Customer Churn Prediction

Bank Churn Prediction:

Introduction to Neural Networks

Contents

Executive Summary

- The third artificial neural network model built is the most-optimized model for this problem statement and gives an overall recall rate of 75% and accuracy of 80%.

- Based on model performance and analysis:
  - Banks can try to take extra care of female customers by providing them with extra benefits and offers as the churn rate of female customers seems to be higher than that of male customers.
  - Banks can also make some promotional offers to keep the customers active as inactive customers seem to churn more than active customers do.
  - Banks should also look into their older customers on high priority, as it seems that older customers do not show high customer satisfaction and seem to churn at a higher rate.

- Banks should also try to keep good relations with their customers and connect with them from time to time to inform them about their usage and the advantage of credit cards, as customers with credit cards seem to churn more. This may happen due to the extra put on amount or service fees of credit cards which customers are not aware of.

Business Problem Overview and Solution Approach

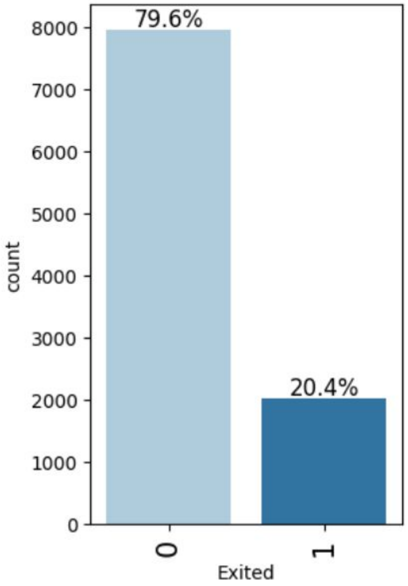
- Terra is a Bank which provides banking services like loans, debit and credit services with high transactions to customers in France and its neighbourhoods. From the past 1 year, Terra has been observing that customers are exiting the bank services or moving to other service providers.

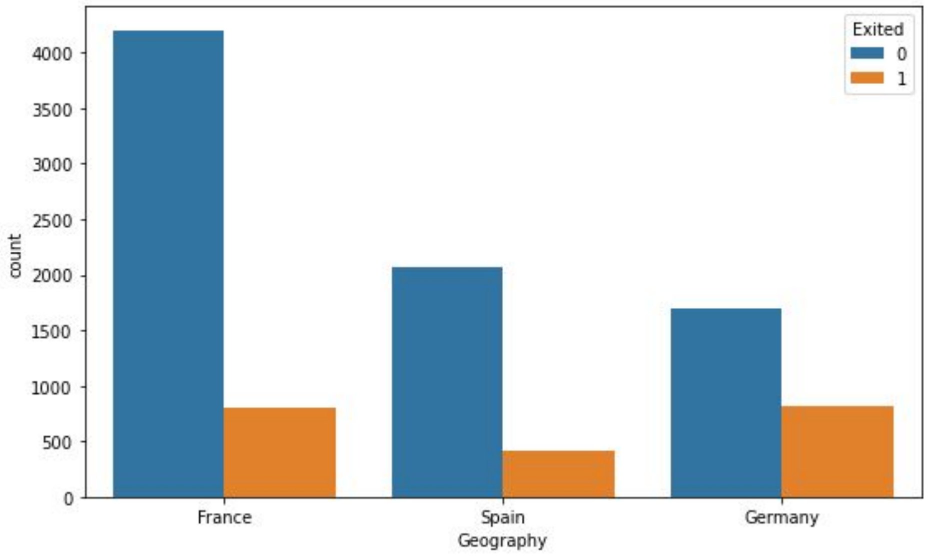
- Businesses like banks which provide services have to worry about the problem of ‘customer churn’ i.e. customers leaving and joining another service provider. So Terra bank wants to understand the factors or aspects of the service that cause a customer to exit from the bank.

- The task at hand is to analyze the data and provide reasonable insights to the bank so that the management can improve their policies or make promotional offers to increase their customer satisfaction rate and decrease the customer churn rate.

EDA Results

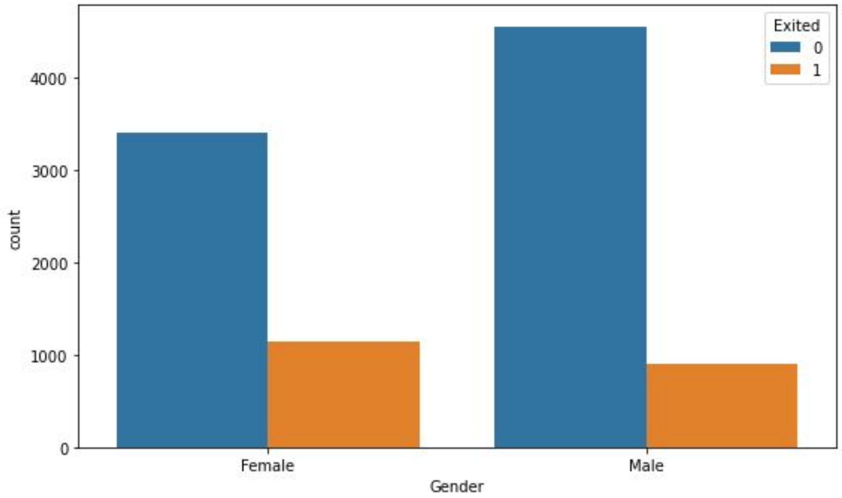
- The dataset is imbalanced and has 80% of its customers retained, whereas 20% of the customers have exited from the bank.
- We also observe that the majority of the customers are from France.

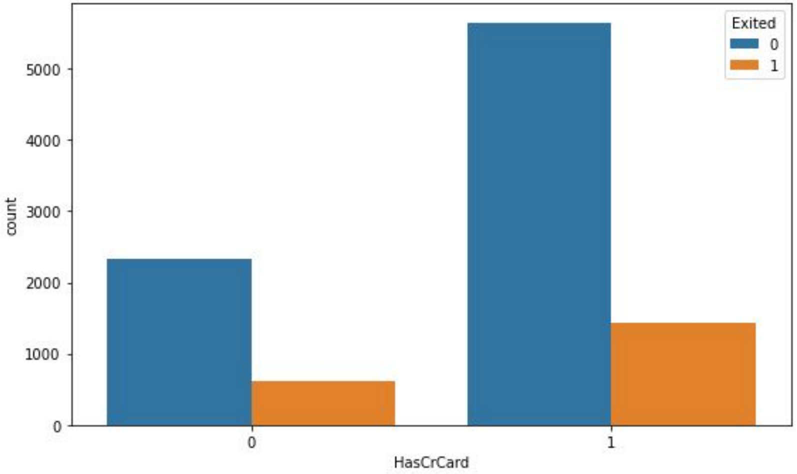




EDA Results

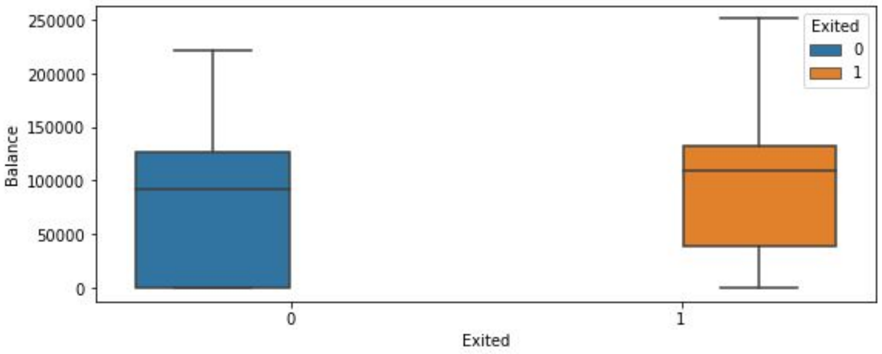
- We also observe that the number of female customers who churn are more than the male customers who churn.
- Also, the majority of the customers who churn use a credit card.

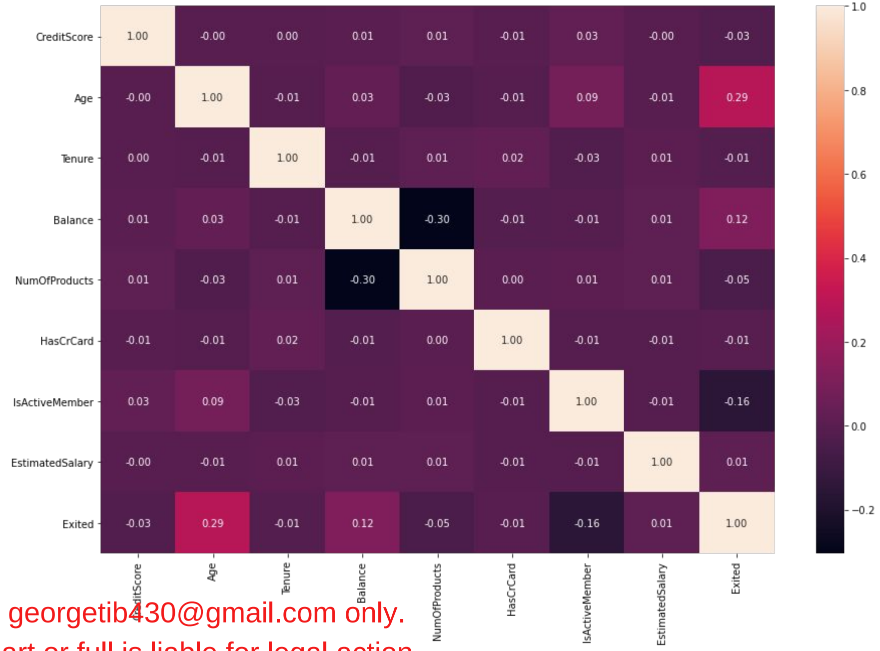




EDA Results

- We observe and it is worrying that most of the customers who churn also have a higher bank balance, so the bank losing these customers is likely to hit their available capital for lending.
- We also observe that there is no multicollinearity in the data.





Data Preprocessing

- There are a total of 14 columns in this dataset, and this data has no null-values.
- There are 3 columns with unique values in this dataset, namely Row number, CustomerId and Surname, so these columns are dropped from the dataset.
- There are 2 categorical columns which are Geography and Gender, and these columns are encoded using the get_dummies() function.
- All numerical columns, namely credit score, age, tenure, balance and estimated salary are scaled using the Standard Scaler.

Model Performance Summary

- We have built 6 models with different combination of optimizers and regularization techniques and below is the performance for each of the model.

Training Performance

Model	Recall
NN with SGD	0.219
NN with Adam	0.672
NN with Adam and Dropout	0.491
NN with SMOTE and SGD	0.742
NN with SMOTE and Adam	0.860
NN with SMOTE, Adam and Dropout	0.825

The recall scores in the tables are for the Target variable 1 (Customer who churned the bank)

Validation Performance

Model	Recall
NN with SGD	0.194
NN with Adam	0.493
NN with Adam and Dropout	0.395
NN with SMOTE and SGD	0.660
NN with SMOTE and Adam	0.673
NN with SMOTE, Adam and Dropout	0.697

Model Performance Summary

- From the above tables, we can observe that The recall score difference between the train and validation sets is notably lower for the NN with SMOTE,Adam and Dropout and also has a higher train and test score.

- Hence, we select this model as the best option since a lower difference suggests improved consistency and generalization across both training and validation datasets.

APPENDIX

Data Background and Contents

- There are a total of 14 columns in this dataset, and this dataset contains 10,000 sample points.

- The columns are namely customerid, surname, credit score, geography, gender, age, tenure, number of products, balance, HasCrCard, estimated salary, isactive, and exited. Exited is the target feature which tells us if the customer has left the bank within the last 6 months.

- A few notable points from the statistical summary and EDA:
    - There is a strong positive correlation between age and target feature exited.
    - The target feature has only 2 values: 0 (customer did not leave the bank) and 1 (customer left the bank).
    - There are no null/missing values in the dataset.
    - The dataset is imbalanced as 20% of the customers have exited the bank and 80% of the customers are retained.

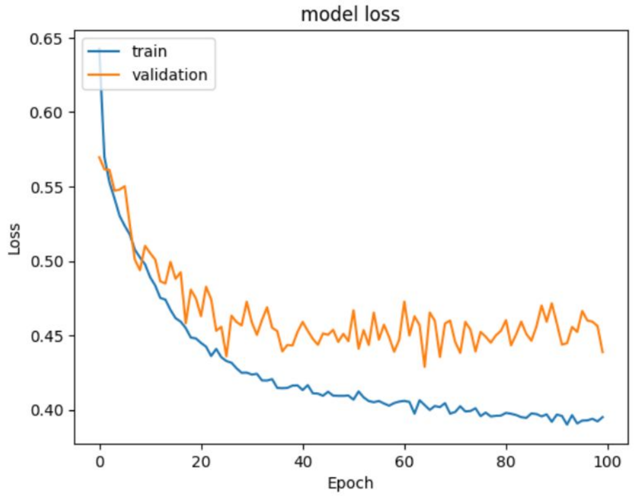
Final Model Summary

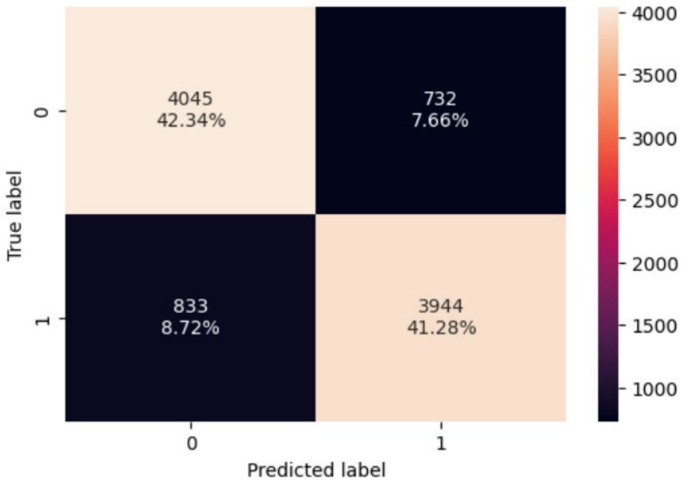
- Several models are trained both on the imbalanced data as well as the data balanced using SMOTE, but the most optimized model is the model is the model with combination of SMOTE, Adam and Dropout. In this model, we have used binary_crossentropy as the loss function, Adam as the optimizer and Dropout as a regularization technique. The structure of the model and the parameters can be observed below:

```
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense (Dense)               (None, 32)                384

 dropout (Dropout)           (None, 32)                0

 dense_1 (Dense)             (None, 16)                528

 dropout_1 (Dropout)         (None, 16)                0

 dense_2 (Dense)             (None, 8)                 136

 dense_3 (Dense)             (None, 1)                 9

=================================================================
Total params: 1057 (4.13 KB)
Trainable params: 1057 (4.13 KB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

Final Model Summary

- We also observe the model loss and confusion matrix of the best model here.
- The Recall value of this model outperformed all other models, and the train and validation loss of the model decreases continuously.





Happy Learning !