PREDICTING CUSTOMER CHURN IN THE TELECOM INDUSTRY

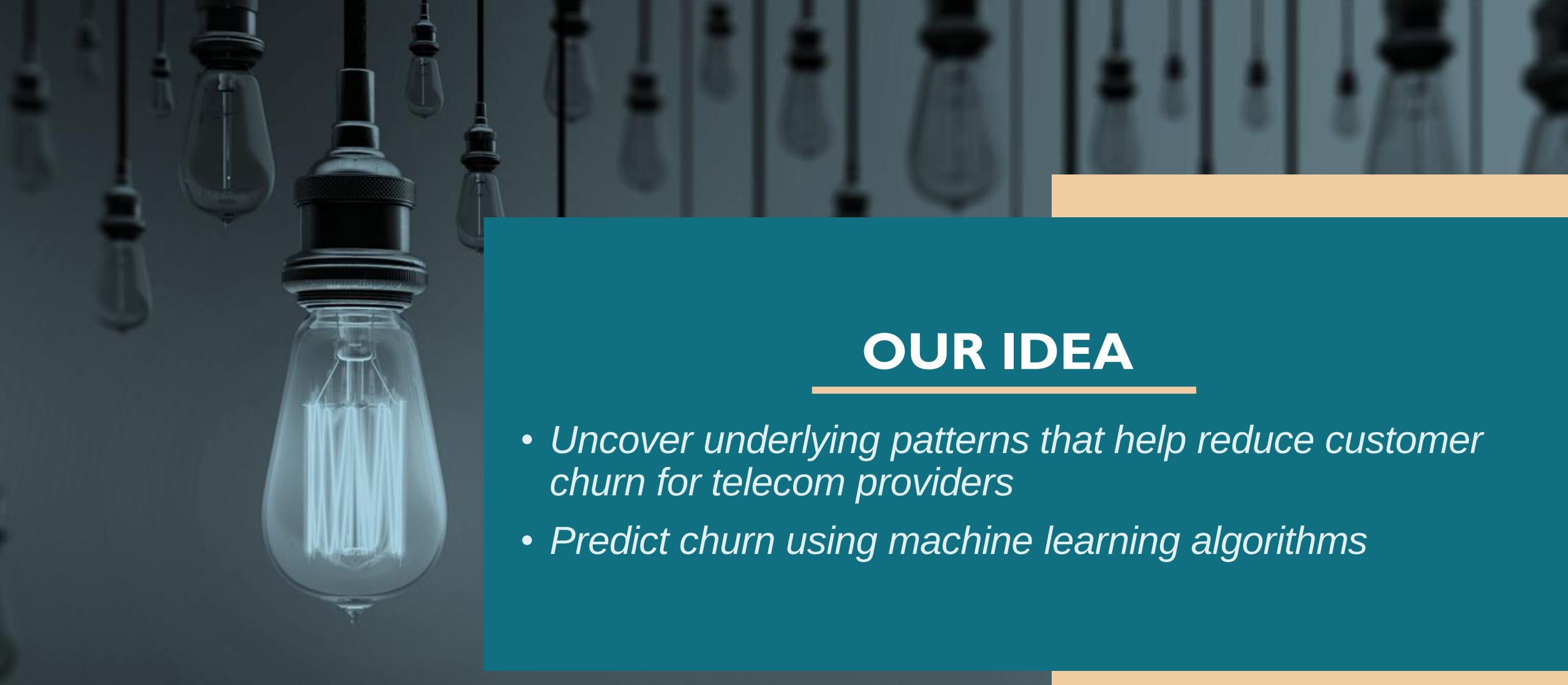
Brainstorming session

PROBLEM

Problem: Customer attrition becoming a major issue for telecoms affecting profits & growth

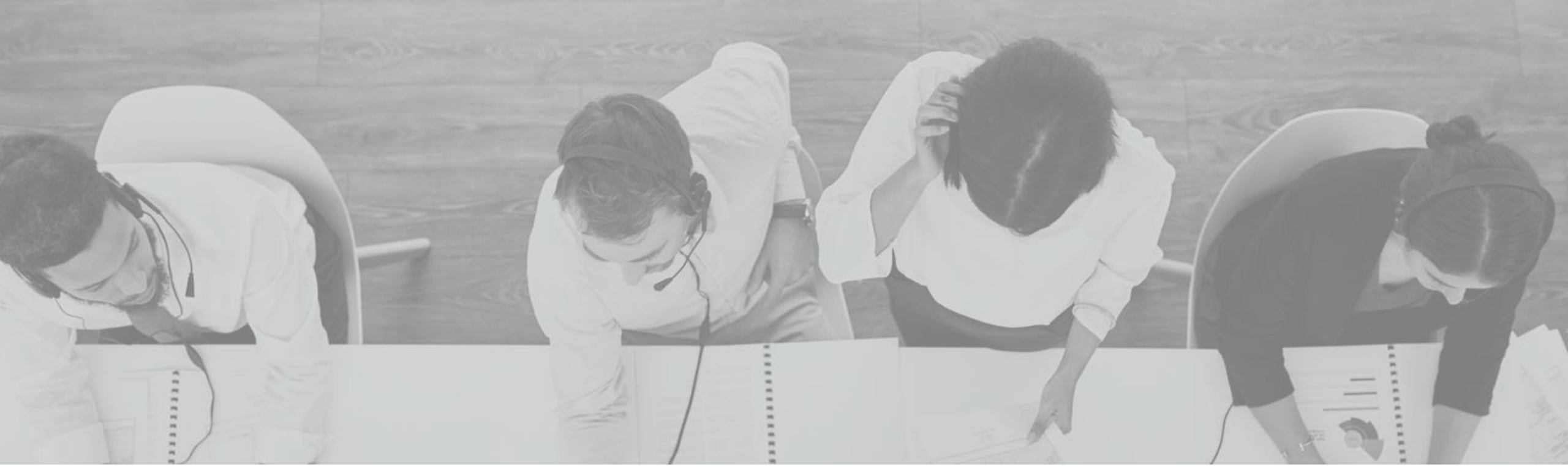
Objective: To identify churn factors and predict at-risk customers

Focusing on the following attributes



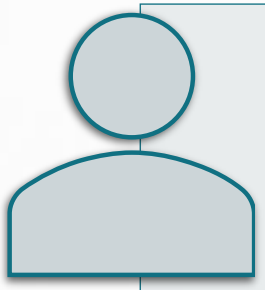
OUR IDEA

- *Uncover underlying patterns that help reduce customer churn for telecom providers*
- *Predict churn using machine learning algorithms*



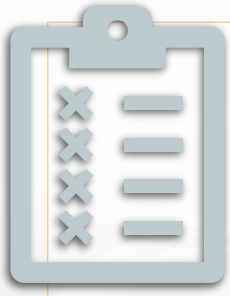
Dataset Overview

DATASET OVERVIEW



Predictor Variable – “Churn”

- Customers who left within the last month is called Churn
- We attempt to classify “Churn” as Yes / No



Unique values

- 7043

Features

- 20

DATASET | FEATURES

Services Opted

- Phone Service
- Multiple Lines
- Internet Service
- Online Security
- Online Backup
- Device Protection
- Streaming TV
- Streaming Movies

Account Information

- Customer ID
- Tenure (Yrs)
- Contract
- Monthly Charges
- Total Charges
- Payment Method
- Paperless Billing

Demographic Info

- Gender
- Age Range
- Senior Citizen?
- Have Partner?
- Have Dependants

DATASET | CATEGORICAL VARIABLES

Services Opted

- Phone Service
- Multiple Lines
- Internet Service
- Online Security
- Online Backup
- Device Protection
- Streaming TV
- Streaming Movies

Account Information

- Contract
- Payment Method
- Paperless Billing

Demographic Info

- Senior Citizen?
- Have Partner?
- Have Dependants

DATASET | FEATURE SELECTION

Services Opted

- Phone Service
- Multiple Lines
- Internet Service
- Online Security
- Online Backup
- Device Protection
- Streaming TV
- Streaming Movies

Account Information

- Contract
- Payment Method
- Paperless Billing

Demographic Info

- Senior Citizen?
- Have Partner?
- Have Dependants

DATASET | FEATURE SELECTION

Services Opted

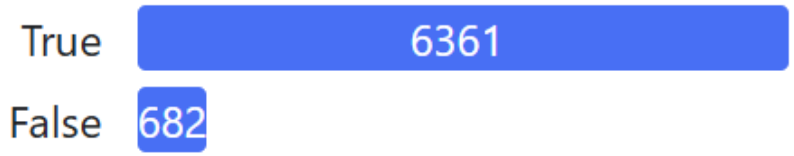
Phone Service

PhoneService

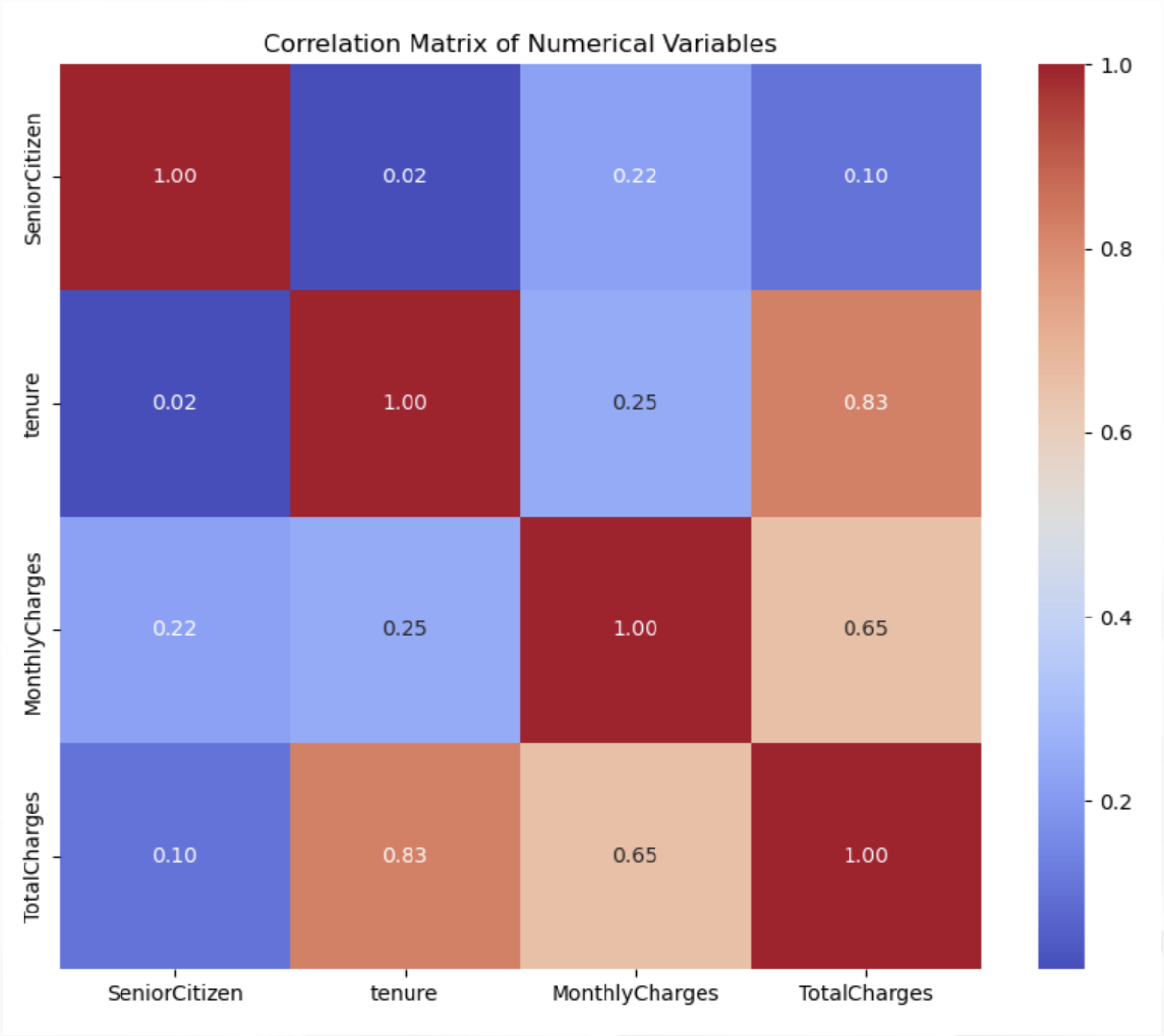
Boolean

High correlation | Imbalance

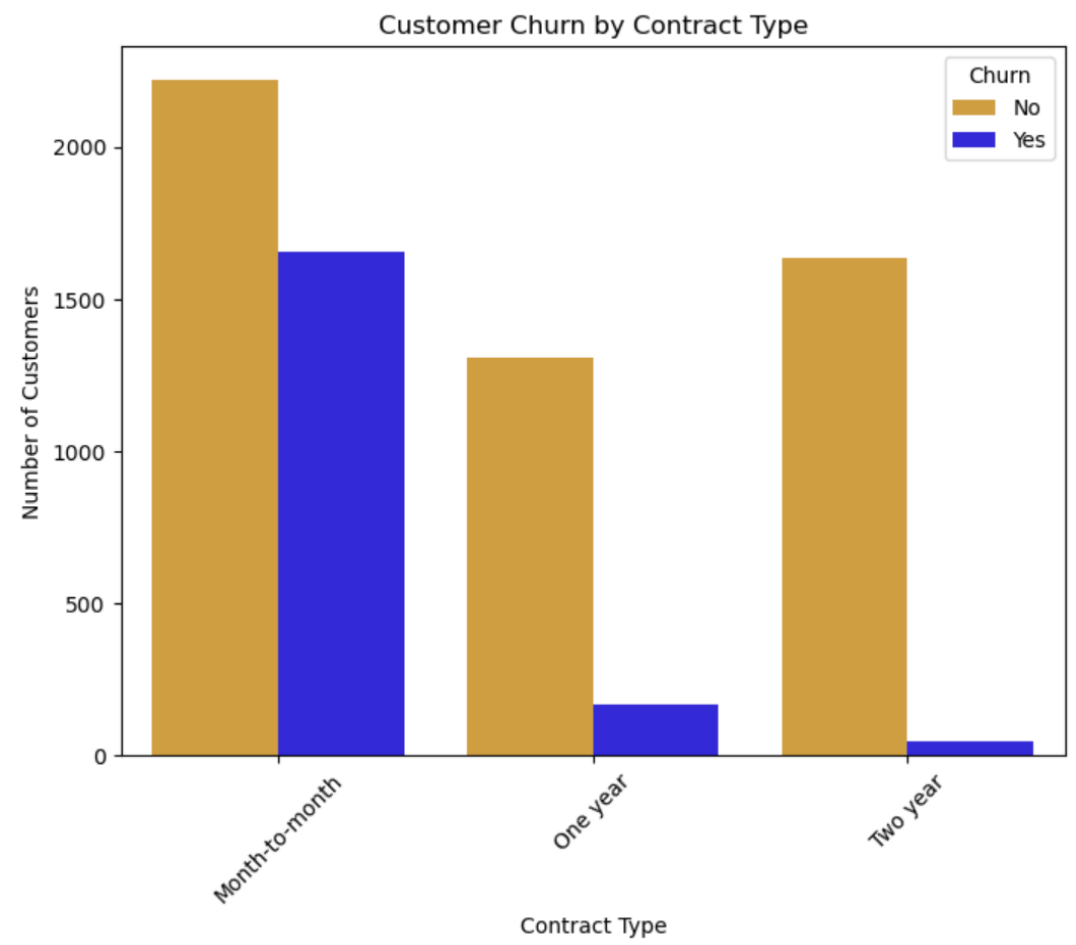
Distinct	2
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	7.0 KiB

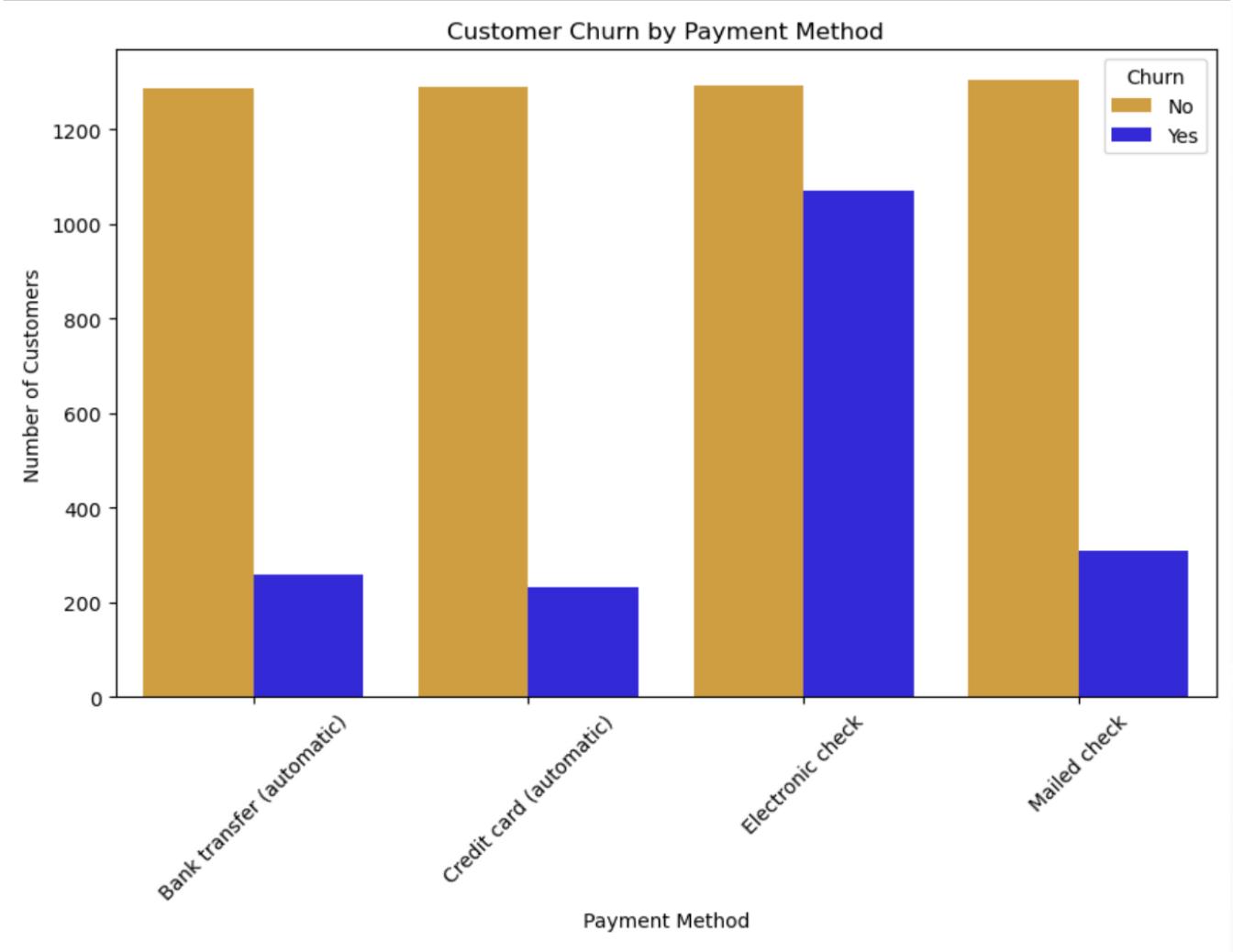


VISUALIZATION EVALUATION | CHURN HEATMAPS



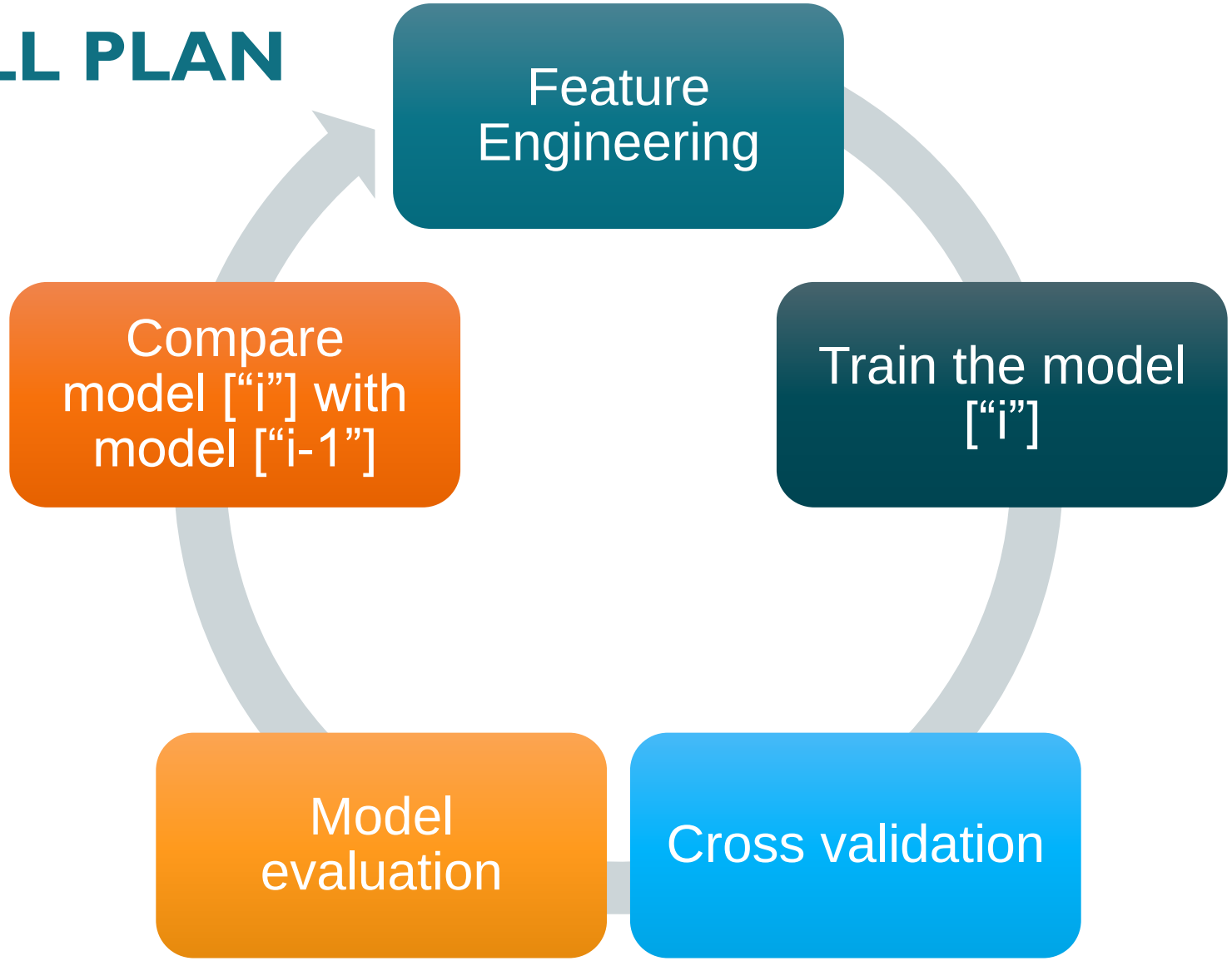
VISUALIZATION EVALUATION | CHURN vs CONTRACT TYPE | PAYMENT METHOD



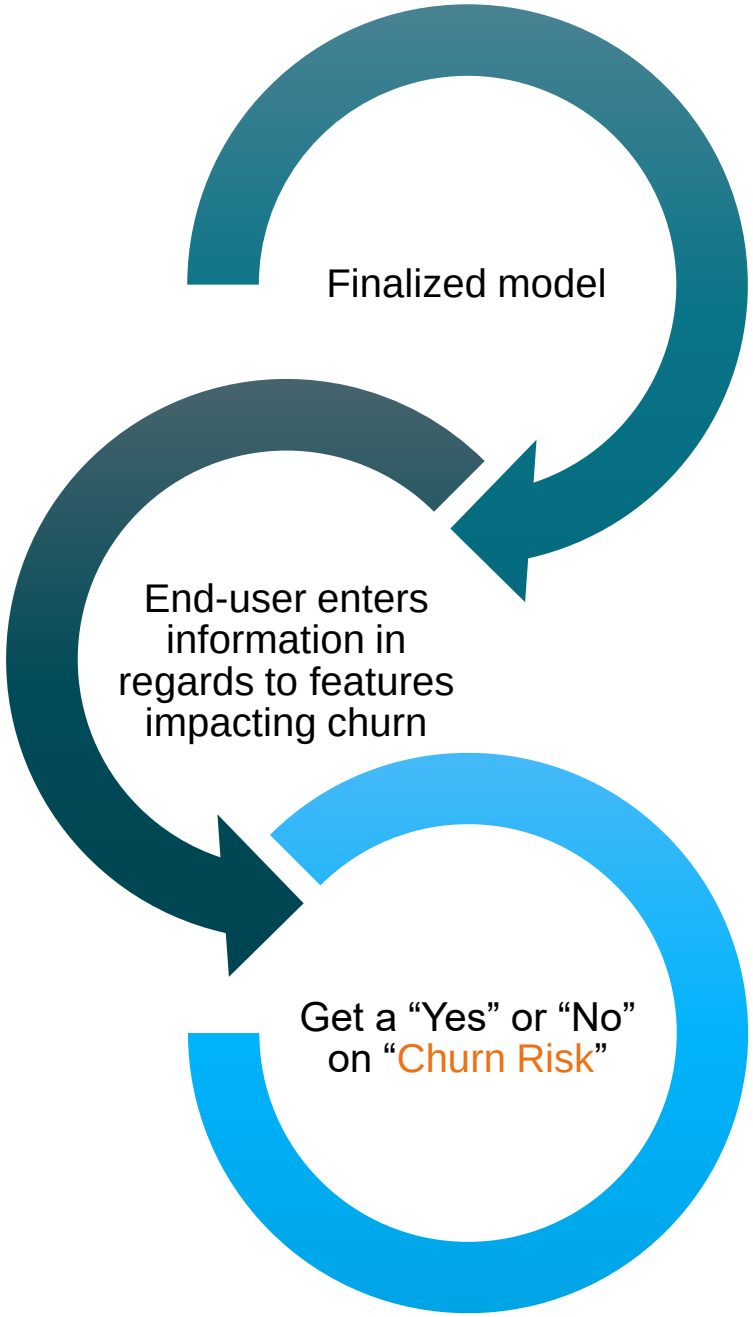


Plan for Analysis and Visualization

OVERALL PLAN



OVERALL PLAN



Exploratory Data Analysis

- Encoding categorical variables

Feature Engineering

- Adding interaction terms
  - Contract type + customer satisfaction
  - Geographic location + churn reason

Churn Risk Segmentation

- Develop a risk score based on
  - Demographics
  - Satisfaction
  - Contract type

Potential Machine Learning Models

Logistic Regression

Acts as base model for classification

Potential Machine Learning Models

SVM

To capture non-linear relationships

Act as benchmark for complex decision boundaries

Potential Machine Learning Models

- **Decision Trees with Random Forest**

  Provide predictions while handling feature interactions

Potential Machine Learning Models

- **Decision Trees with Random Forest**

  Provide predictions while handling feature interactions

Potential Machine Learning Models

- Decision Trees with CatBoost

  Handling categorical variables on tabular data

  Reduce overfit with minimal parameter tuning

Thank you!