Background

What is Customer Churn?

Customer churn refers to the natural business cycle of losing and acquiring customers. Every company no matter the quality of its products or customer service experiences churn. Generally speaking, the less churn we have, the more customers we keep.

Why is customer churn important?

Understanding customer churn is essential to evaluating the effectiveness of our marketing efforts and the overall satisfaction of the customers. It's also easier and cheaper to keep customers we already have versus acquiring new ones. Due to our subscription business models, it's critical for us to understand where, how, and why our customers may be churning.

Objective

The primary objective of the customer churn predictive model is to retain customers at the highest risk of churn by proactively engaging with them. Knowing in advance which customers may churn soon, especially in the case of high revenue or long-time customers, can help the company to focus exactly on them and develop an efficient strategy to try to convince them to stay. The approach can include a call to such clients with a special offer of a gift, discount, subscription upgrading for the same price, or any other customized experience.
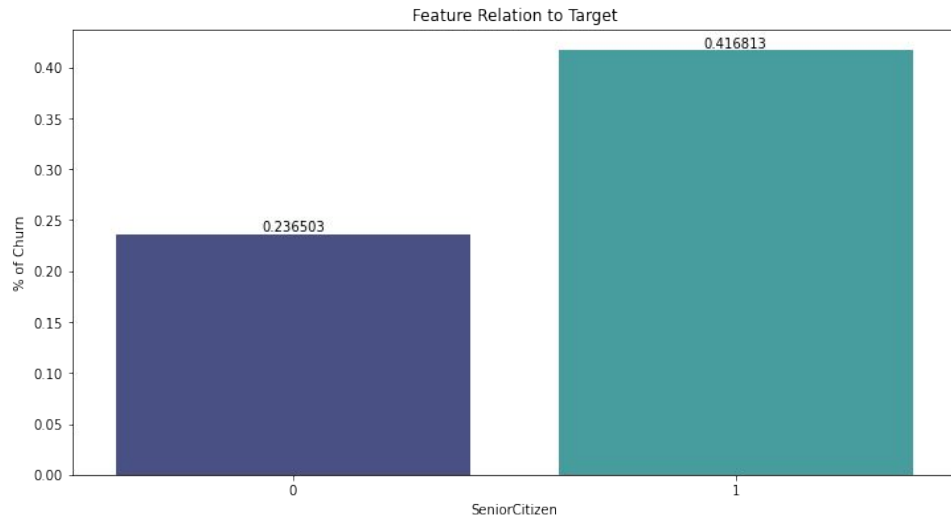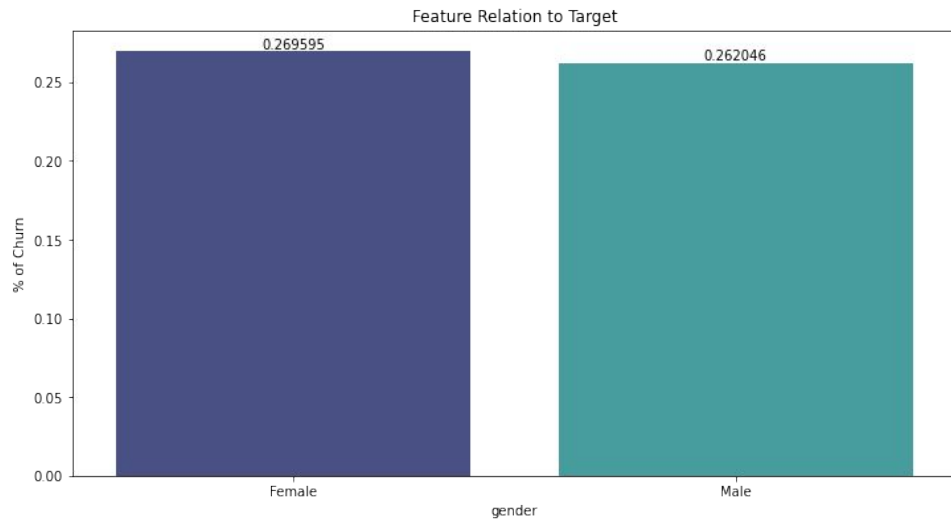
Outline

| Exploratory Data Analysis | Data Preprocessing | Model Result & Improvement | Conclusions & Suggestion |

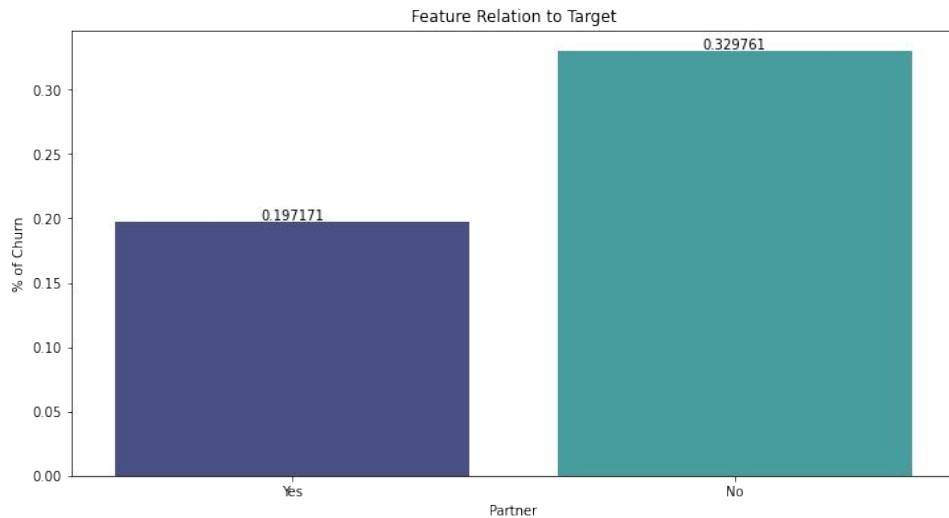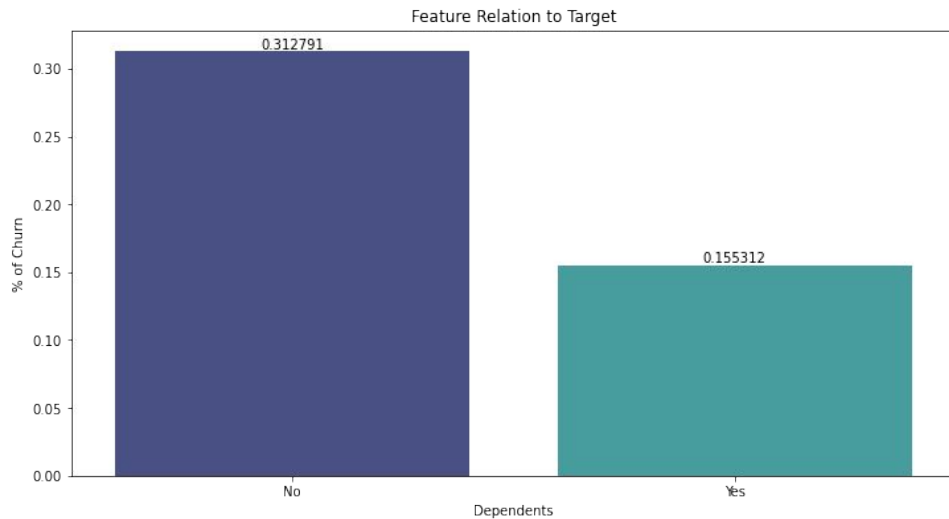# Exploratory Data Analysis
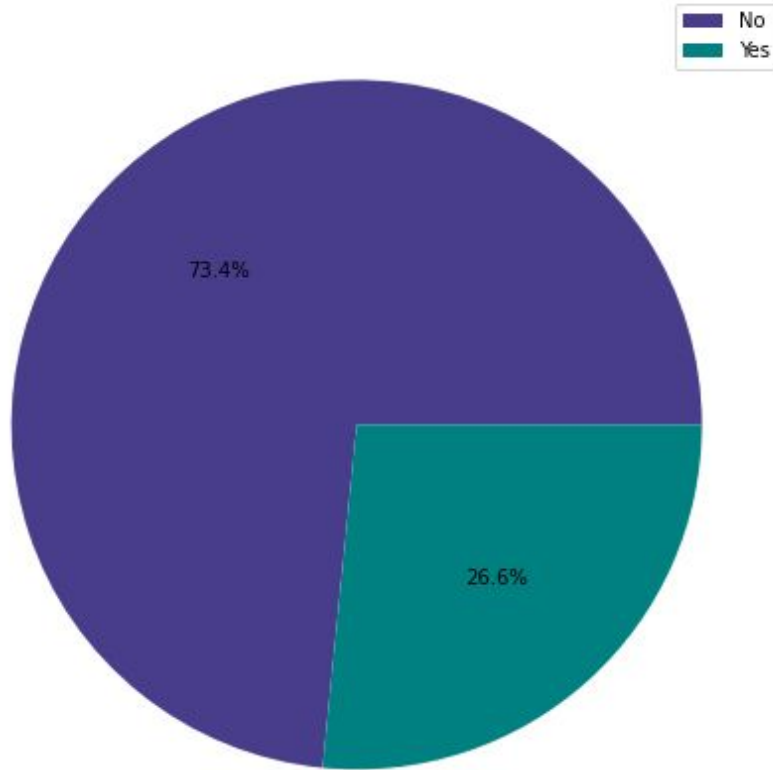
Customer Churn

27% of our customers leave our products/service. It's an alarming number for our business.

Customer Demographic

There is no relationship found between gender and churning. Both men and women have the same tendency to churn.

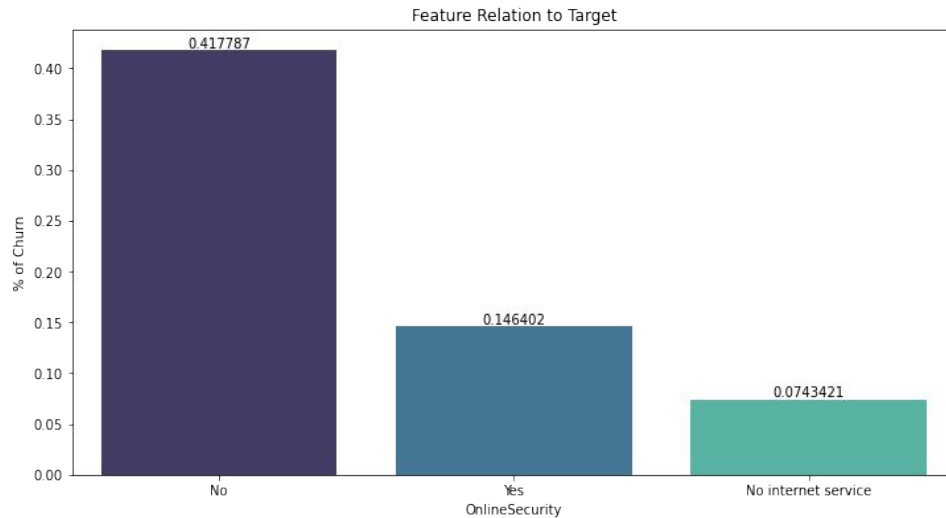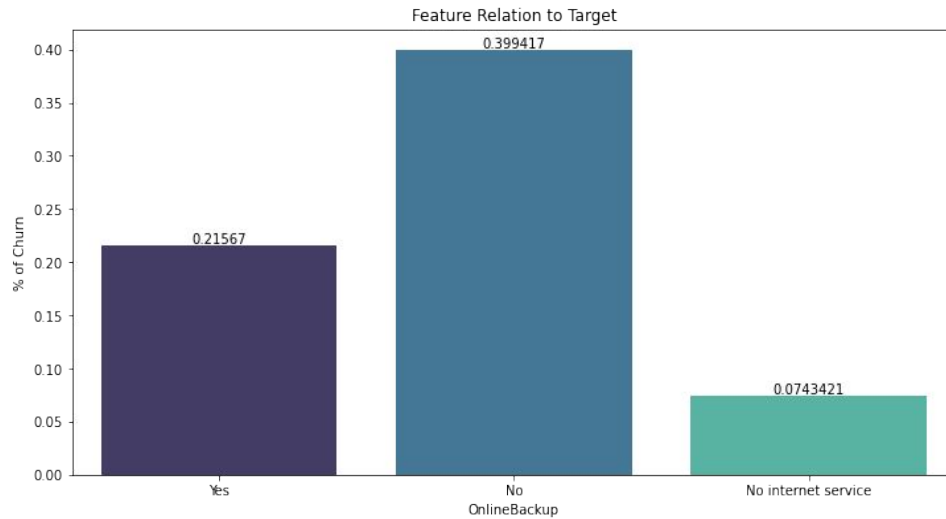Customers with age more than 65 are more likely to churn.

Customer Demographic

Customers who already have partners and dependents are less likely to churn compared to a customer whose single and have no dependents.

Feature Relation to Target
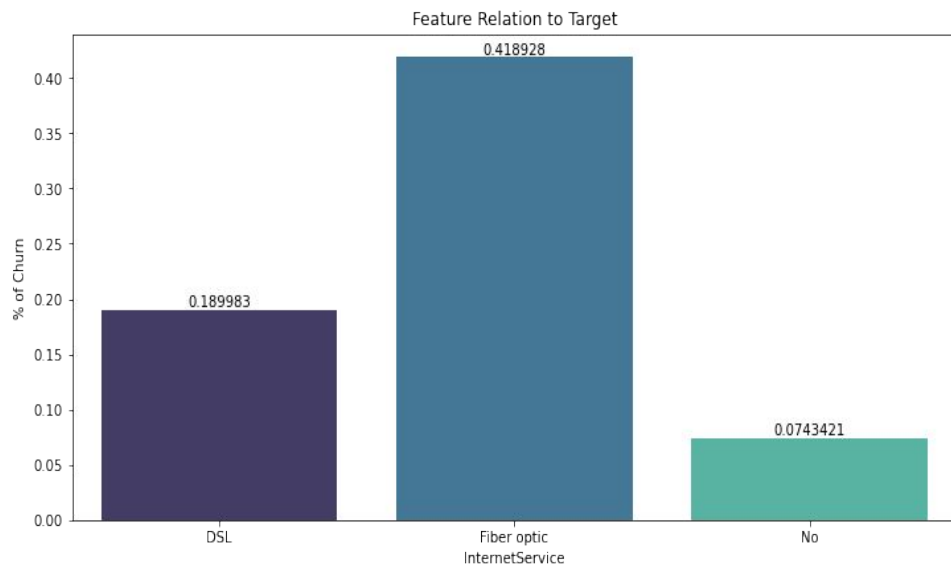
Internet Service

Customers that use fiber optic for internet service are more likely to churn than others.

Feature Relation to Target (OnlineBackup)

- Yes: 0.21567
- No: 0.399417
- No internet service: 0.0743421

Feature Relation to Target (OnlineSecurity)

- No: 0.417787
- Yes: 0.146402
- No internet service: 0.0743421

Additional Services

Besides the main product, our company has additional services for the customer who wants them. Namely Online Backup and Online Security.

Customers who haven't used our additional services are more likely to churn than others.
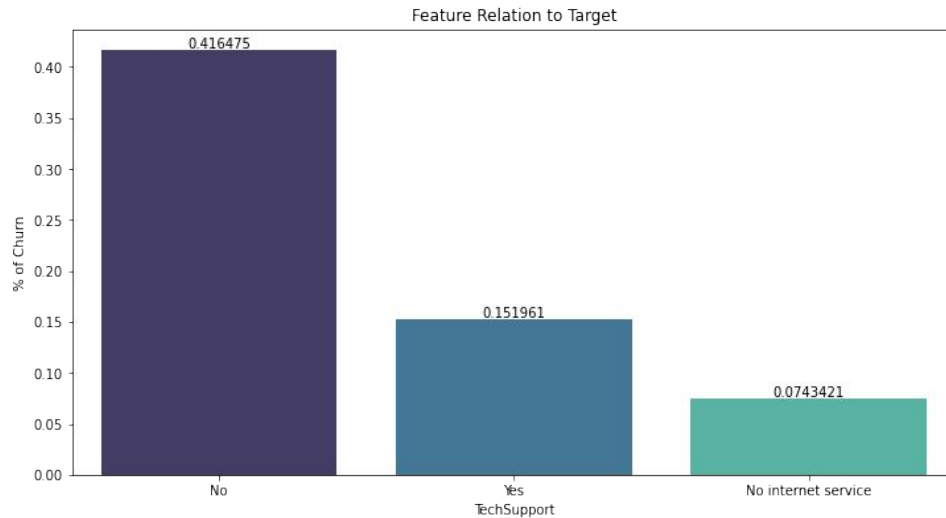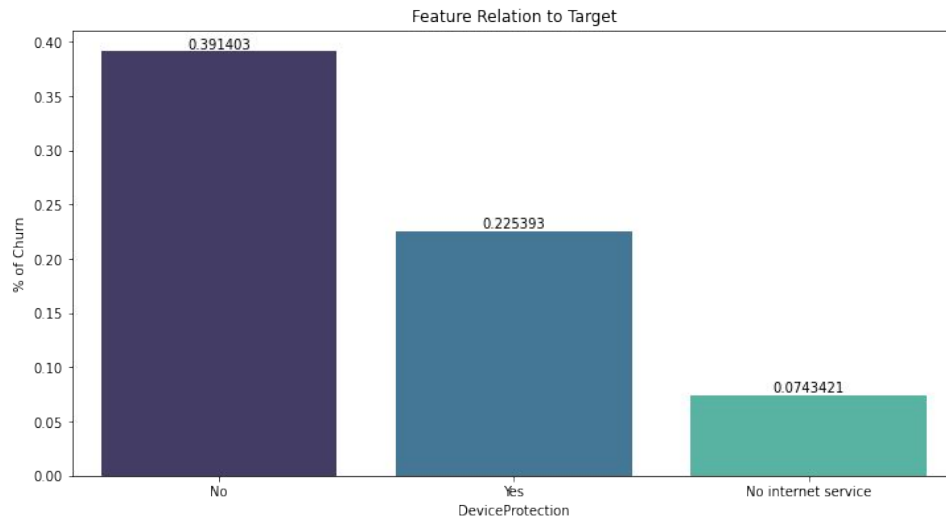
Additional Services

Besides the main product, our company has additional services for the customer who wants them. Namely Device Protection and Tech Support.

Customers who haven't used our additional services are more likely to churn than others.

Feature Relation to Target (Contract)

Feature Relation to Target (tenure_range)
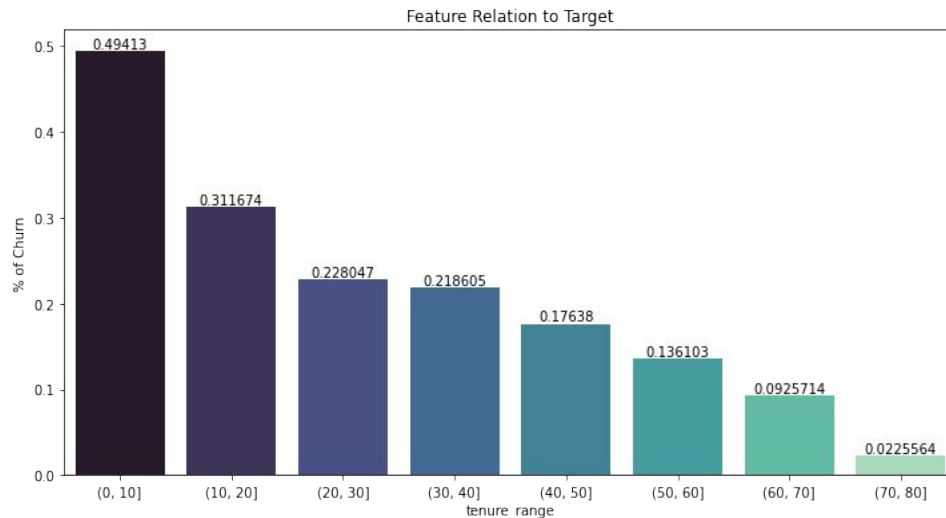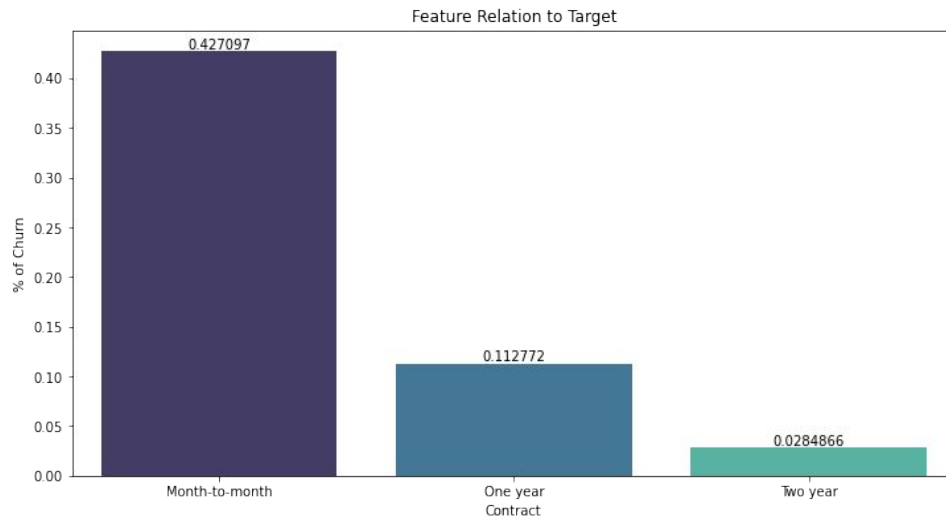
Subscription Type and Payment

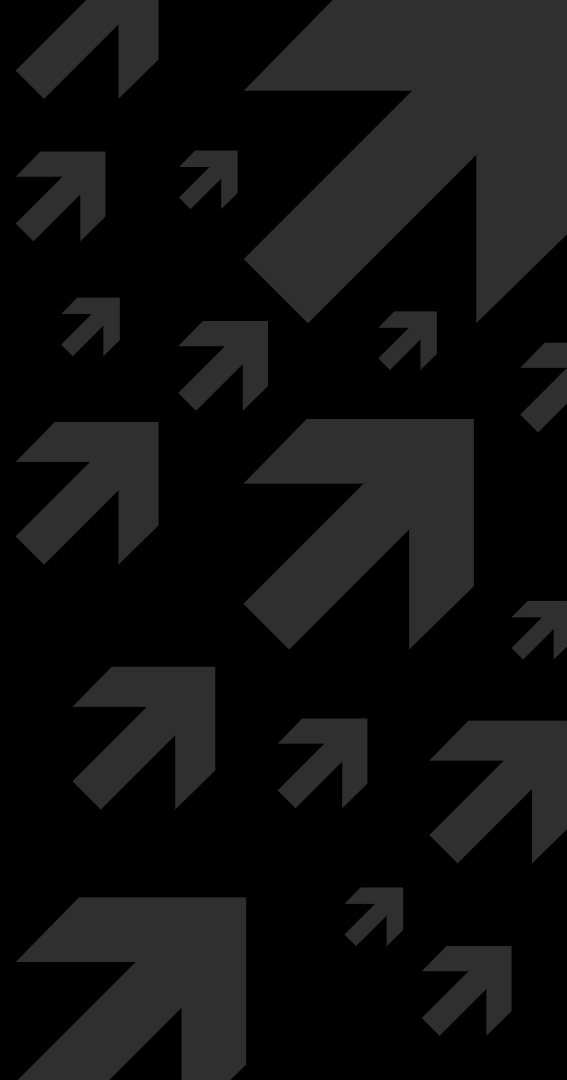Customers whose on a monthly contract are more likely to churn than others.

Customers that used our product for less than a year are still prone to leaving our product.
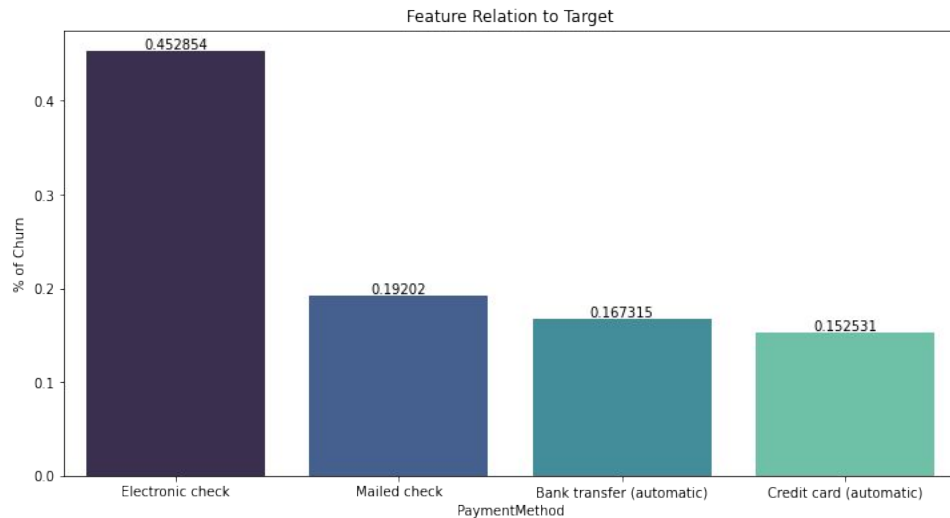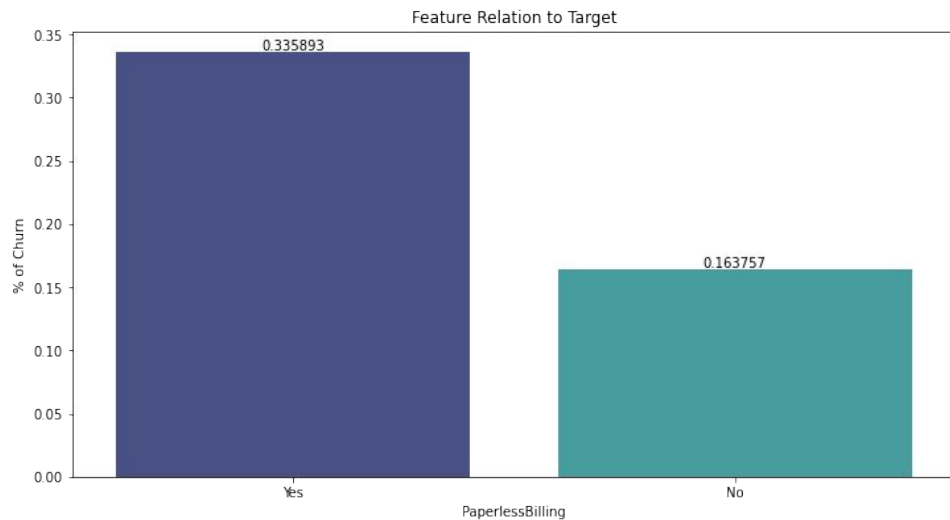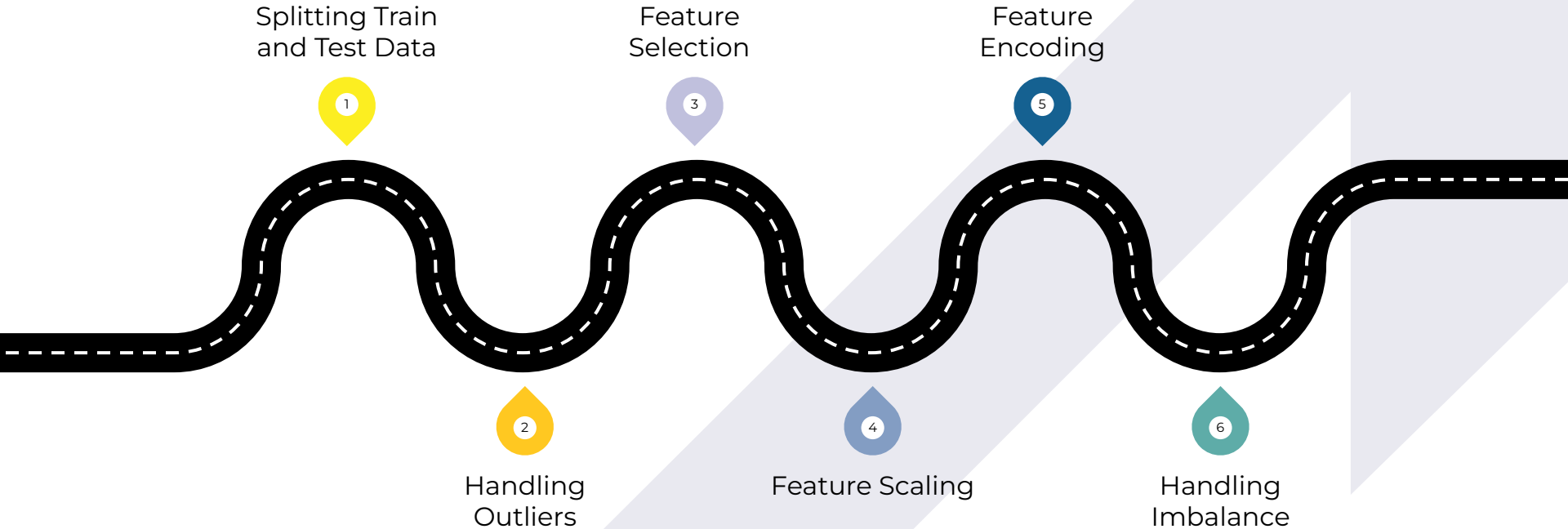
Subscription Type and Payment

Customers who use paperless billing and electronic check for payment are more likely to churn than others.
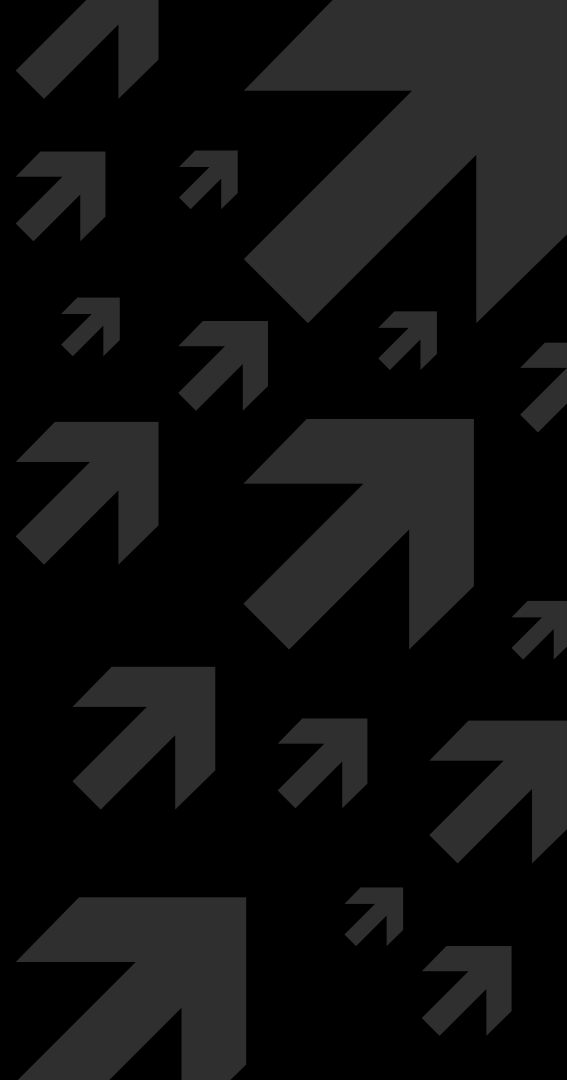
Data Preprocessing

Preprocessing Workflow

Splitting Train and Test Data

1

Handling Outliers

2

Feature Selection

3

Feature Scaling

4

Feature Encoding

5

Handling Imbalance

6

Model Result & Improvement

Sequential Base Model Evaluation

Training and Validation Accuracy

Training and Validation Losses

Functional Base Model Evaluation

Training and Validation Accuracy
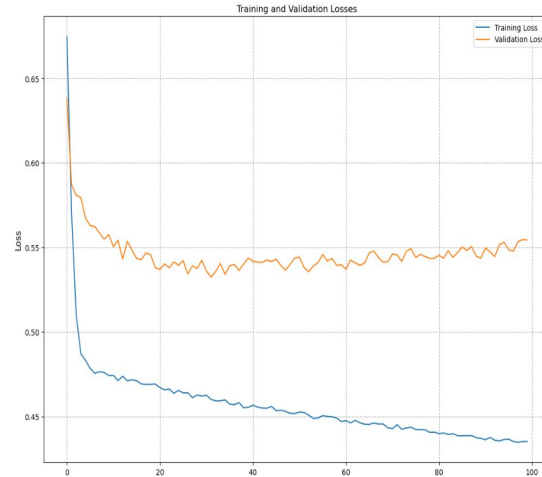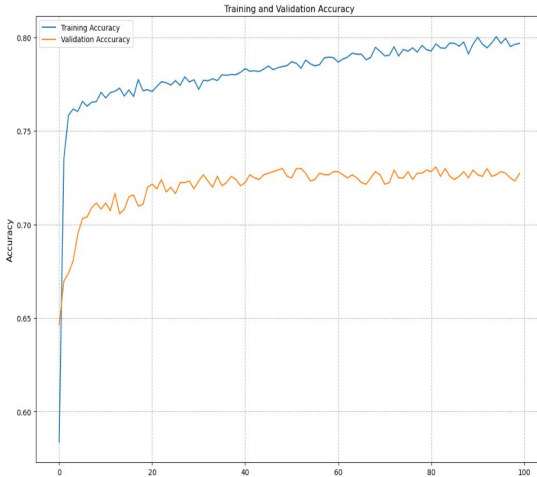

Training and Validation Losses

Base Model

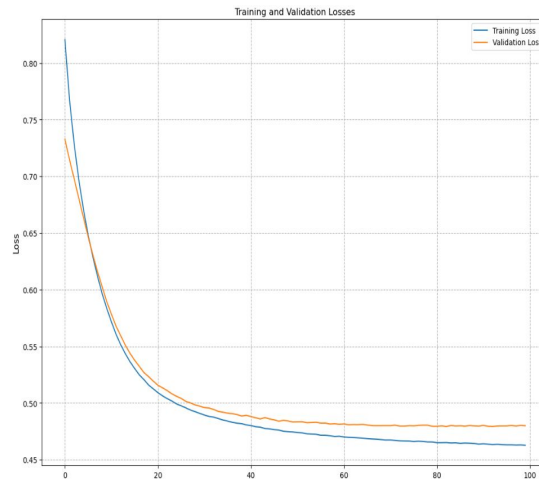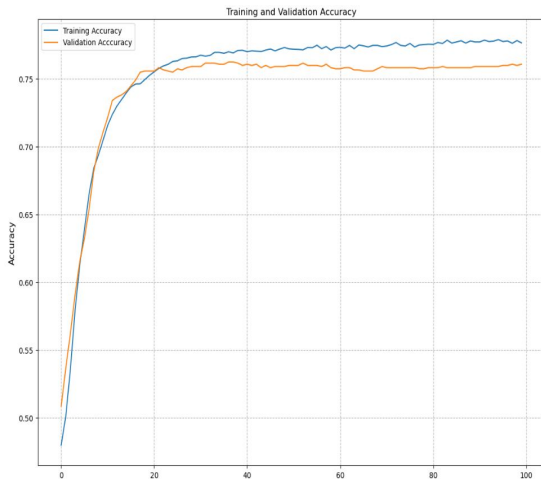Hyperparameters for both base models are:

- 1 hidden layer with 16 neuron
- Activation function : Relu
- Optimizer : Adam

As we can see on the graphic, the gap between data train and validation is still distant. So we could say that the models obtained are still overfitting and need to be improved.
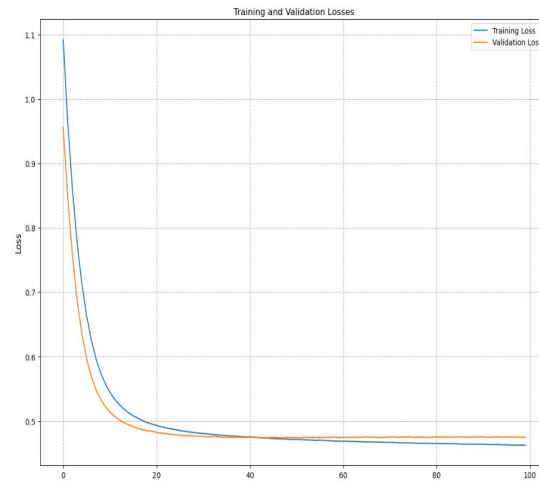
17

Sequential Method Evaluation

Functional Method Evaluation
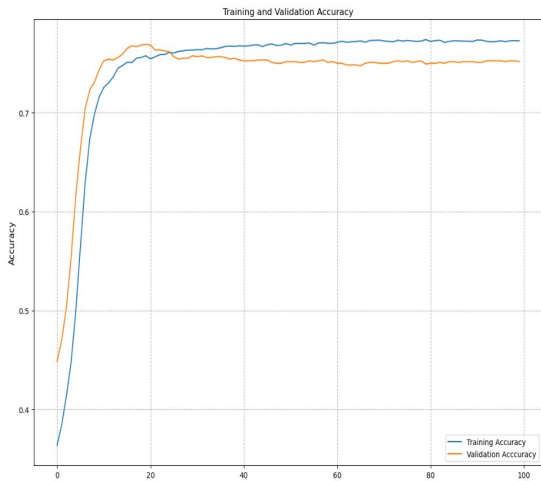
Improvement

We improve the model basically by manually training it with multiple hyperparameter combinations. These are the hyperparameters that we got after a few tries:

- 2 hidden layer with 16 and 8 neurons.
- Activation function for both hidden layer are Elu.
- Optimizer Adam with learning rate 0.0001.

As we can see on the graphic, the gap between data train and validation is getting closer. Or we could say that the models obtained are fit now.

18

Model Evaluation Metric

Model	Accuracy	Recall 1	AUC-Score
ANN Sequential	0.73	0.86	0.77
ANN Functional	0.71	0.86	0.76

Based on the model evaluation metric, we choose **ANN Sequential** as the best model for predicting customer churn.

Conclusion & Suggestion
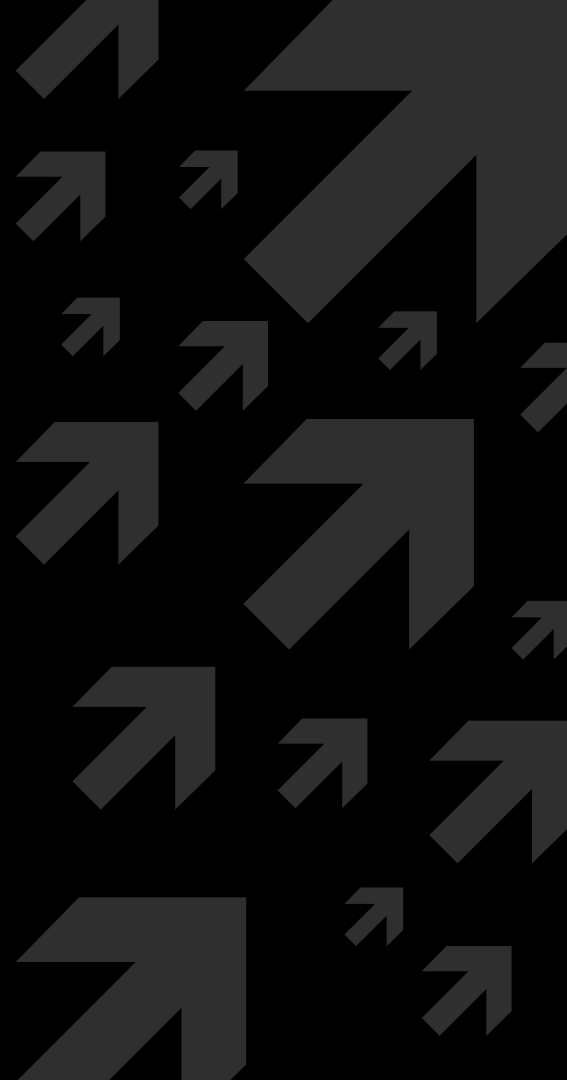
Conclusion

- The performance of the ANN Sequential model is good enough based on the accuracy score. Especially in accurately predicting customer churn with a recall value of 0.86.

- Of all customers who will churn, the ANN Sequential model accurately predicts 86% of churn customers. And only 14% of churn customers are considered to be non-churn customers.

- Mistakes in predicting customer churn to be non-churn will harm the company because not only we lose revenue source, but it will also hinder our company's potential for growth because the cost of acquiring new customers is more expensive than retaining the existing ones.

Suggestion

- The marketing team can be more focused on the customers whose still single by creating new products or promotional programs that are suitable for them.

- The sales team can be more intensive in offering our additional service besides the main product because the customers who use our additional service are less likely to churn.

- Improve the quality of the fiber optic network that we own because the churned customer predominantly uses our product with fiber optic.

- Also, we can attempt to convert our customers whose on monthly contracts to yearly contracts by offering them more affordable prices but with an extended period of time.

Thanks!

Any questions?

For more details on this project, please visit:

→ https://github.com/H8-Assignments-Bay/p2---ftds-001-hck--ml1-dimitriasta