

# Classifying Kiva Loans

Robert Fang



**Dreams are universal, opportunity is not.**

Lend as little as \$25 to create opportunity  
for people around the world.

# Crowdfunding Loans



**219,435**

Borrowers supported  
in conflict zones

**526,197**

Farmers supported

**749,149**

Borrowers supported in  
least developed countries

**28,234**

Education loans

**65k**

Borrowers gained  
access to  
clean energy

# Crowdfunding Loans



**219,435**

Borrowers supported  
in conflict zones

**526,197**

Farmers supported

**749,149**

Borrowers supported in  
least developed countries

**28,234**

Education loans

**65k**

Borrowers gained  
access to  
clean energy



# Objectives for Analysis


- Identify key characteristics
- Improve the probability of loans get funded



**22%** funded  
27 days left \$775 to go

Total loan: \$1,000  
Powered by 9 lenders

**Maria**

 Becora, Timor-Leste / General Store

\$25 ▼ Lend now

A loan of \$1,000 helps to buy additional stock for her general store and a refrigerator to keep drinks cold for customers.

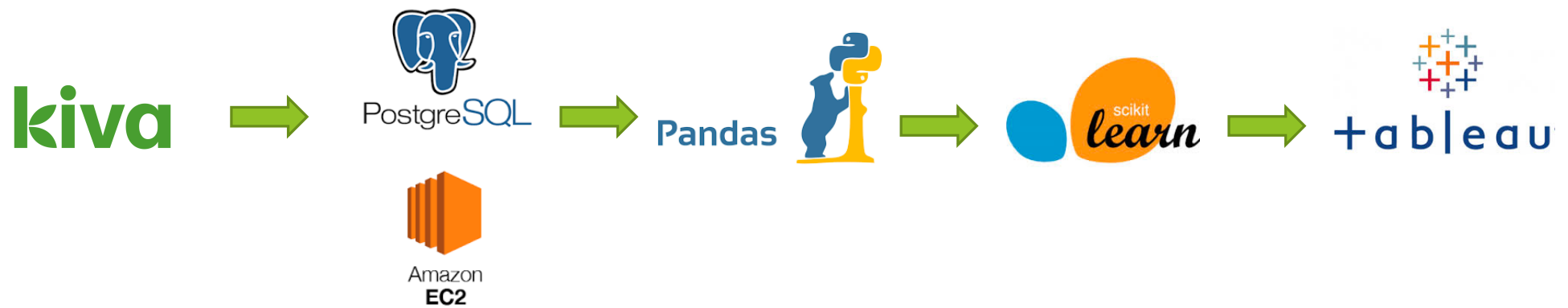
## Challenge:

- Large Dataset
- Imbalanced Classes:
  - 1.6mm loans funded vs 85k unfunded loans

## Approach:

- Oversample Minority Class
- Optimize for Recall
- Use Ensemble Modeling

# Tools



# Error Metrics for Models

Models	Accuracy	Precision	Recall	F1 Score	AUC
kNN	93%	94%	88%	91%	66%
Logistic Regression	62%	51%	16%	21%	46%
<b>Naive Bayes</b>	<b>46%</b>	<b>41%</b>	<b>98%</b>	<b>58%</b>	<b>69%</b>
Linear SVC	52%	24%	37%	29%	-
Decision Tree	76%	71%	64%	67%	77%
<b>Random Forest</b>	<b>96%</b>	<b>95%</b>	<b>96%</b>	<b>95%</b>	<b>93%</b>



Highest Recall

# Model Selection Process

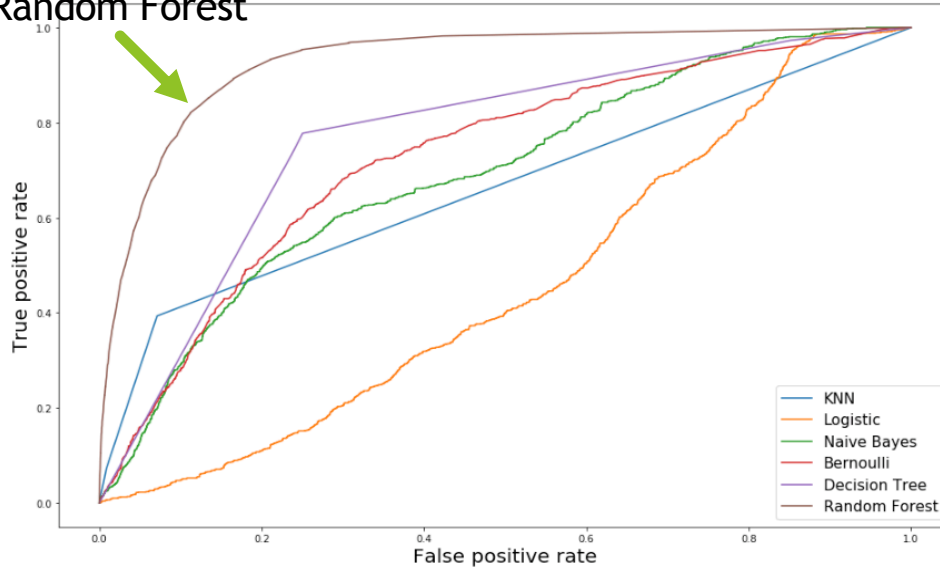
Models	Accuracy	Precision	Recall	F1 Score	AUC
kNN	93%	94%	88%	91%	66%
Logistic Regression	62%	51%	16%	21%	46%
Naive Bayes	46%	41%	98%	58%	69%
Linear SVC	52%	24%	37%	29%	-
Decision Tree	76%	71%	64%	67%	77%
Random Forest	96%	95%	96%	95%	93%

Highest AUC

Highest Recall

Random Forest

ROC Curve





# Model Selection

Models	Accuracy	Precision	Recall	F1 Score	AUC
kNN	93%	94%	88%	91%	66%
Logistic Regression	62%	51%	16%	21%	46%
Naive Bayes	46%	41%	98%	58%	69%
Linear SVC	52%	24%	37%	29%	-
Decision Tree	76%	71%	64%	67%	77%
Random Forest	96%	95%	96%	95%	93%

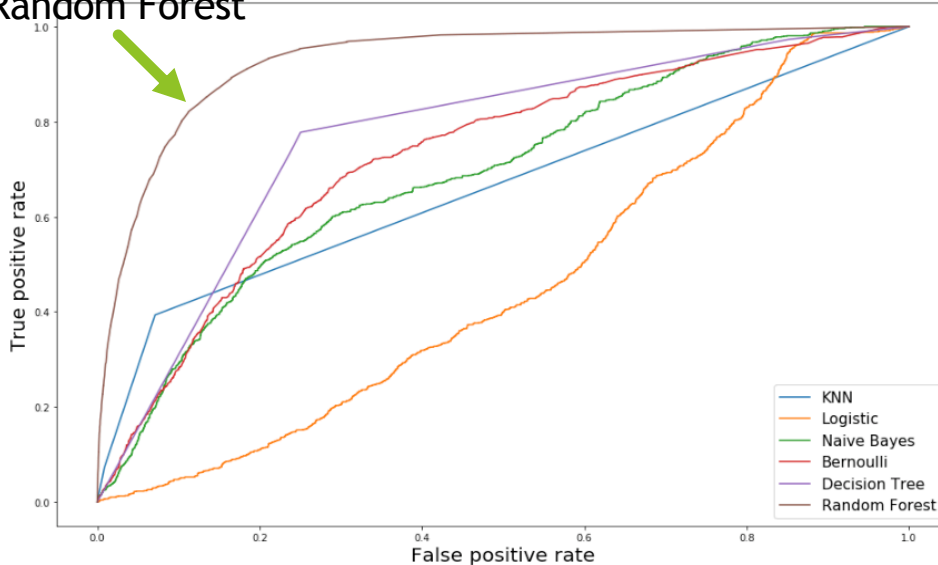
Ensemble Recall on Test Set: 46%

Highest AUC

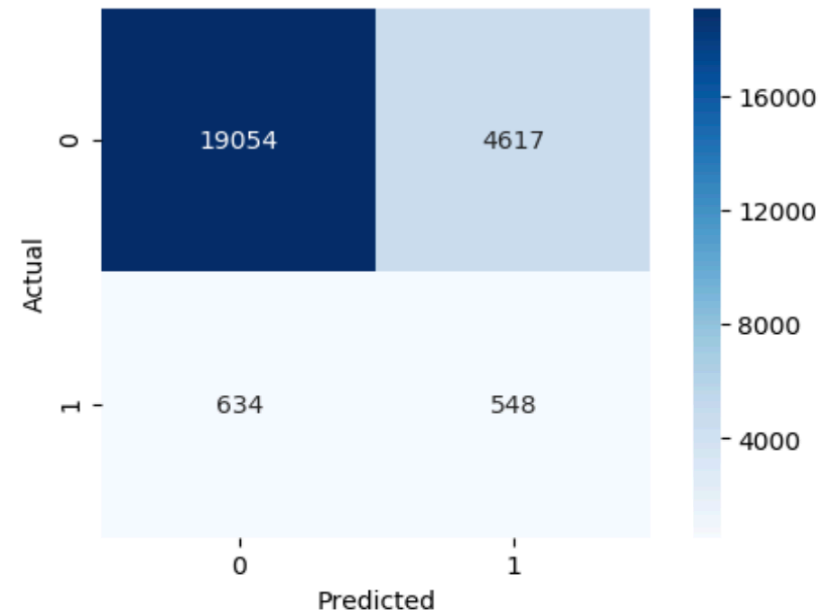
Highest Recall

Random Forest

ROC Curve



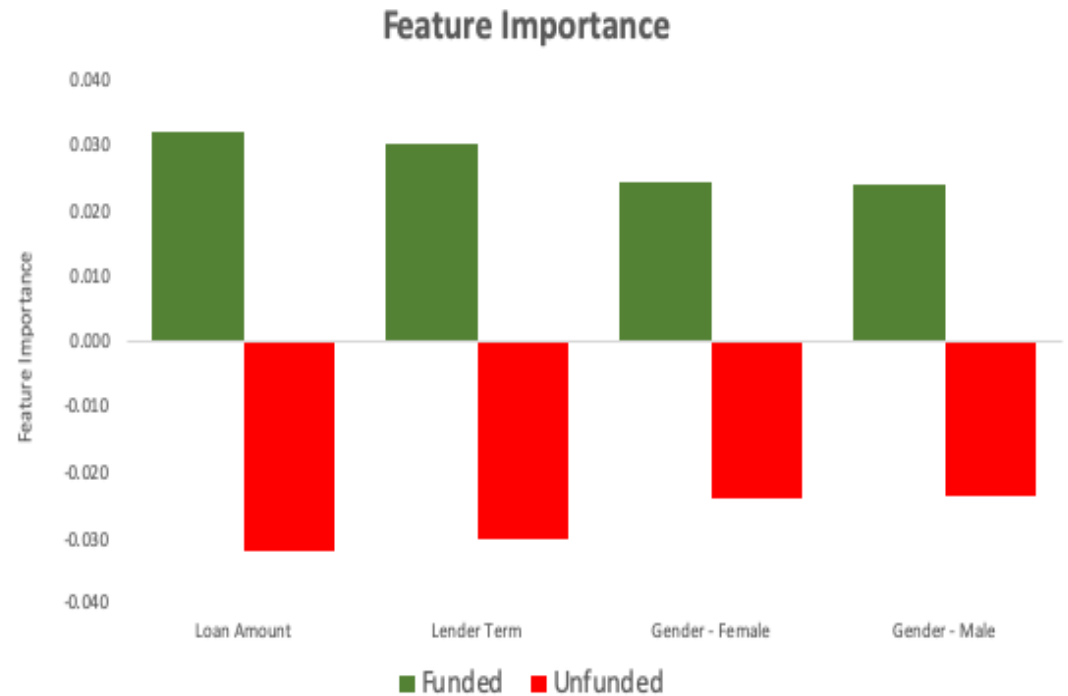
Ensemble confusion matrix



# Conclusion

## Key Features:

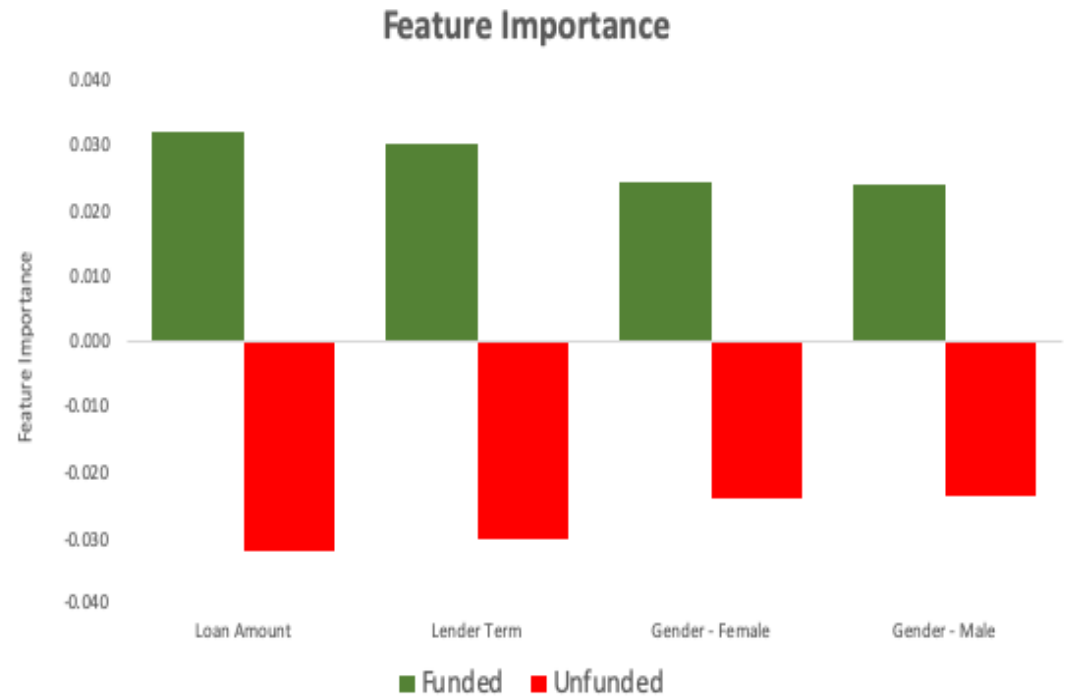
- Size of Loan
- Length of Loan
- Gender of Borrower



# Conclusion

## Key Features:

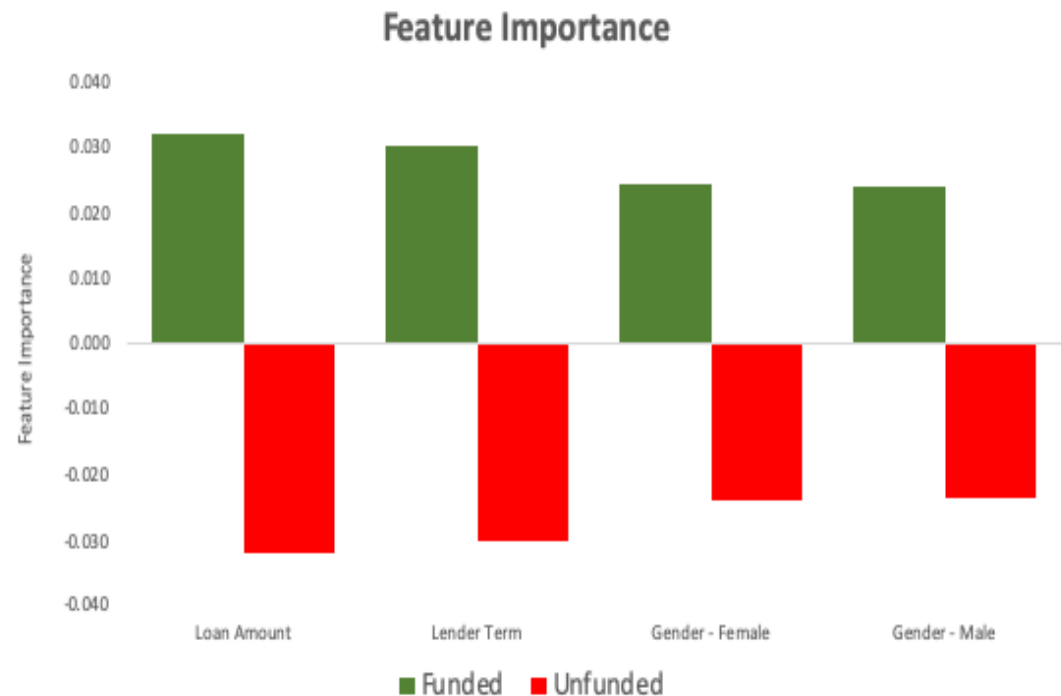
- Size of Loan
- Length of Loan
- Gender of Borrower
- Sectors to Avoid
  - Entertainment
  - Art
  - Wholesale



# Conclusion

## Key Features:

- Size of Loan
- Length of Loan
- Gender of Borrower
- Sectors to Avoid
  - Entertainment
  - Art
  - Wholesale

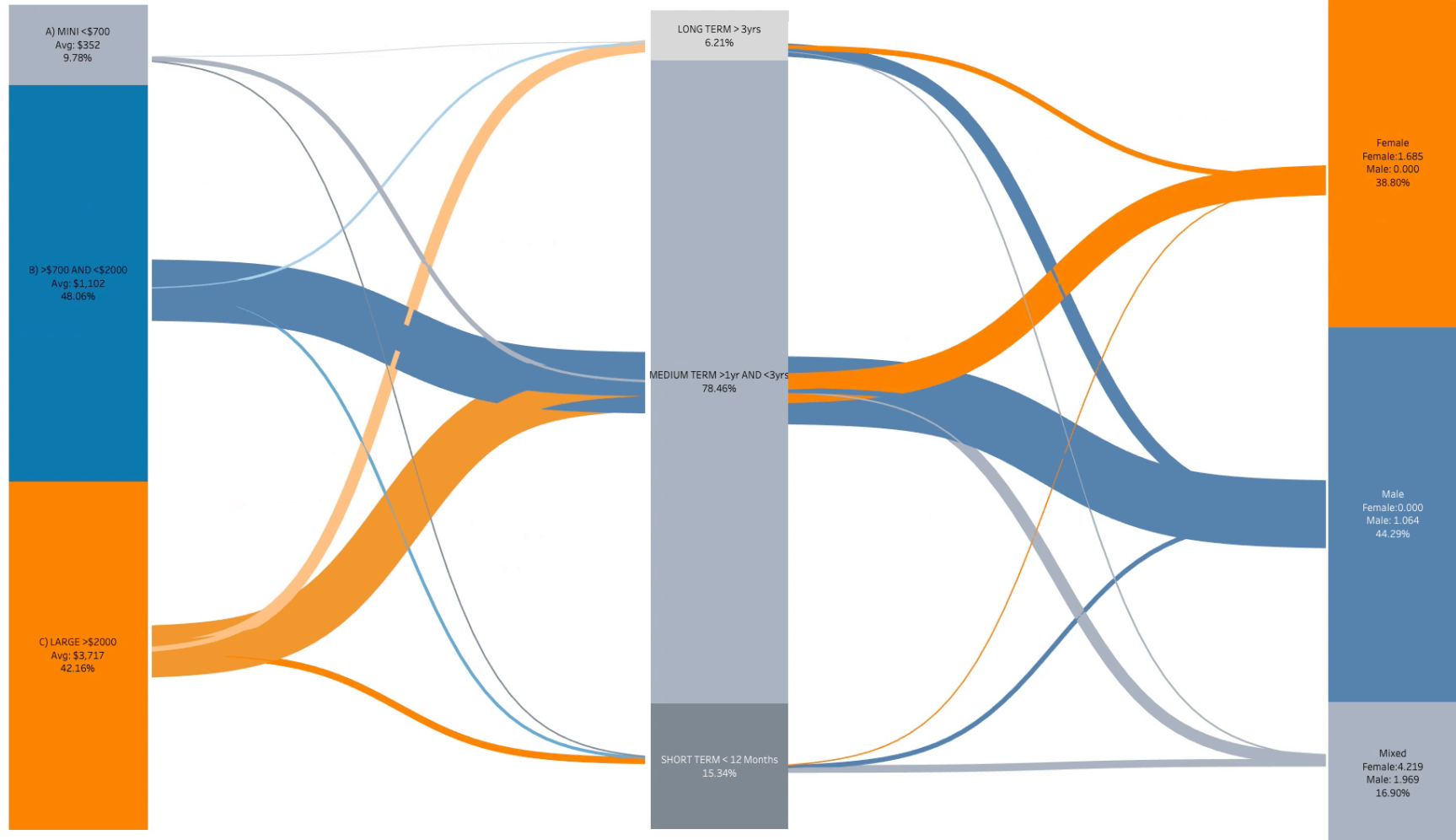


**INCREASE PROBABILITY USE SMALLER TRanches WITH SHORTER TERMS**

Loan Amount

Length of Loan

Borrower's Gender



# Thank you

---



## Questions?

<https://www.kiva.org/>

**"Kiva is a simple concept that can change a person's life."**

- Oprah

THE  
HUFFINGTON  
POST

*"In lieu of a financial profit from the money you lend, you get something that, in my opinion, is more valuable: the chance to make a tangible difference in the world and to impact a stranger's life in a positive, permanent way. Kiva is not about investing money in the traditional sense; it's about using your excess cash to jumpstart a life somewhere else in the world."*



*"The loans typically back individual economic activity to help alleviate poverty through support for entrepreneurial ventures in places where other sources of funding are hard to come by."*