Predicting Client Churn for *Shield International Bank*

What is Client Churn?

What is Client Churn?

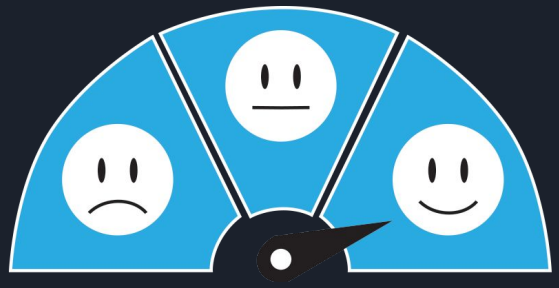
A Client stops using your services

=

A Client Churns

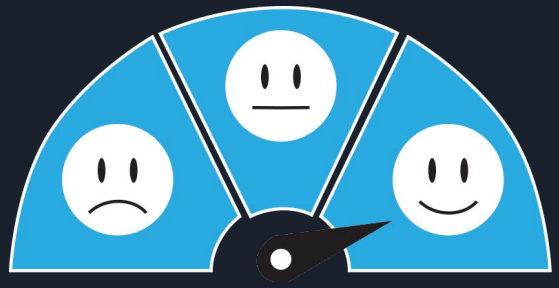
Why Predict Client Churn?

Why Predict Client Churn?



Win clients back ***before*** they leave

Why Predict Client Churn?

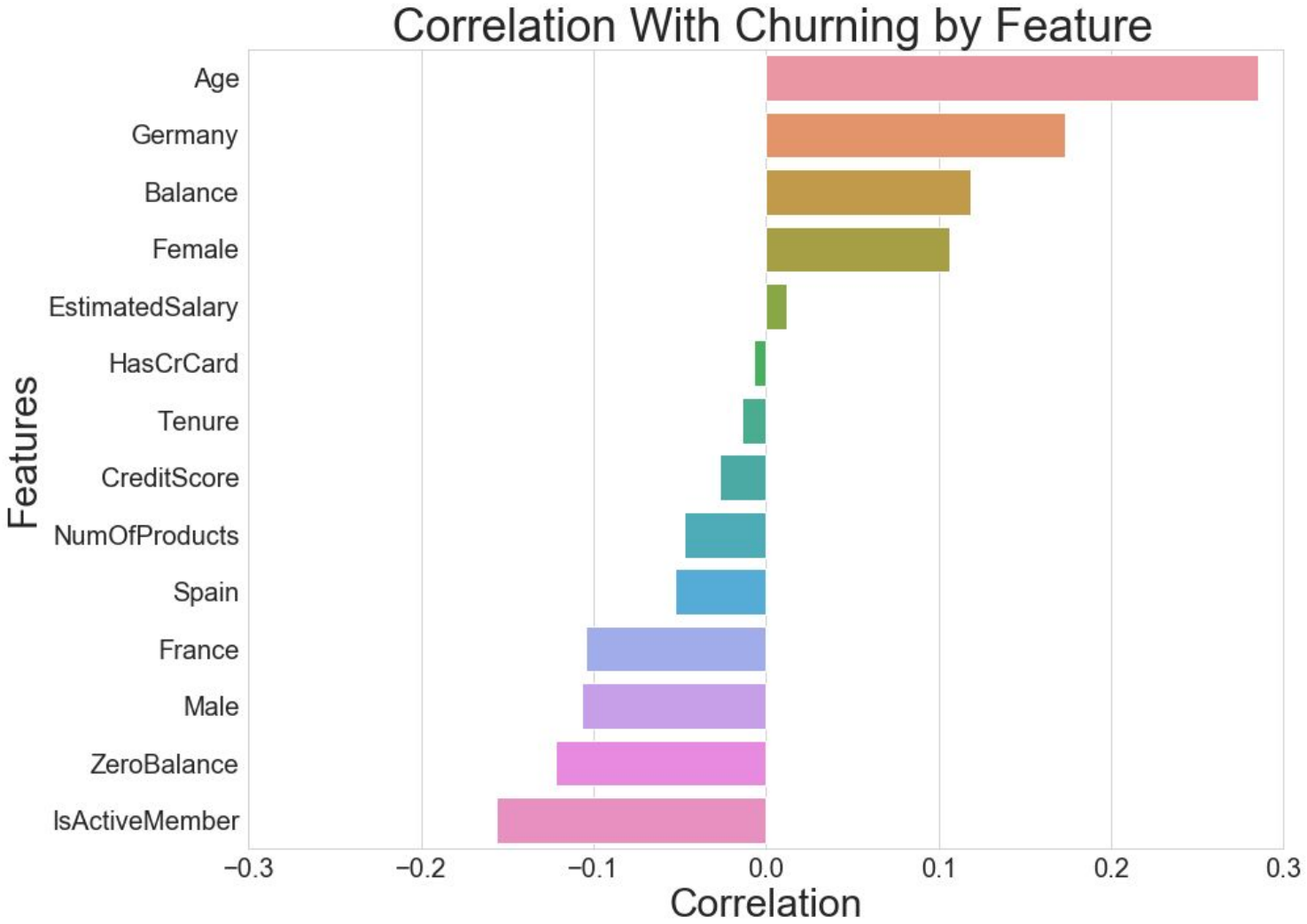


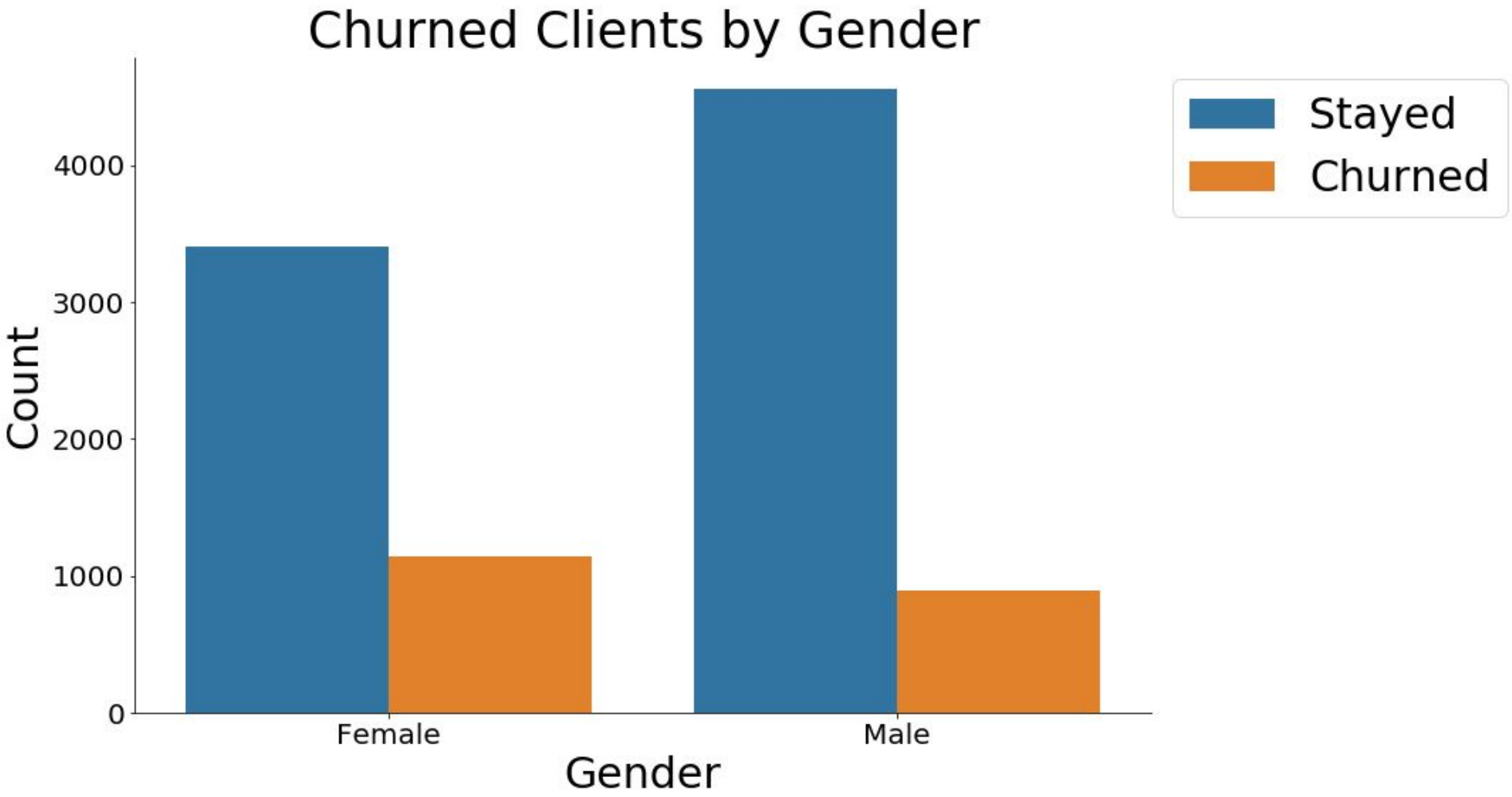
Win clients back ***before*** they leave

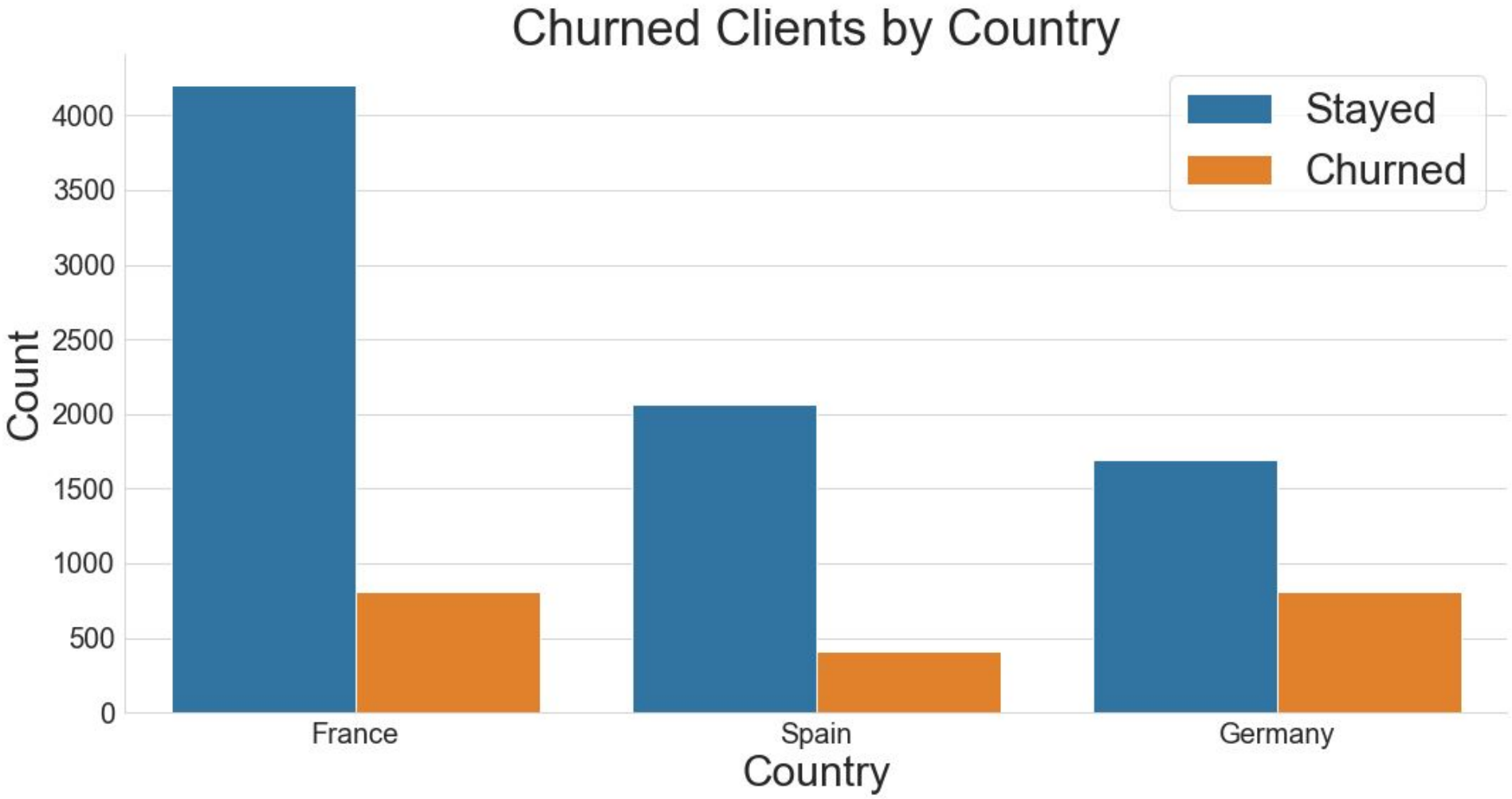
Increase average lifetime value per customer

Rule of Thumb:

Acquiring a new customer is ***5x more expensive*** than keeping one


Correlation With Churning by Feature
Age
Germany
Balance
Female
EstimatedSalary
HasCrCard
Tenure
CreditScore
NumOfProducts
Spain
France
Male
ZeroBalance
IsActiveMember
Features
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3
Correlation


Churned Clients by Gender
Stayed
Churned
Count
4000
3000
2000
1000
0
Female
Male
Gender


Churned Clients by Country
Stayed
Churned
Count
4000
3500
3000
2500
2000
1500
1000
500
0
France
Spain
Germany
Country

Predicting Churn

Fit a Random Forest Model to the given data

- Handles non-linear relationships well

Predicting Churn

Evaluation Metrics:

- Recall, Precision, and F-beta score

Recall vs. Precision

Recall:

- What percent of the churning clients did we identify?

Recall vs. Precision

Recall:

- What percent of the churning clients did we identify?

Precision:

- What percent of the clients we identified as churning were correct?

Recall vs. Precision

Recall:

- **What percent of the churning clients did we identify?**

Precision:

- What percent of the clients we identified as churning were correct?

Predicting Churn

Assumed recall was ***5x more valuable*** than precision based on ‘Rule of Thumb’

Predicting Churn

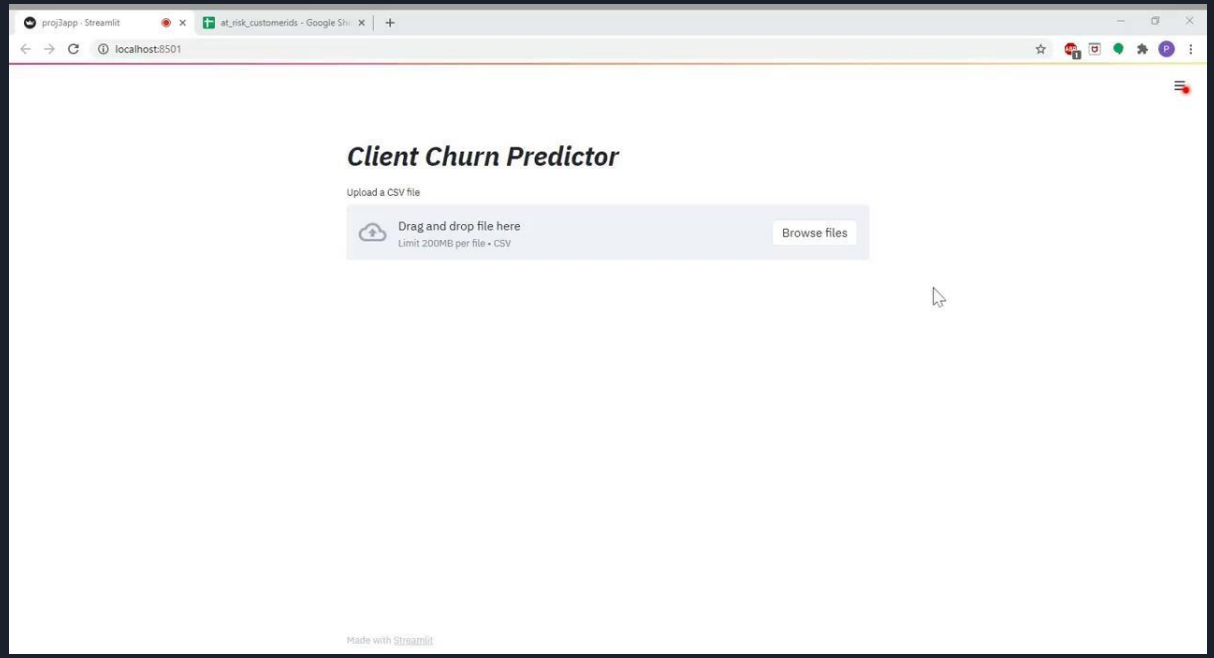
Assumed recall was ***5x more valuable*** than precision based on ‘Rule of Thumb’

Used F-beta score to optimize parameters of model

- *Beta* is chosen such that recall is *beta* times more important than precision

Results

App Demo


proj3app · Streamlit
at_risk_customerids - Google Sh
localhost:8501
Client Churn Predictor
Upload a CSV file
Drag and drop file here
Limit 200MB per file • CSV
Browse files
Made with Streamlit

Results

1250 clients out of 2000 in holdout set labeled as at-risk of churning (63%)

Results

1250 clients out of 2000 in holdout set labeled as at-risk of churning (63%)

Model Recall: **92%**

Model Precision: **30%**

Results

Compare to Random Selection of Clients:

Results

Compare to Random Selection of Clients:

Recall: 61% --- 31% lower than model

Precision: 18% --- 12% lower than model

Future Work

Improve model precision without lowering recall

Future Work

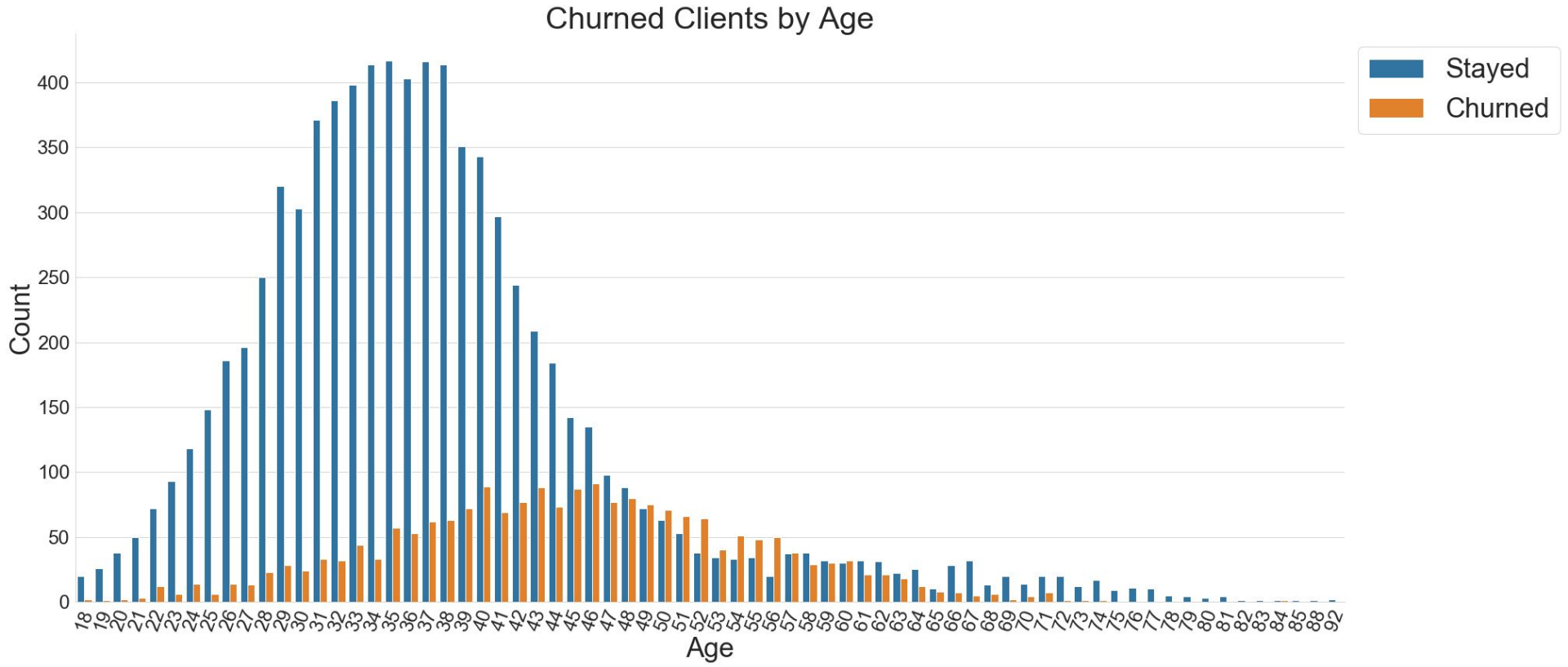
Improve model precision without lowering recall

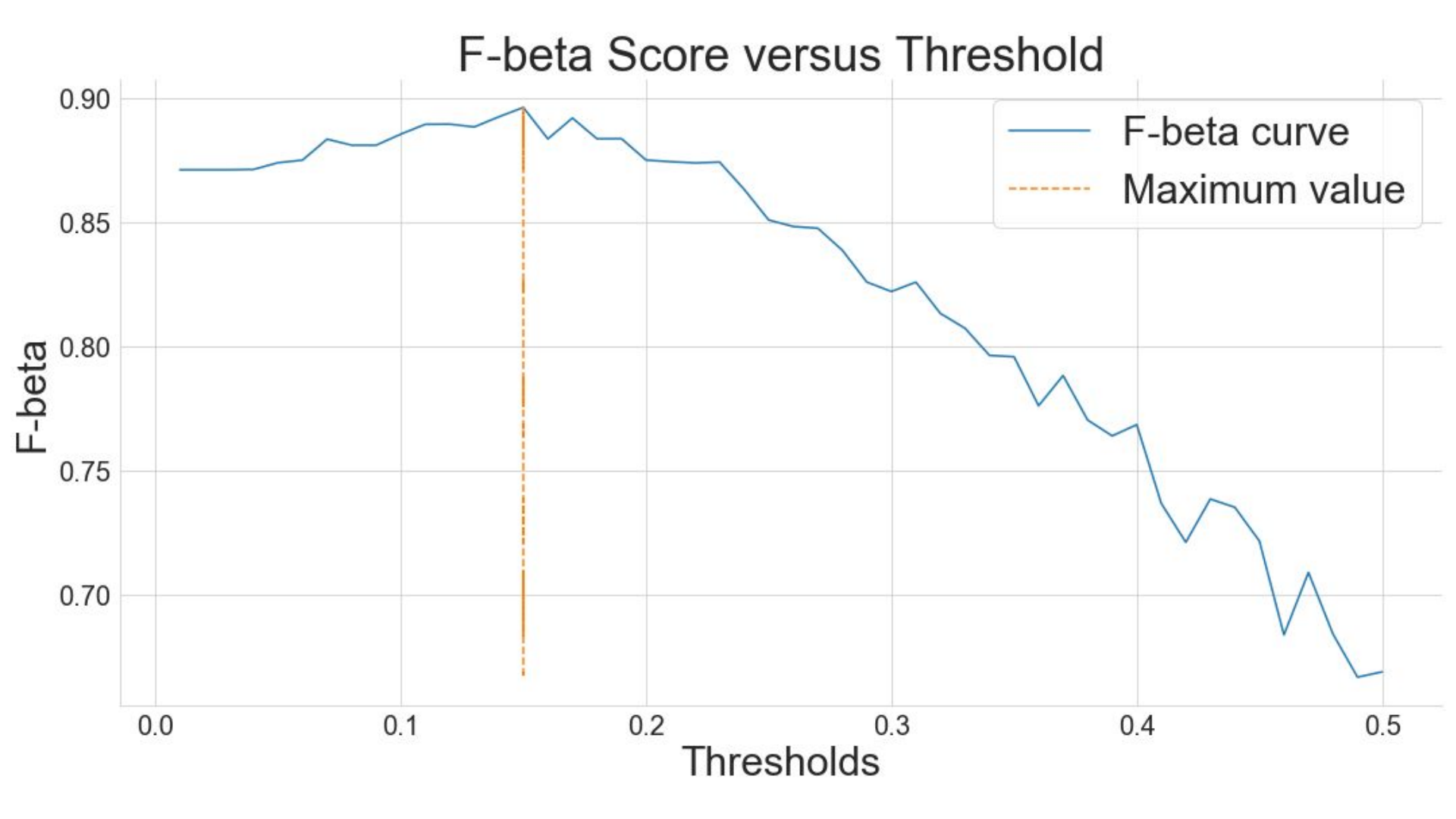
- Add features
- Neural Networks/Deep Learning

Thank you!

Any questions?

Appendix


Churned Clients by Age
Stayed
Churned
Count
0
50
100
150
200
250
300
350
400
Age
18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 88 92


F-beta Score versus Threshold
F-beta curve
Maximum value
F-beta
0.90
0.85
0.80
0.75
0.70
0.0
0.1
0.2
0.3
0.4
0.5
Thresholds

Results