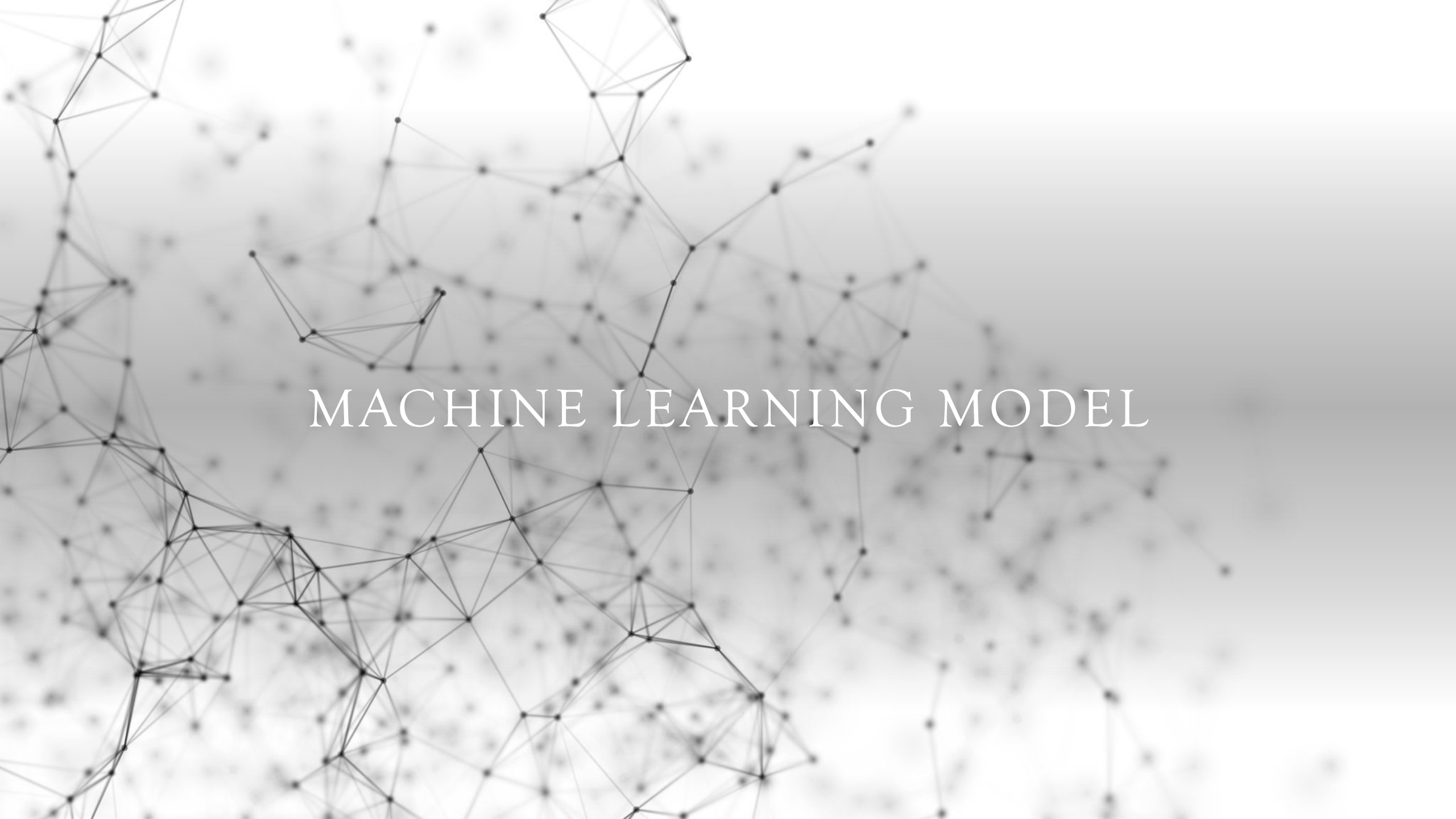

MACHINE LEARNING MODEL

10 000 INSTANCES
BANK CUSTOMERS
PREDICT CHURNS

GRADIENT BOOSTING

TEST RECALL : 84%

RANDOM FOREST

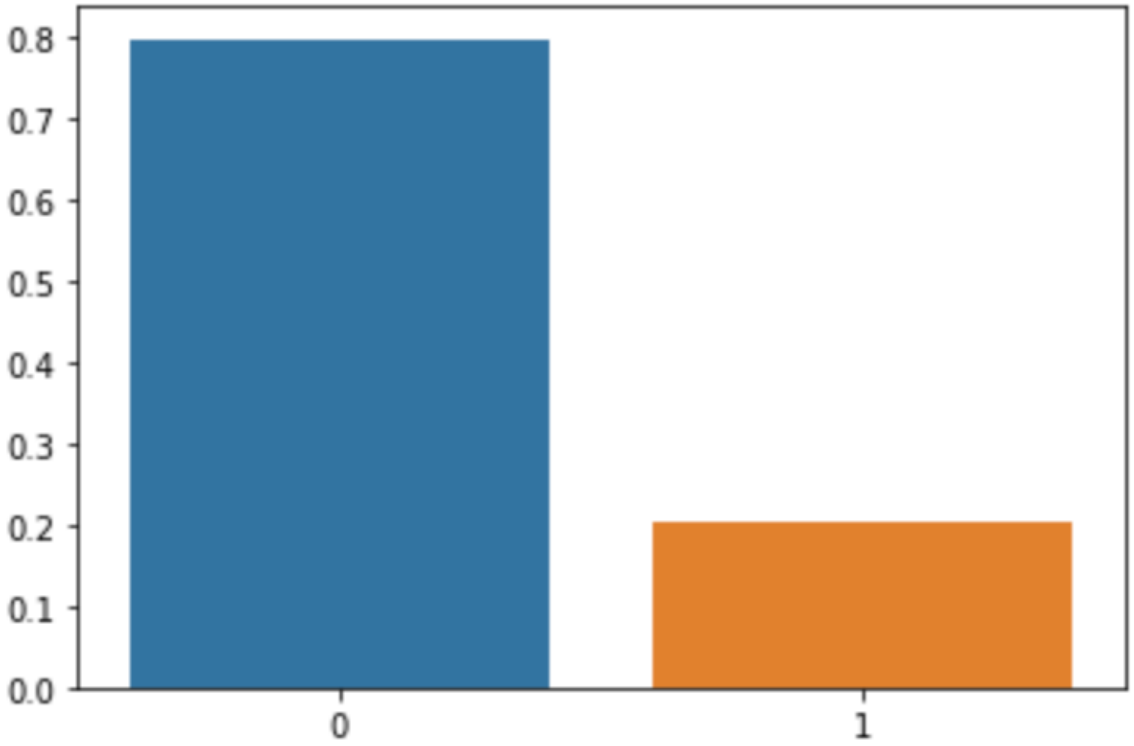
TEST RECALL: 74%

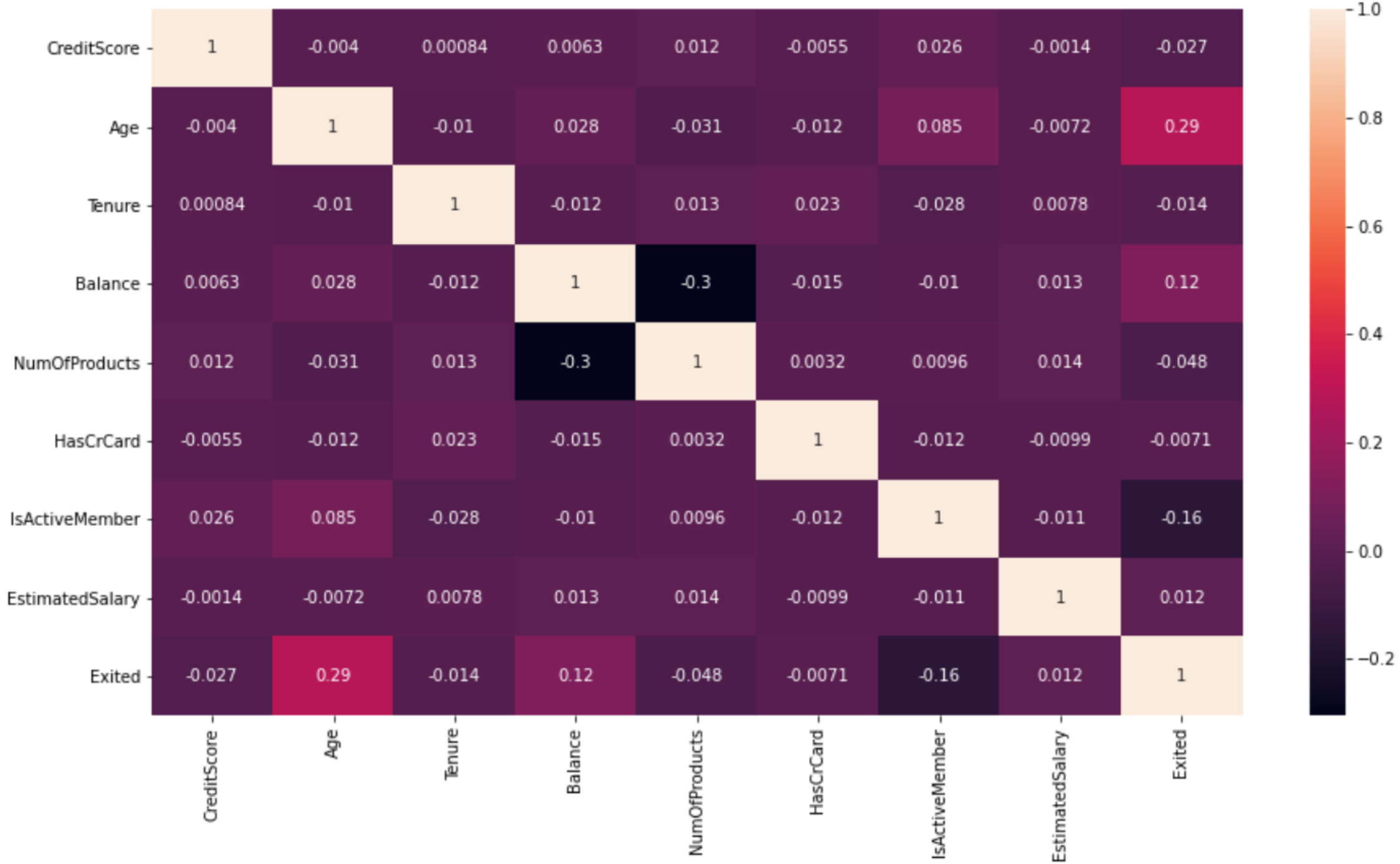
DATA

- Name, age, country, tenure, id
- Creditscore, balance, estimated salary
- Credit card, active member, Num of products

10 000 INSTANCES 20,37% CHURNS

```
0    0.7963
1    0.2037
Name: Exited, dtype: float64
```




CreditScore 1 -0.004 0.00084 0.0063 0.012 -0.0055 0.026 -0.0014 -0.027
Age -0.004 1 -0.01 0.028 -0.031 -0.012 0.085 -0.0072 0.29
Tenure 0.00084 -0.01 1 -0.012 0.013 0.023 -0.028 0.0078 -0.014
Balance 0.0063 0.028 -0.012 1 -0.3 -0.015 -0.01 0.013 0.12
NumOfProducts 0.012 -0.031 0.013 -0.3 1 0.0032 0.0096 0.014 -0.048
HasCrCard -0.0055 -0.012 0.023 -0.015 0.0032 1 -0.012 -0.0099 -0.0071
IsActiveMember 0.026 0.085 -0.028 -0.01 0.0096 -0.012 1 -0.011 -0.16
EstimatedSalary -0.0014 -0.0072 0.0078 0.013 0.014 -0.0099 -0.011 1 0.012
Exited -0.027 0.29 -0.014 0.12 -0.048 -0.0071 -0.16 0.012 1
CreditScore
Age
Tenure
Balance
NumOfProducts
HasCrCard
IsActiveMember
EstimatedSalary
Exited
1.0
0.8
0.6
0.4
0.2
0.0
−0.2

TASK IS TO FIND THE ONES PRONE TO CHURN(EXIT). WE WILL BE PARTICULARLY INTERESTED IN FALSE NEGATIVES (AND OF COURSE TRUE NEGATIVES). THAT IS; THOSE PREDICTED TO STAY BUT THEY´RE ACTUALLY LEAVING. A POSITIVE IN THIS CASE (1) IS EQUAL TO A CHURN. THUS, RECALL IS OF BIGGEST INTEREST.

80/20-SPLIT TWICE

Dataset → 80% development, 20 % test
Development → 80% train, 20% validation

DECISIONTREECLASSIFIER
RANDOMFORESTCLASSIFIER
SVC
KNEIGHBORSCLASSIFIER
GAUSSIANNB
LOGISTICREGRESSION
GRADIENTBOOSTINGCLASSIFIER

RANDOMFOREST & GRADIENTBOOSTING

GRIDSEARCH FOR THESE TWO

GRADIENT BOOSTING

DEV RECALL : 81%
VAL RECALL: 84%

RANDOM FOREST

DEV RECALL : 100%
VAL RECALL: 74%

GRADIENT BOOSTING	RANDOM FOREST
TEST RECALL : 84%	TEST RECALL: 74%

THANK YOU!