Dayananda Sagar Academy of Technology and Management

Accredited by National Assessment & Accreditation Council (NAAC) with ‘A’ Grade, an Autonomous Institute Affiliated to VTU, Approved by AICTE & UGC, NIRF RANK 2024

Department of Information Science and Engineering

INNOVISION

Presents

TECH TRIAD

12 hours Datathon

TEAM DETAILS

- TEAM NAME: The Gradient Descendants
- THEME : Industry , Innovation and Infrastructure
- TEAM MEMBERS DETAILS:

NAME:	EMAIL:	PHONE NO:	USN:
Yashwanth B L	yashwanthbloo27@gmail.com	9731895061	1JT23CS190
Mehaboob S	sameernaseema175@gmail.com	8050425367	1JT23CS090

PROBLEM STATEMENT

THEME : Customer Retention in Telecom
SDG-9 Industry, Innovation and Infrastructure

Predict whether a customer will churn (leave the service) based on demographic details, service subscriptions, billing information, and engagement patterns.

Recognizing key data points of customer activity

Analysing the key trends

An ML model trained to CLASSIFY customers based on the available infrastructure and their living conditions.

Customer retention using machine learning.

DATASETS OVERVIEW

- The provided dataset is of a Telecom organization where each row represents demographic information, account details and service usage patterns of each customer.
- In the provided dataset “Churn” is the target variable and the remaining 20 features are the predicting variables.
- CustomerID is a unique ID provided to each customer and can be removed as it has no correlation between the churn rate and CustomerID.
- We have observed that the “MutipleLines” feature is dependent on the “PhoneLine”. So if PhoneLine is false then the MultipleLines feature is false by default.
- We have observed that features like “TechSupport”, “OnlineSecurity”, “StreamingTV”, “OnlineBackup”, “DeviceProtection”, “StreamingMovies” are all dependent on “InternetService” feature so if InternetService is false then the above mentioned features are also false for the respective customers.
- The categorical features need to be encoded in numerical form for statistical predictions.
- About 5000 customers have True value for Churn and only 2000 customers having False value for Churn. We can say that the given dataset is skewed.

PROPOSED SOLUTION

- We’re building an intelligent Decision Support System for telecom organizations that predicts which customers are most likely to migrate to other service helping teams with smart, data-driven recommendations to enhance service delivery and optimize infrastructure investments.

- Build a classification model which predicts whether a customer will be retained or migrate to a different service based on his/her usage pattern of demographics, account details, and infrastructure availability.

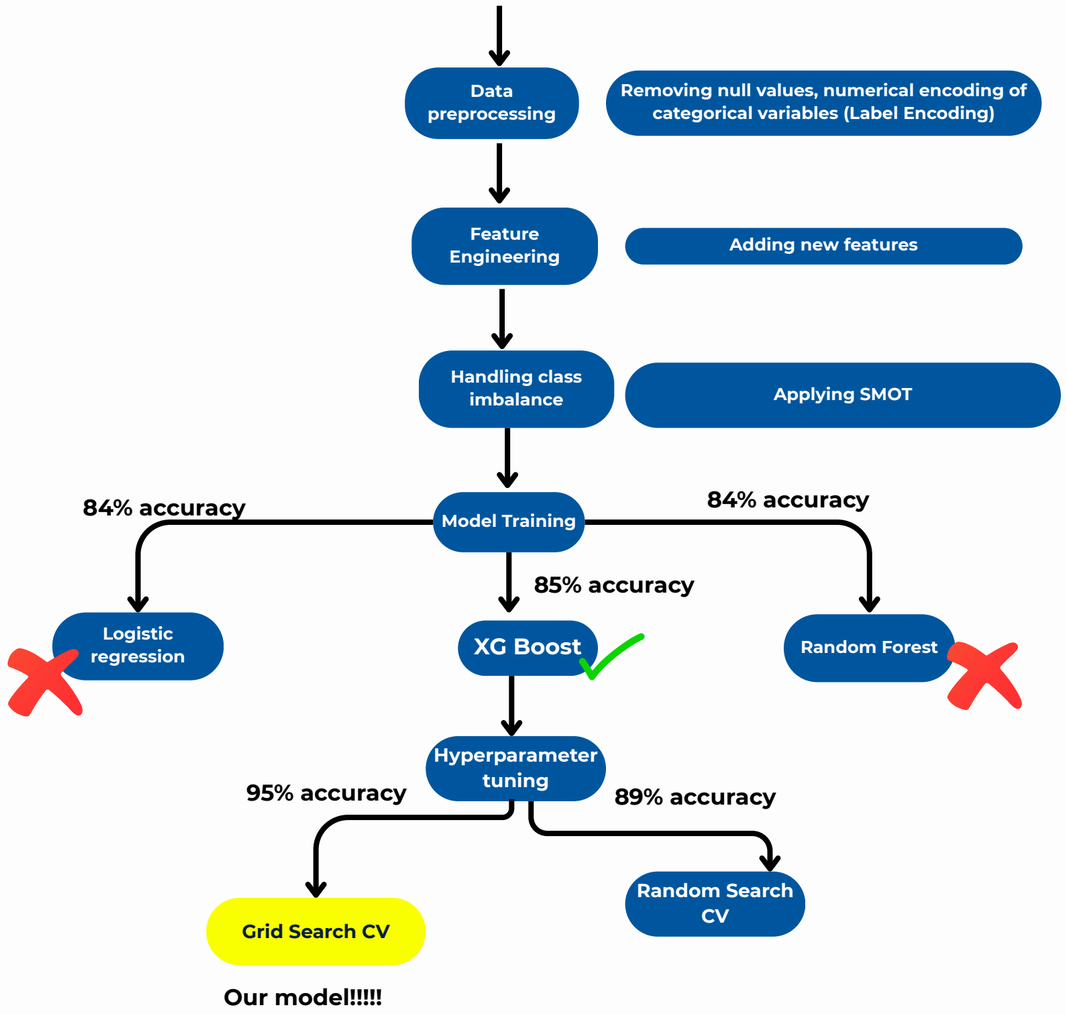
- Build a web based interface where an Telecom Organization enter the customer demographic details and reason if the customer will use the same service or not.

- We will be building an alert system highlighting the next crucial step to take in order to retain a customer.

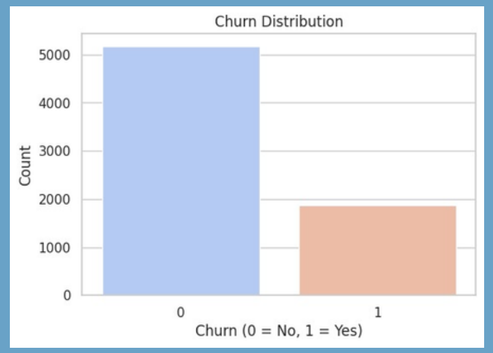
DATA PIPELINE/WORKFLOW

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService
0	7590-VHVEG	Female	0	Yes	No	1	No	No	DSL
1	5575-GNVDE	Male	0	No	No	34	Yes	No	DSL
2	3668-QPYBK	Male	0	No	No	2	Yes	No	DSL

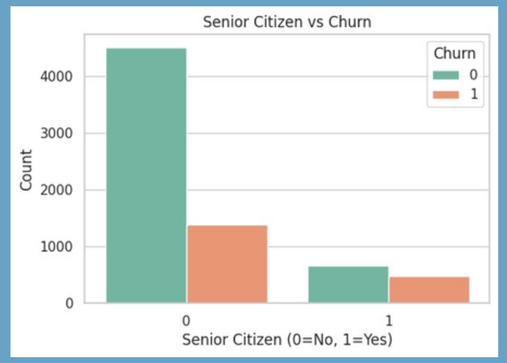
Quarterly datasets of customers



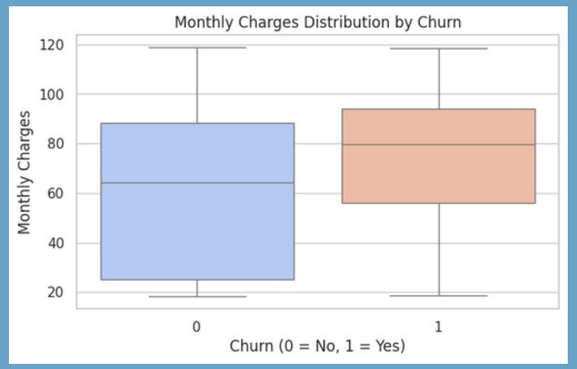
EDA INSIGHTS



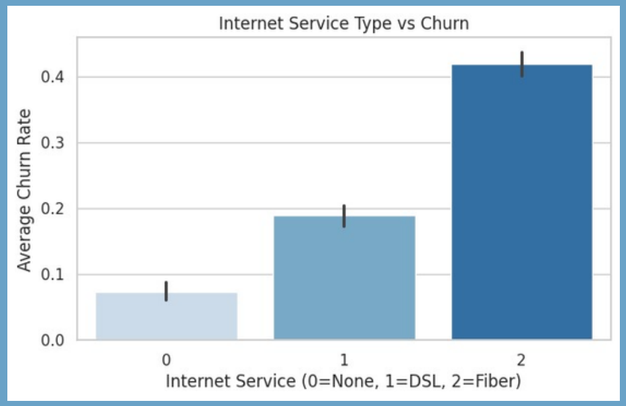
Skewed Dataset observed



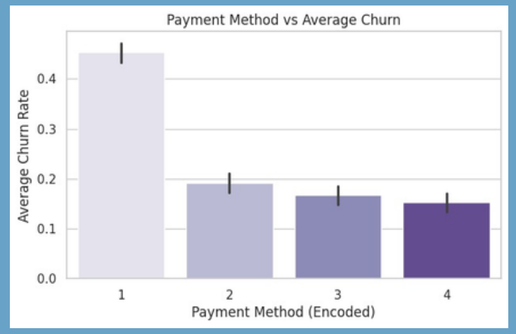
Barchart between senior citizens and Churn rate



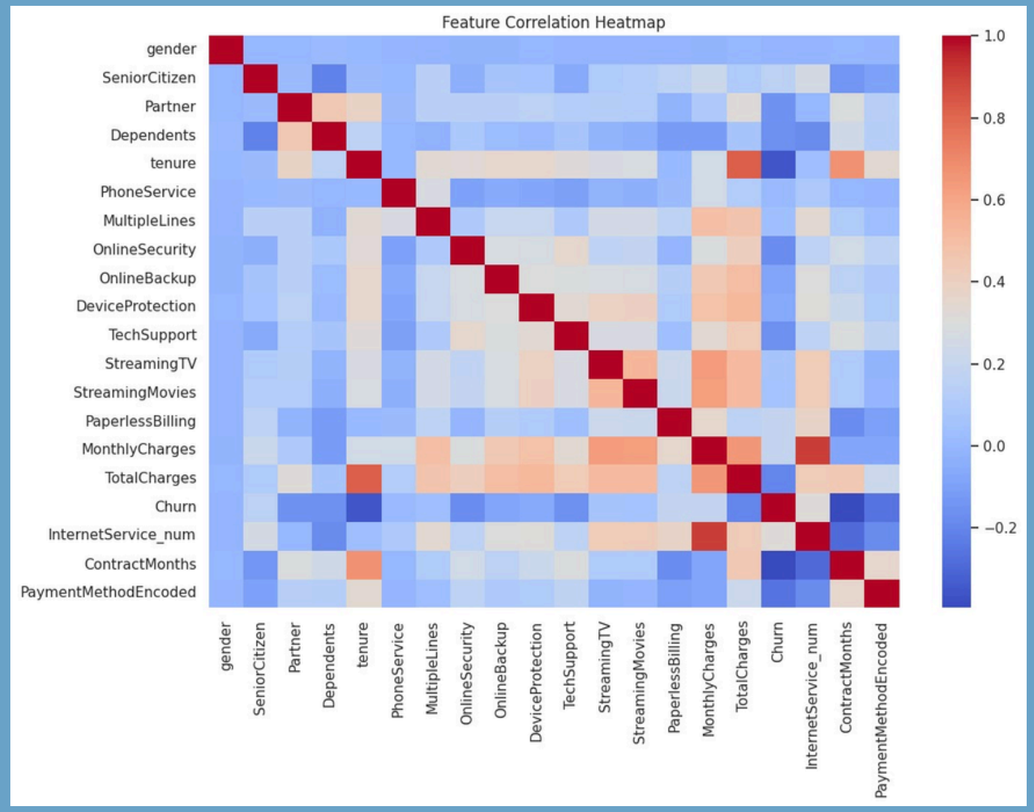
Boxplot between monthly charges and Churn rate



Bargraph between internet service and churn rate

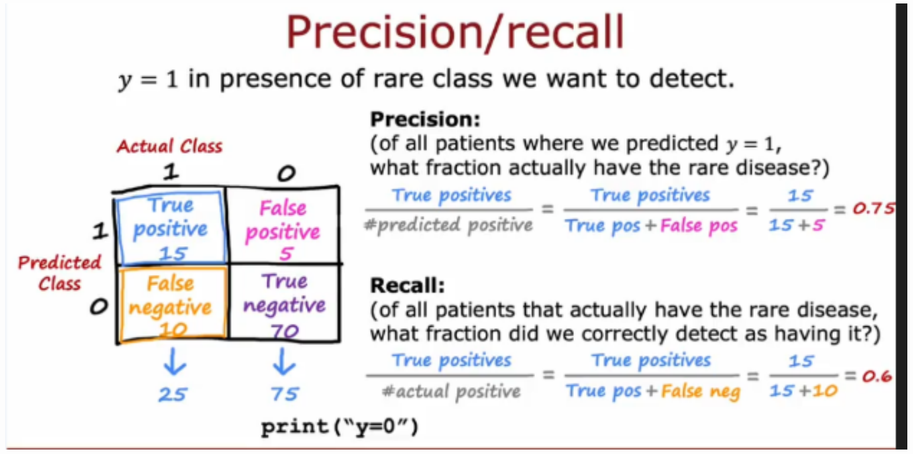


Bargraph showing customers and their payment methods

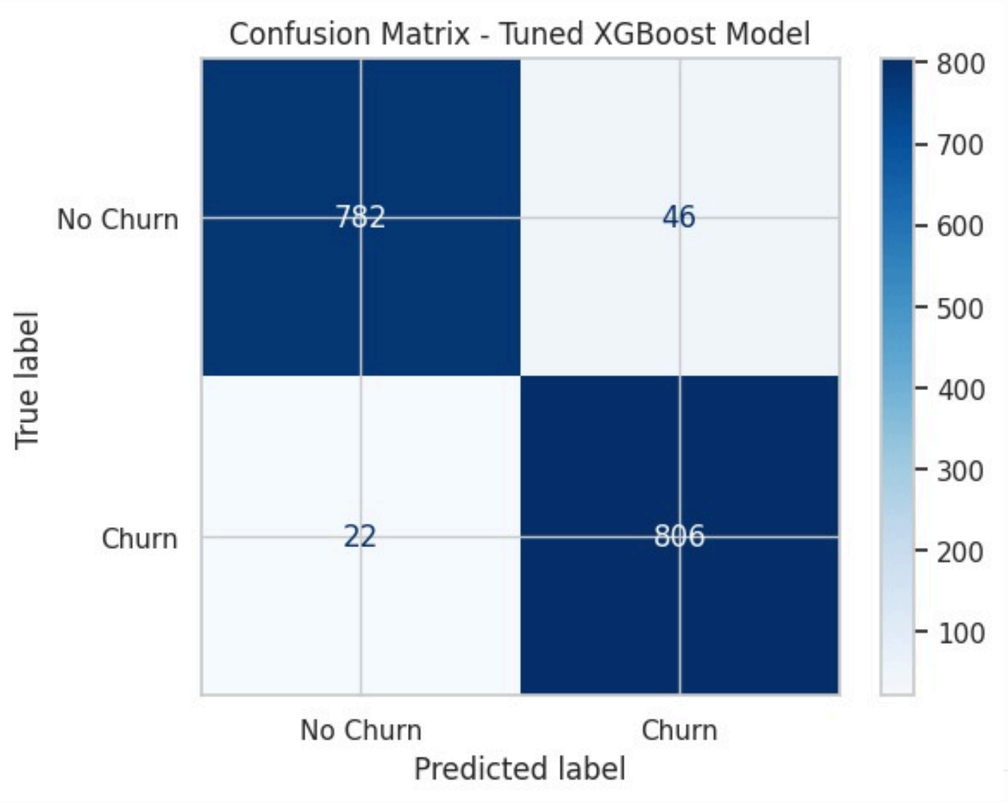


The final Heatmap showing the correlations between target variables and features.

MODEL ANALYSIS AND RESULTS



- **Goal is to maximize the f1 value for a classification model**



- **Our model performance confusion matrix**

TECH STACK

Frontend

Backend

IMPACT

(How does your idea create value?)

Predicts churn early – act before customers leave
Data-driven insights – smarter business decisions
Optimizes infrastructure – invest where it matters most
Improves customer experience – reduce complaints, increase satisfaction
Provides actionable recommendations – clear next steps for teams
Enables targeted campaigns – focus on high-risk customers
Learns continuously – improves accuracy over time