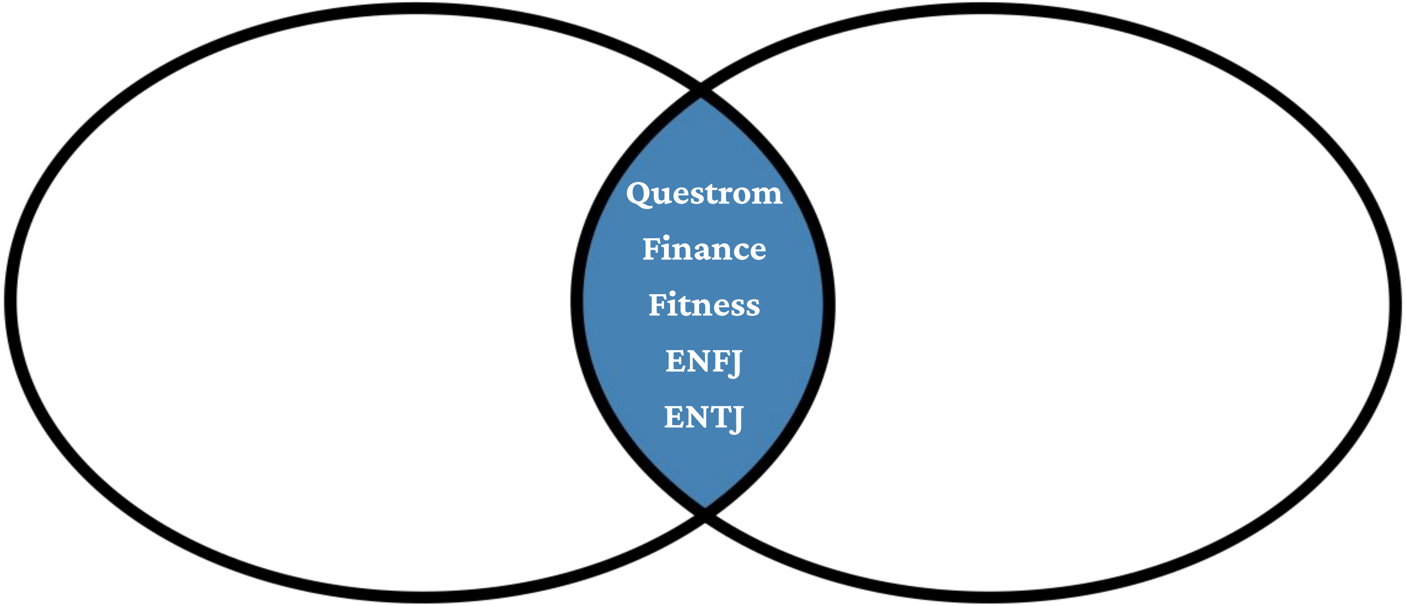
Bank Customer Churn Prediction with ANN

Using Artificial Neural Networks to predict bank customer churn

Sungwoo Noh & Jaejoong Kim

1. Our Team

Sungwoo Noh ∩ Jaejoong Kim



2. Problem Statement

1. Customer Churn is Costly
2. Acquisition < Retention
3. Actionable Insight

BAIN & COMPANY   Industries ▾   Consulting Services ▾   Digital   Insights ▾   About ▾   Careers

Article

Retaining customers is the real challenge

By increasing retention by as little as 5%, profits can increase by 25% to 95%

January 20, 2006 • 2 min read

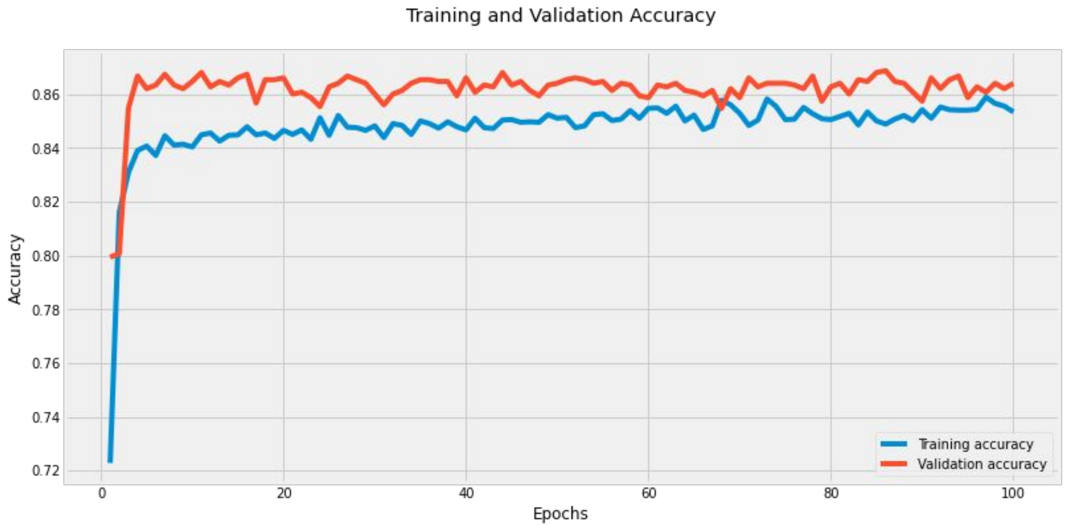
3. Method: ANN Model Architecture and Training

Base Neural Network Model

- ➔ Input layer with 10 features
- ➔ 2 dense layers & 2 dropout layers
- ➔ 2 batch normalization layers, and output layer

Training Parameters

- ➔ Loss function: Binary Cross-Entropy
- ➔ Optimizer: Adam
- ➔ Epochs: 10



- ➔ Peak Accuracy on Training Dataset:
  - ◆ **0.8644** (Epoch 69)
- ➔ Peak Accuracy on Validation Set:
  - ◆ **0.8680** (Epoch 85)

4. Experiments

Feature Selection

- ➔ Lasso Regularization
  - ◆ Credit Score, Geography, Gender, Age, Balance, IsActiveMember
  - ◆ Validation Accuracy : **0.848**
- ➔ Forward Selection
  - ◆ Credit Score, Gender, Age, Balance, IsActiveMember
  - ◆ Validation Accuracy: **0.846**

Data Balancing

- ➔ Resolve Data Imbalance Using SMOTE
  - ◆ Validation Accuracy: **0.8008**

Hyperparameter Tuning

- ➔ Using GridSearchCV to tune HyperParameters
- ➔ Validation Accuracy: **0.8756**

Best Parameters:

```
'batch_size': 32
'model__activation': 'relu'
'model__dropout_rate': 0.1
'model__neurons': 32
'model__optimizer': 'rmsprop'
```

Model: "sequential_468"

Layer (type)	Output Shape	Param #
dense_1400 (Dense)	(None, 32)	352
dropout_932 (Dropout)	(None, 32)	0
batch_normalization_932 (BatchNormalization)	(None, 32)	128
dense_1401 (Dense)	(None, 32)	1,056
dropout_933 (Dropout)	(None, 32)	0
batch_normalization_933 (BatchNormalization)	(None, 32)	128
dense_1402 (Dense)	(None, 2)	66

Total params: 3,334 (13.03 KB)
Trainable params: 1,602 (6.26 KB)
Non-trainable params: 128 (512.00 B)
Optimizer params: 1,604 (6.27 KB)

5. Results and Evaluation

87.56%

Accuracy

A high accuracy score reflects the model's ability to correctly classify customers.

76.35%

Precision

Represents the proportion of correctly predicted churned customers among all predicted churned customers.

6. Challenges

1. Hypothesis Testing

- ➔ One of our Experiments
- ➔ Train two versions of the model where one model includes has_credit feature and the other model does not include the feature.
- ➔ We've been having difficulty obtaining the p-value from the Neural Network model.
- ➔ Implement regression model to obtain p-value along with SHAP(SHapley Additive exPlanations) values to see if having a credit card is an important factor in predicting churn.

2. Worse Performance with Experiments

- ➔ Expected models with experiments to outperform baseline model
- ➔ The models have actually shown worse performance compared to the baseline.

"Churn prediction is not just about identifying who will leave, but also why and when.

7. Conclusion and Fun Facts 💡

Age

Older customers were more likely to churn.

Balance

Customers with higher balances were more prone to churn.

Geography

German customers exhibited higher churn rates.

Thank you !

Sungwoo Noh & Jaejoong Kim