DSFT-07 PHASE 3 PROJECT PRESENTATION

PROJECT TITLE:Detecting customers who are likely to churn

By Mainye Gesanda Peter

PROJECT OVERVIEW

- This Dataset contains information about customers of a Telecom company and whether they churned or not.
- Churn refers to the phenomenon of customers leaving a service or product. The dataset includes various features that can be used to predict churn behavior

Problem statement

- The problem we aim to address in this project is predicting customer churn in a telecom dataset.
- By analyzing customer data such as usage patterns, demographics, and customer service interactions, we aim to develop a predictive model that can accurately identify customers who are likely to churn.
- This model will enable telecom companies to proactively intervene and implement retention strategies to reduce churn rates

DATA

- Data Source: https://www.kaggle.com
- Target variable: Churn
- Target audience: Telecom company

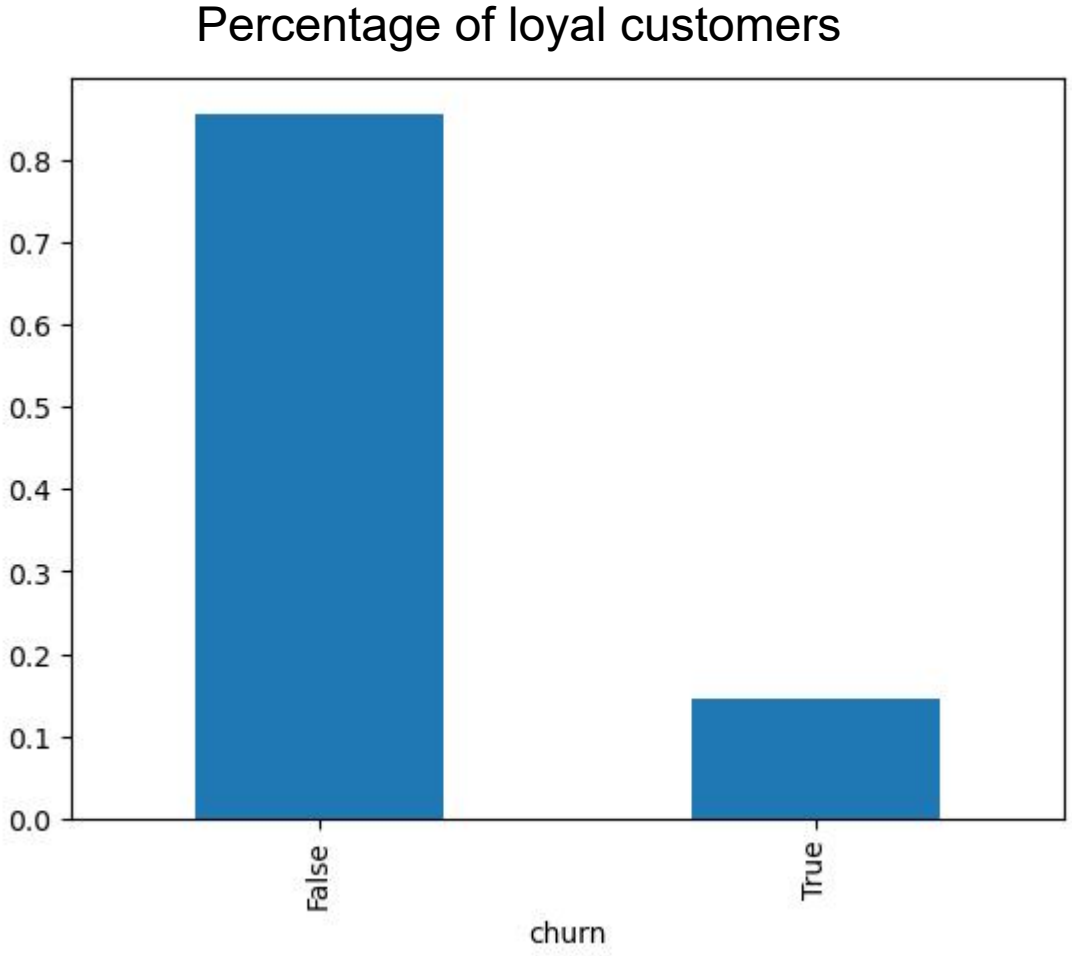
PROCEDURES

- Importing all necessary packages
- Loading the dataset
- Cleaning the data
- Exploratory Data Analysis
- Modelling approach

Main Objective

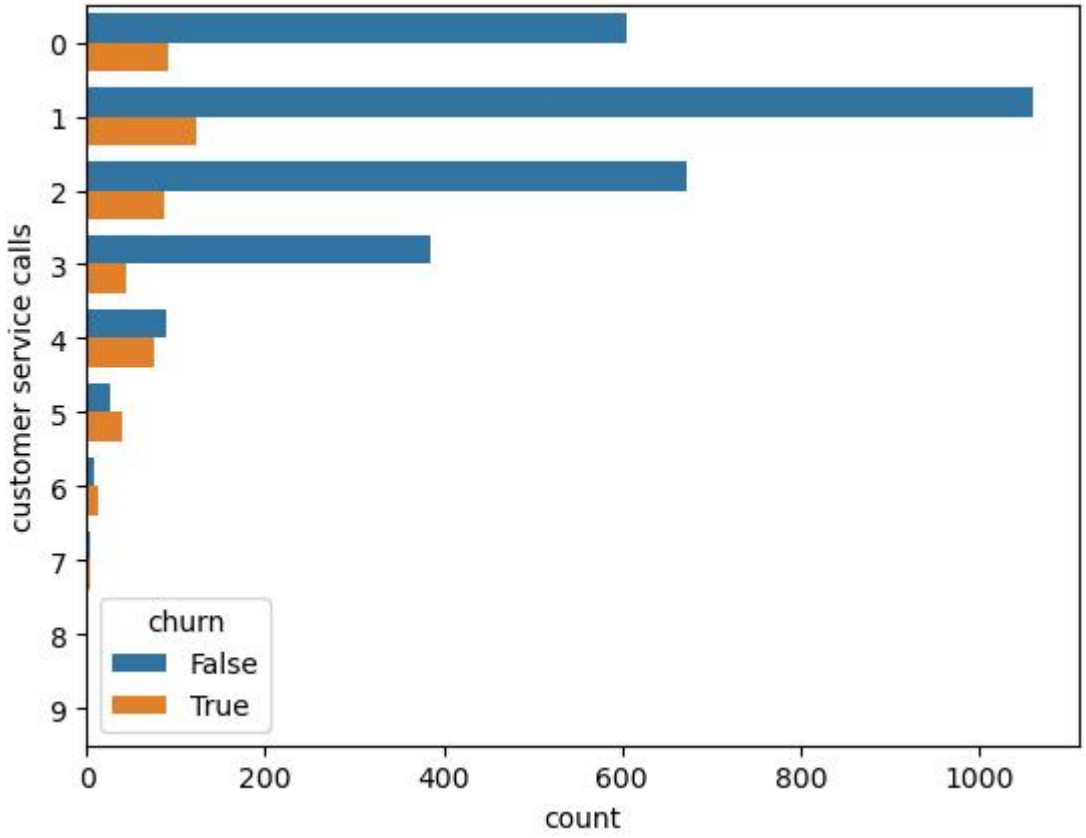
- Build a classifier to predict whether a customer will ("soon") stop doing business with SyriaTel, a telecommunications company

VISUALIZATIONS OF THE FINDINGS



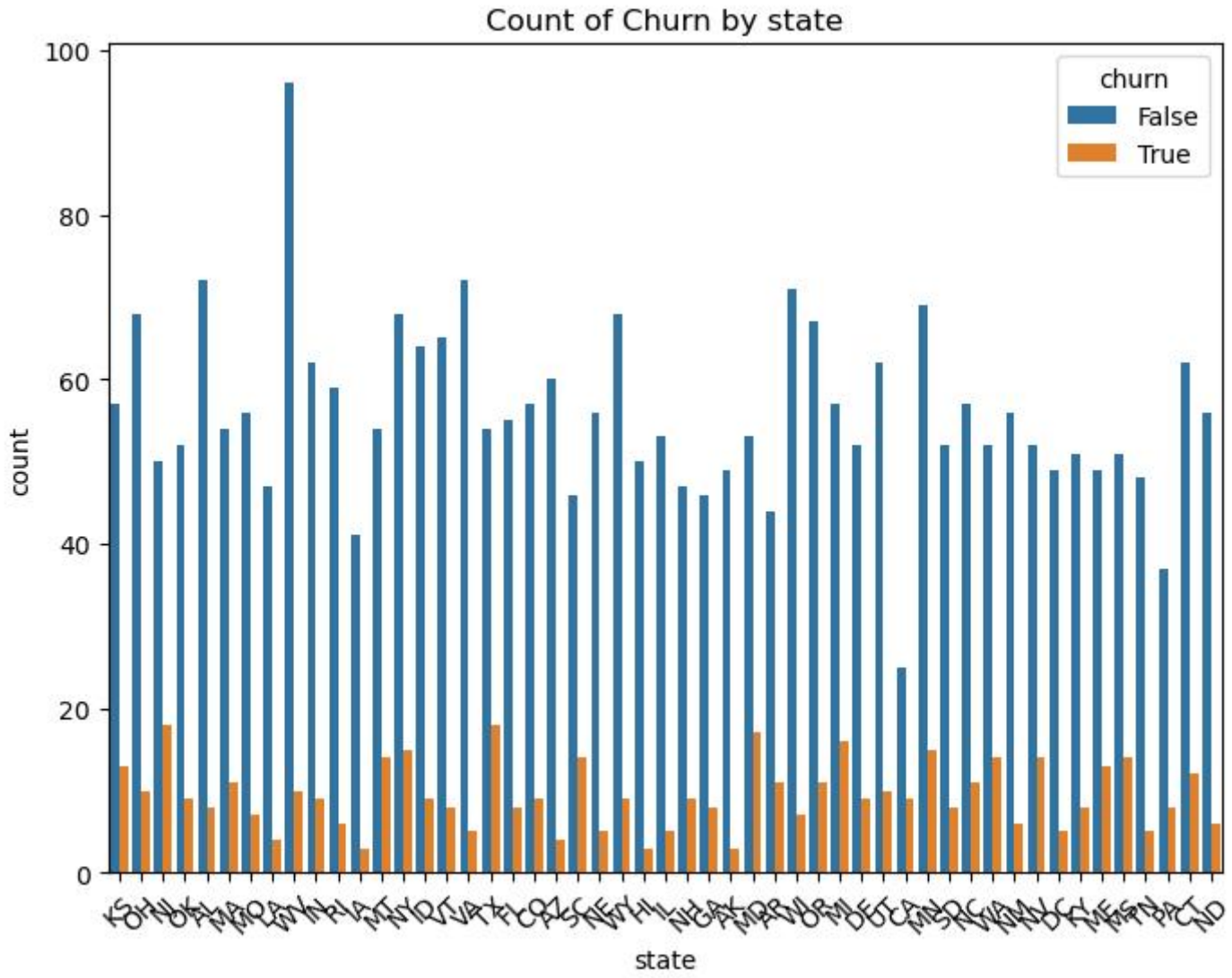
Call center and churn

Relationship between the number of calls to the call center and Churn



CHURN AND STATE

Co



CONCLUSION

- Precision: For the class labeled as False, the precision remains the same as before hyperparameter tuning, indicating that 91% of the customers predicted to not churn actually did not churn. For the class labeled as True, the precision remains perfect at 100%, meaning that all customers predicted to churn actually did churn.
- Recall: After hyperparameter tuning, the recall for the class labeled as False remains perfect at 100%, indicating that all actual instances of not churning were correctly identified by the model. However, the recall for the class labeled as True has improved to 46%, compared to the previous recall of 39%. This suggests that the model is now better at identifying actual instances of churn.
- F1-score: The F1-score for both classes has improved slightly, with the F1-score for the class labeled as True increasing from 0.56 to 0.63.
- Accuracy: The overall accuracy of the model has also improved slightly from 91% to 92%

RECOMENDATION

- As a recommendation, it's advisable to deploy the refined model into a production environment for real-world usage.


THANK
YOU !