



Step 2

- Cleaned, combined, and summarized data sources using Alteryx



Step 3

- Built a classification neural network model to predict the probability of a business being minority owned. Used predicted values to fill in NAs within the minority column

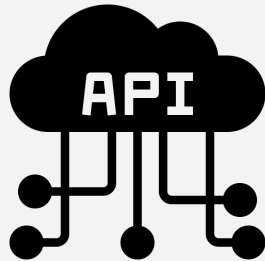
Web Scraping Process

Utilized SerpAPI to
web-scrape Google



Read each article,
validated the
partnership, and
generated a list of 41
banks that had
Fintech partnerships

237K



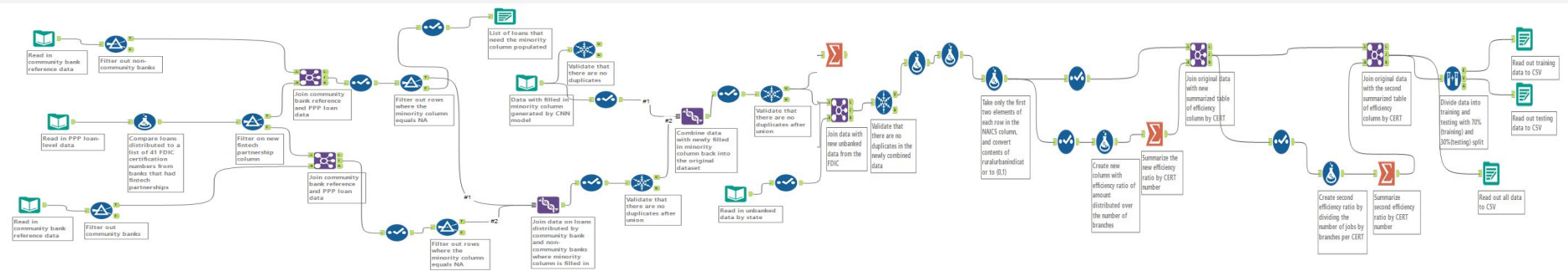
Collected over 100
links that mentioned
Fintech + Community
Bank partnerships



Identified 237K loans
that were likely
distributed by
Community Bank +
Fintech partnerships

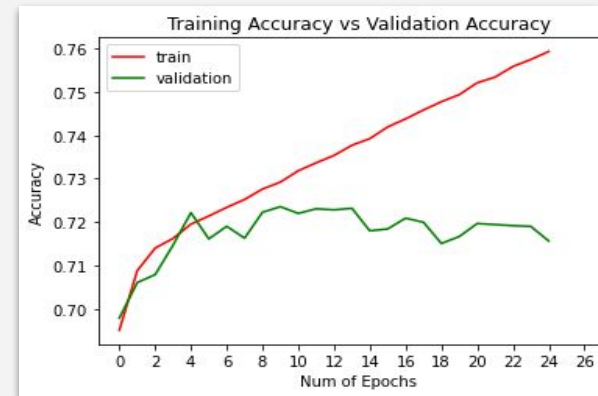
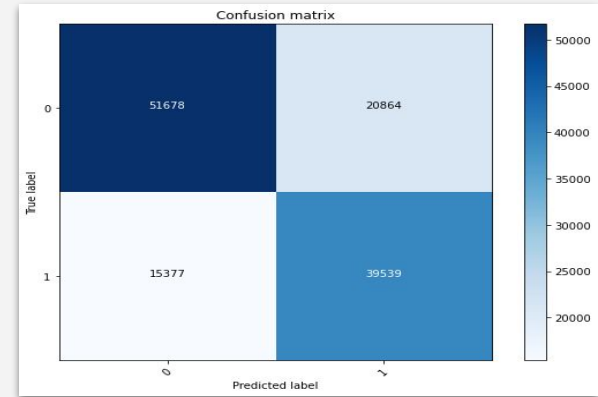
Alteryx: Preprocessing and Data-Cleaning

- Cleaned Data
 - Combined various data sources
 - Cleaned data
 - Filled in NAs
 - Summarized data
 - Split data into train and test sets



Classification Neural Network Model

- Problem:
 - The minority column contained a large amount of empty rows
- Feedforward model:
 - Experimented with different number of layers and nodes
 - Experimented with various learning and momentum rates
- Final model:
 - 6 hidden layers (varying number of nodes within each layer)
 - Learning rate of 0.2
 - Momentum rate of 0.7
 - Accuracy of **72%**





05

**Minority
Model
Methodology**

Logistic Regression Assumptions:

- #1: Linearity of the Logit
- #2: Absence of Multicollinearity
- #3: Lack of Strongly Influential Outliers
- #4: Independence of Errors

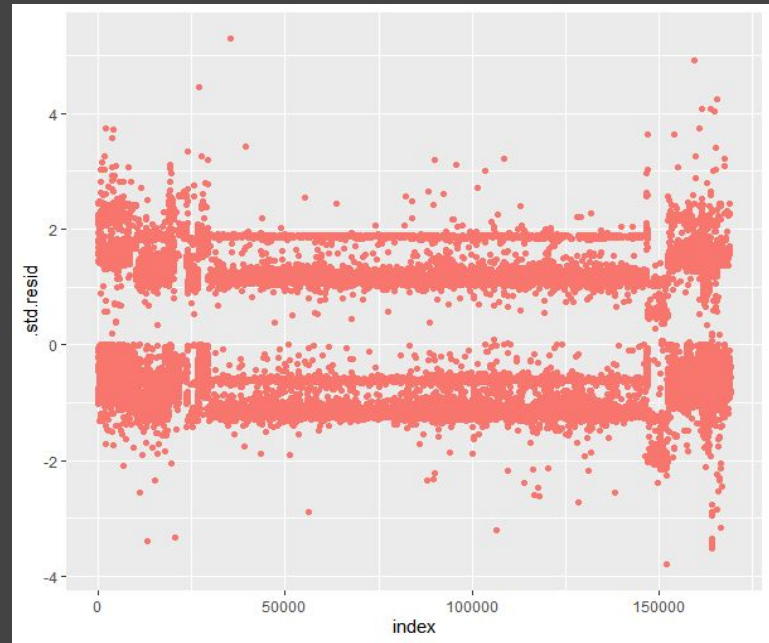


Figure 5

Classification Random Forest Minority Model

- Utilized the number of trees with highest accuracy: **300**
- Overall accuracy of **66.83%**

Random Forest Models	
Number of Trees	Accuracy
100	66.73%
200	66.76%
300	66.83%
400	66.77%
500	66.77%

Figure 6

Importance Report: Minority Model Results

- Fintech Partnership has the lowest importance in predicting PPP Borrower as Minority
- Fintech Partnership doesn't contribute to increasing Minority outreach

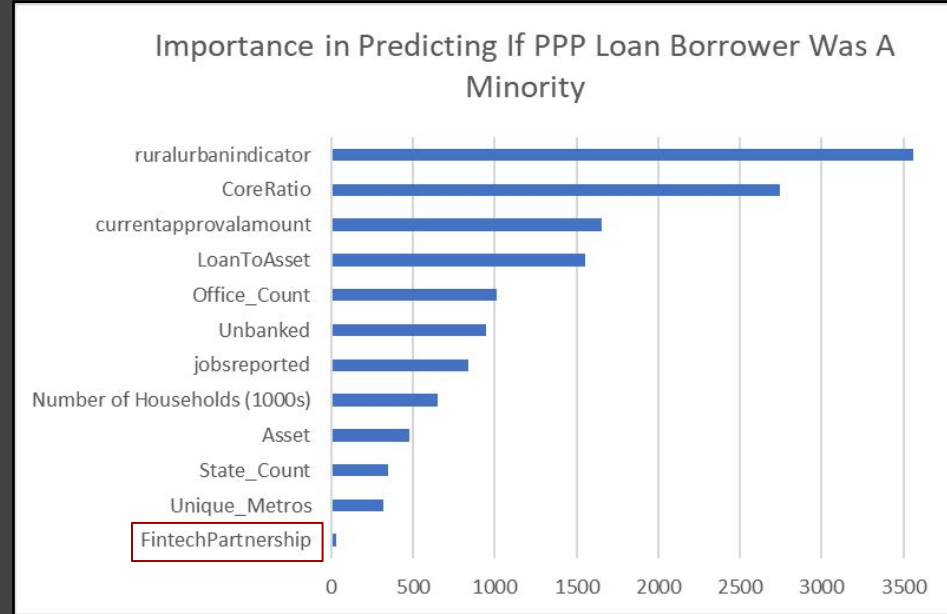


Figure 7



06

**Low and
Moderate
Income (LMI)
Model
Methodology**

Classification Random Forest LMI Model

- Utilized the number of trees with highest accuracy: **100**
- Overall accuracy of **69.864%**

Random Forest Models	
Number of Trees	Accuracy
100	69.864%
200	69.859%
300	66.856%
400	55.859%
500	66.857%

Figure 8

Importance Report: LMI Model Results

- Fintech Partnership has the lowest importance in predicting PPP Borrower as LMI
- Fintech Partnership doesn't contribute to increasing LMI outreach

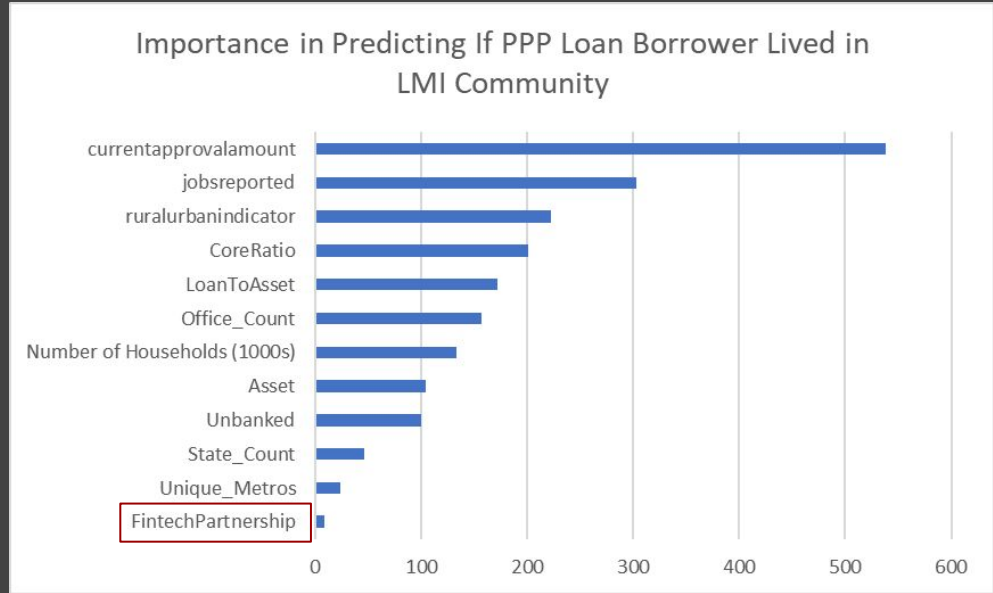


Figure 9



**Further
Discussion &
Conclusion**

07

Further Possibilities

1

2

3

4

1. Further data gathering & cleaning

- Lack of publicly available data
- Limited options of metrics and predictors

2. Web scraping

- “Brute forced” list of Fintechs from scrapings
- Potential user error and data leakage

3. Minority prediction using ANN classifier

- Dependent on predictive library’s validity
- Accuracy cap: 72%

4. Neo-bank classification

- Dependent on FDIC’s ruling update
- Good precedent for future evolution of banks

Implications

- Fintech partnership was not a significant factor for community banks handling PPP loans
- However, may help acceleration with banks' digitalization
- Caution for hasty digitalization without establishing security measures and/or compliance protocols
- Safety measures and guidelines should be mandated by governmental bodies

Conclusion



Hypothesis

- Community bank's partnership with fintech did not help to reach minority and LMI population with PPP loans in 2020

Fintech Partnership

- Help to diversify the banking ecosystem
- Need for adequate regulations is ever-present
- More research is needed

Best Model

- Random forest classification
- Accuracy: **70%**



Thank you

Sources

1. Bowman, M. W. (2020, September 30). *Speech by governor Bowman on community banks rising to the Challenge*. Board of Governors of the Federal Reserve System. Retrieved May 2, 2022, from <https://www.federalreserve.gov/newsevents/speech/bowman20200930a.htm>
2. Conference of State Bank Supervisors. (2020, December 11). *Community banks play outsized role in PPP lending*. Conference of State Bank Supervisors. Retrieved May 2, 2022, from <https://www.csbs.org/newsroom/community-banks-play-outsized-role-ppp-lending>
3. Erel, I., & Liebersohn, J. (2020). Does Fintech substitute for banks? evidence from the Paycheck Protection Program. *NATIONAL BUREAU OF ECONOMIC RESEARCH*. <https://doi.org/10.3386/w27659>