

End of Month Churn Review

A Statistical Analysis of Telcom Churn Data





Agenda

- Data Overview
 - Key Features
- Exploratory Data Analysis
 - Feature Engineering
- Model Performance
 - Metrics
- Conclusions
 - Areas of Focus





Data Overview

What is churn?

- Individuals who have left within the last month.
- Binary classification problem
- Inherently imbalanced

7043 Customers

- 5174 Retained
- 1869 Churned (26.5%)

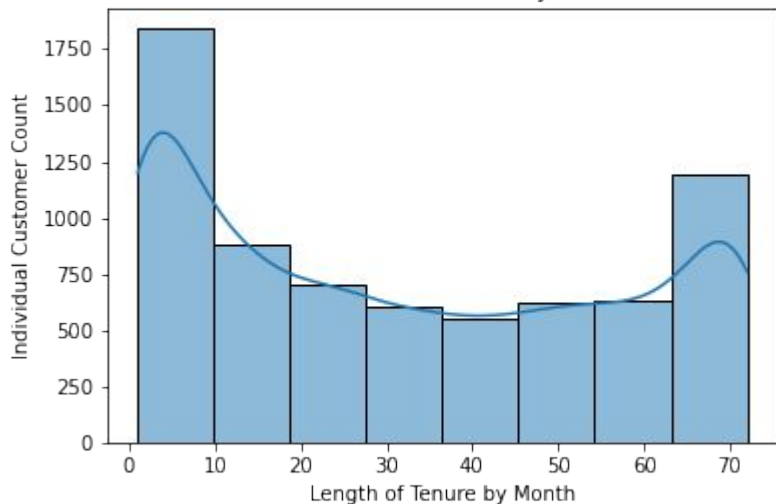
Key Features

- Tenure & Contract
 - Negative Coefficients
 - Significant Z-Scores
- Services
 - Streaming
 - Online Support
 - Technical Support
 - Internet Service
 - Phone Service

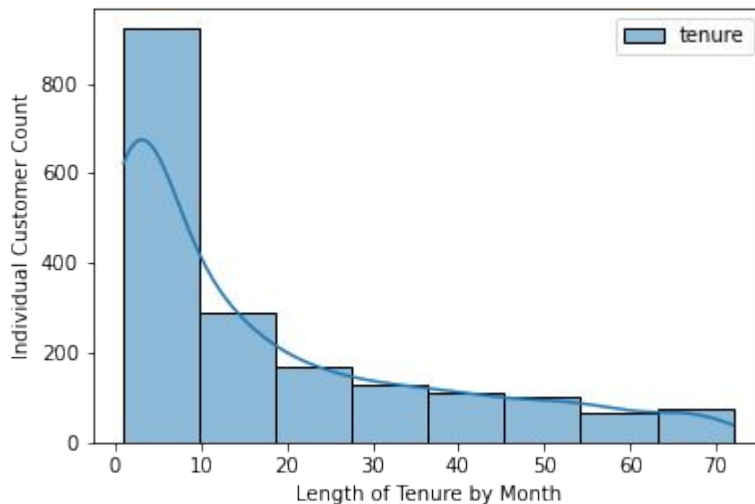


Drilling into Tenure

Distribution of Tenure by Month



Distribution of Churned Customer Tenure





Exploratory Data Analysis - Feature Engineering

Dropped

- Gender
- MultipleLines
- PaymentMethod
- Partner

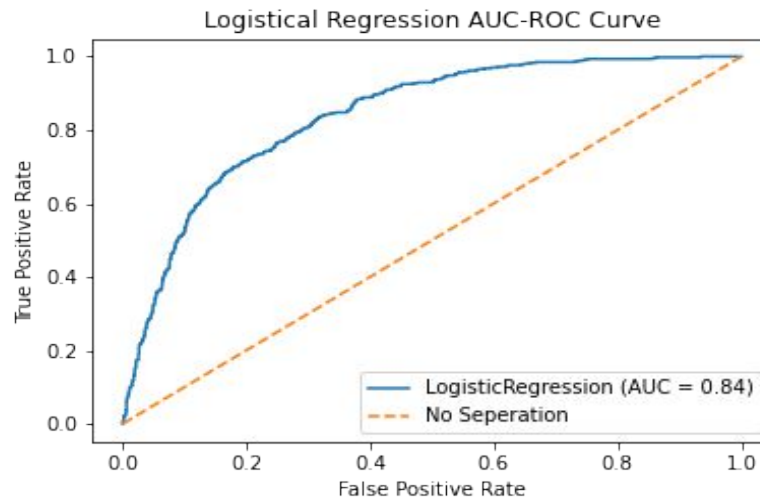
Combined

- OnlineSecurity
- OnlineBackup
- DeviceProtection
- TechSupport
- StreamingTV
- StreamingMovies
- InternetService
- Contract



Model Performance - Logistic Regression

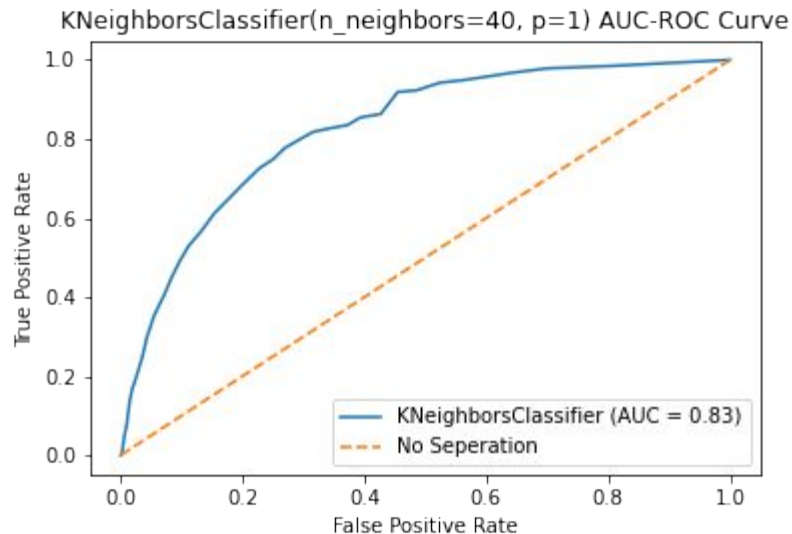
- Original
 - Training Accuracy: 80.39%
 - Testing Accuracy: 80.60%
 - AUC: 0.84
 - Recall: 0.81
- Feature Engineered
 - Training Accuracy: 80.05%
 - Testing Accuracy: 80.15%
 - AUC: 0.84
 - Recall: 0.80





Model Performance - K-Nearest Neighbors

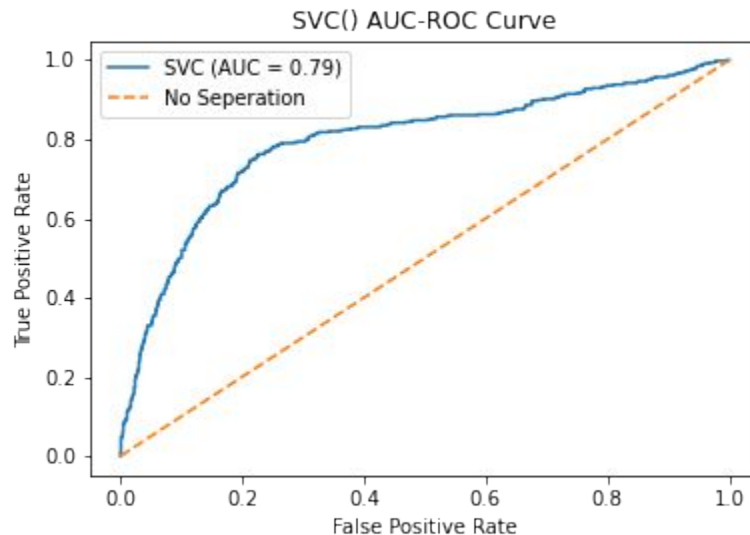
- Original
 - Training Accuracy: 80.38%
 - Test Accuracy: 78.78%
 - AUC: 0.83
 - Recall: 0.79
- Feature Engineered
 - Training Accuracy: 83.92%
 - Test Accuracy: 76.28%
 - AUC: 0.82
 - Recall: 0.78





Model Performance - Support Vector Machine

- Original
 - Training Accuracy: 81.82%
 - Test Accuracy: 80.03%
 - AUC: 0.80
 - Recall: 0.80
- Feature Engineered
 - Training Accuracy: 80.41%
 - Test Accuracy: 79.86%
 - AUC: 0.79
 - Recall: 0.79





Conclusions

- Modern modelling methods such as Support Vector Machines do not outperform Logistic Regression models on predicting this set of churn data.
- We should examine the account details of the senior citizen customers who have churned over the last quarter to expand upon our findings regarding that demographic.
- Streaming services is our lowest performing service, which may be due to competition like Netflix and Hulu. We should do a cost analysis on maintaining our Streaming services going forward.
- Prioritize customer contact during the first 12 months, regardless of contract type, but especially if they are on month-to-month.