

Survival Analysis on Telco Customer Churn

Prepared by:

Steve Song MBA

April 22, 2020



About customer churn

- What is customer churn?
the percentage of customers that stop using a company's products or services;
- Why is it critical to the business?
it usually costs more to acquire new customers than it does to retain existing ones.

What is the project goal to achieve?

- Identify the importance of each predictor;
- Build the predictive model using survival analysis;
- Predict risk to churn on unseen new customers
- Launch customer retention program to improve ROI

Telco customer churn data

- size of data : 7,043 unique customers

- number of features: 21 features

tenure, churn: tenure-event as target variables

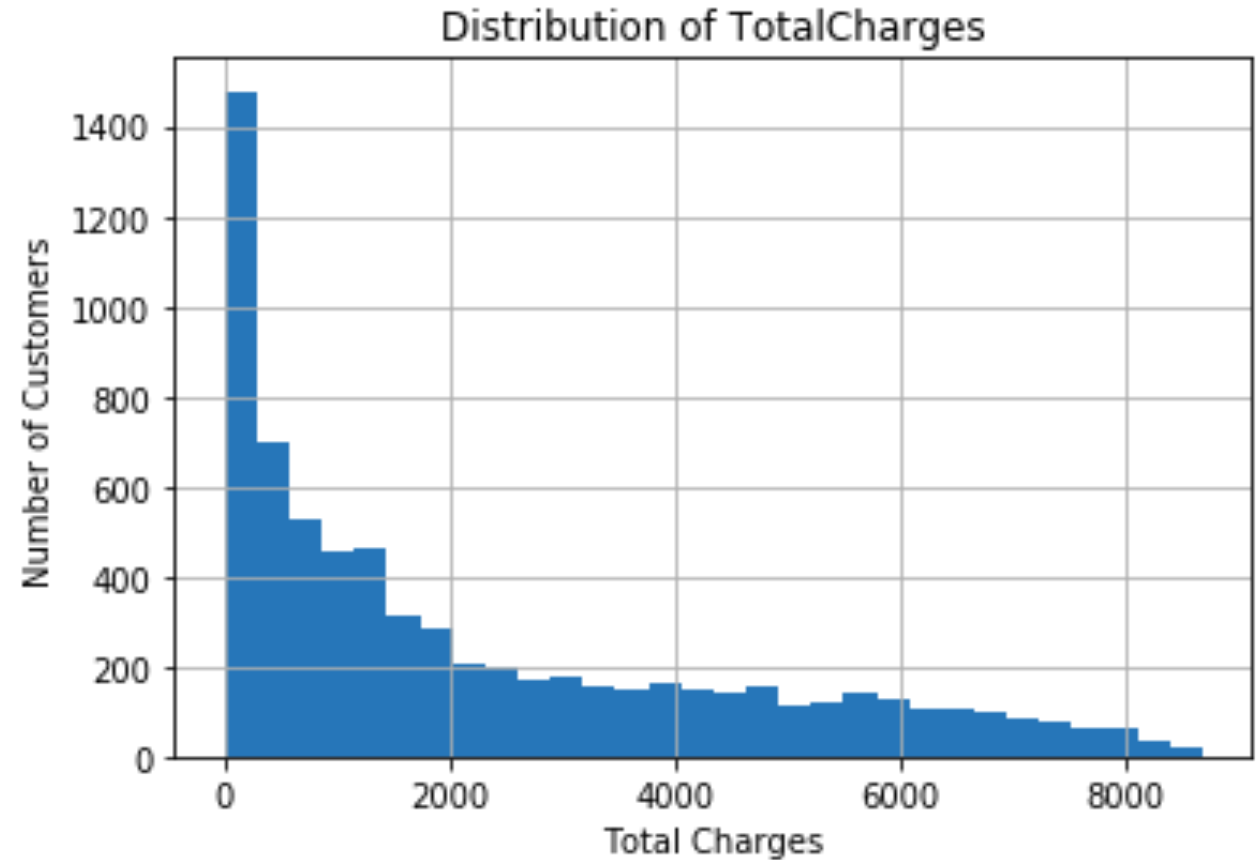
customer attributes: gender, SeniorCitizen, partner, Dependents

consumer behavior: PhoneService, MultipleLines, InternetService
OnlineSecurity, OnlineBackup, DeviceProtection,
TechSupport, StreamingTV, StreamingMovies,
Contract, PaperlessBilling, PaymentMethod
MonthlyCharges, totalCharges

customerID: random serial number, no value for prediction

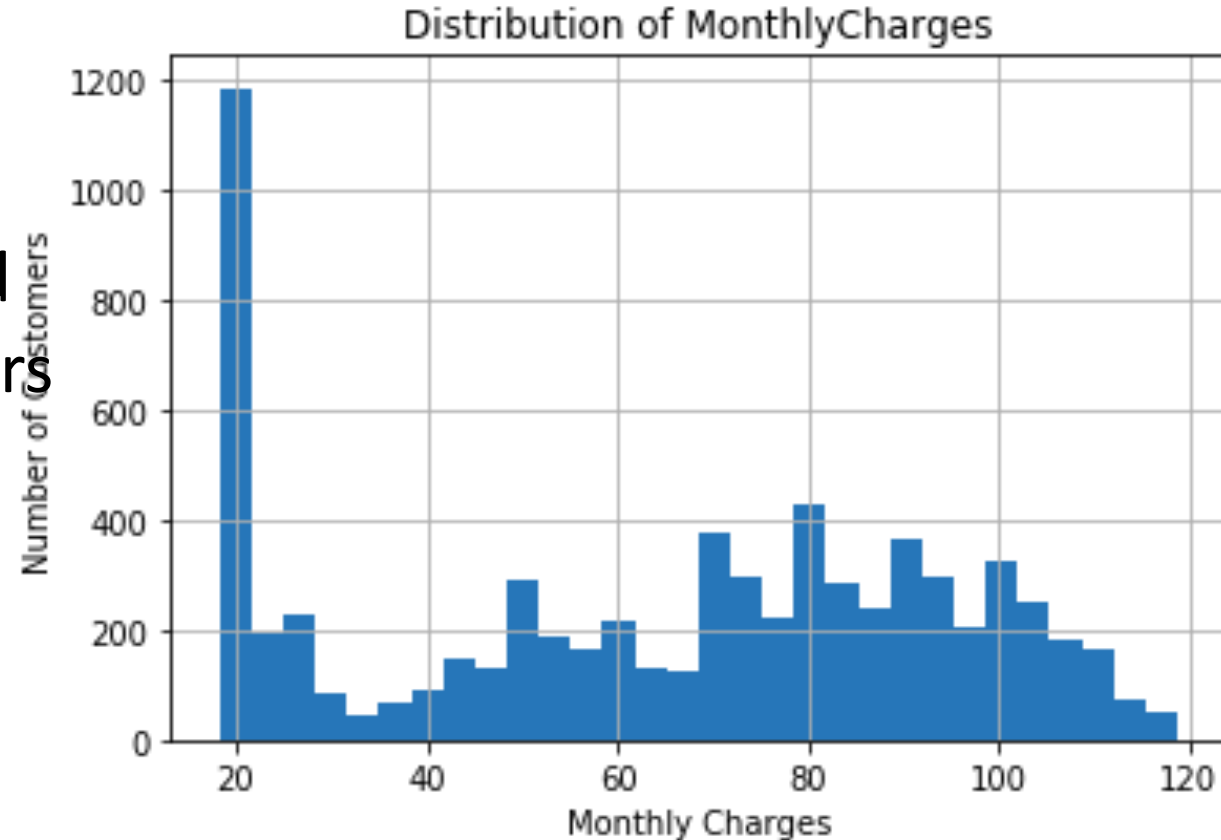
Distribution of Total Charges

- The distribution positively skewed.
- The majority of customers charged under \$2000.
- This feature is unknown at the time of prediction. We will discard it to avoid data leakage



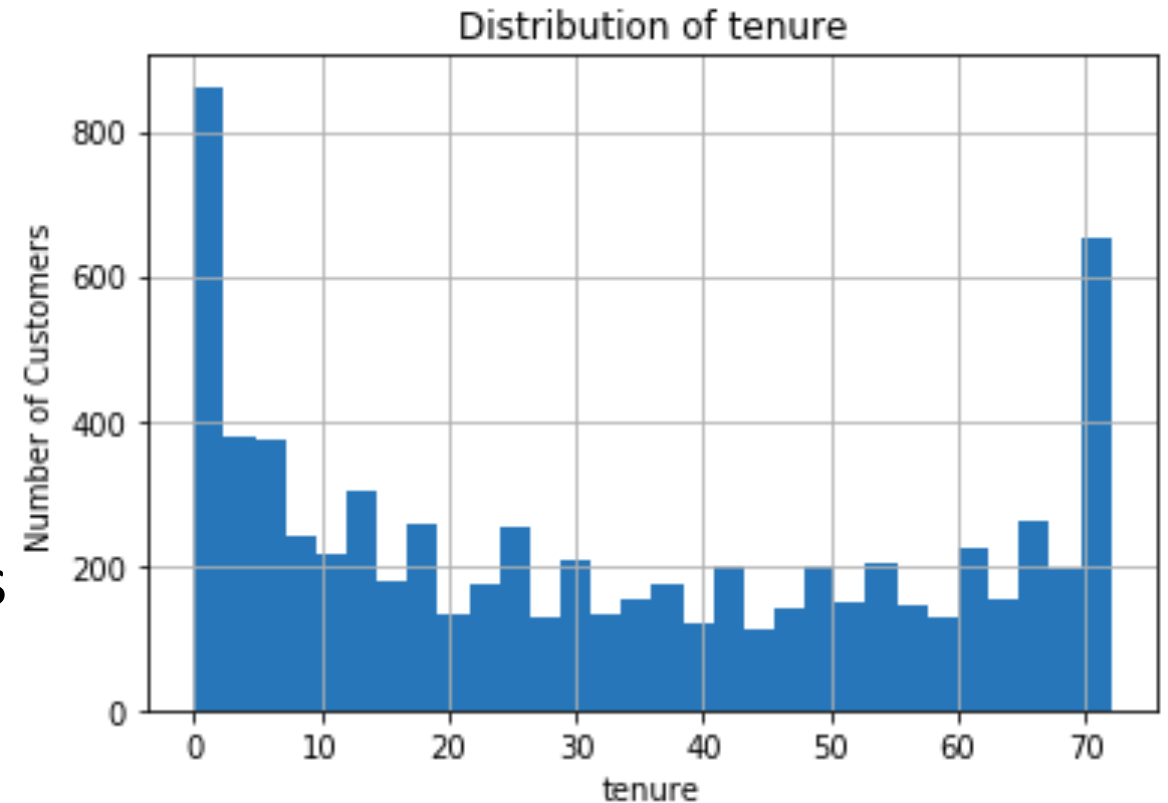
Distribution of Monthly Charges

- Normally distributed above \$30 beyond
- Significant number of low-end customers spent less than \$30 monthly

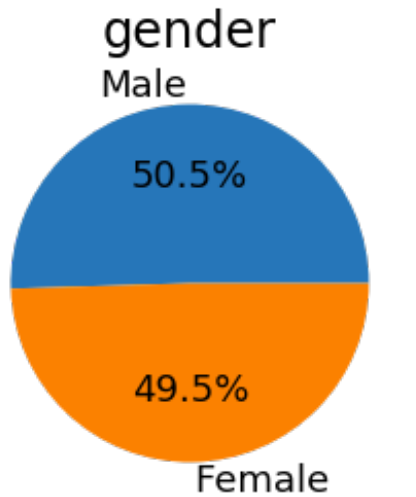


Distribution of tenure

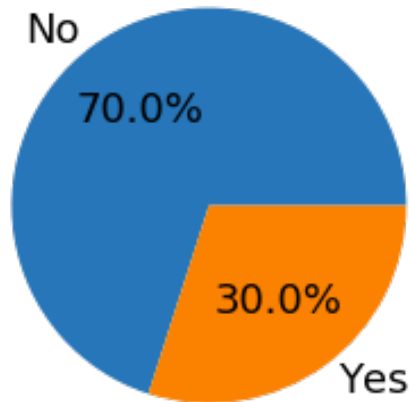
- relatively stable trend between 10 months and 60 months.
- significant number of customers stayed with business for 10 months (new customers) or above 60 months (loyal customers).



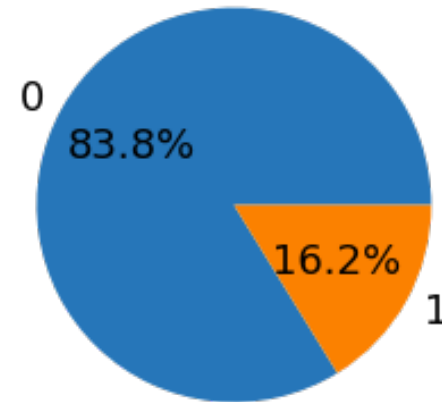
Customer attributes analysis



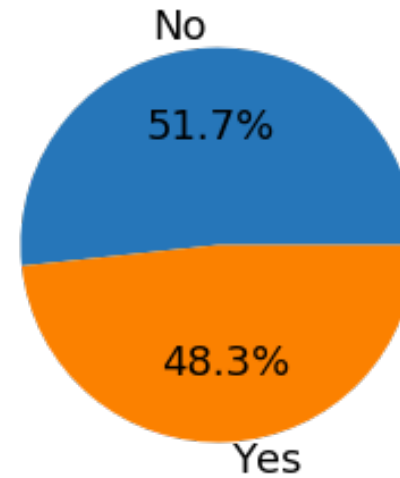
Dependents



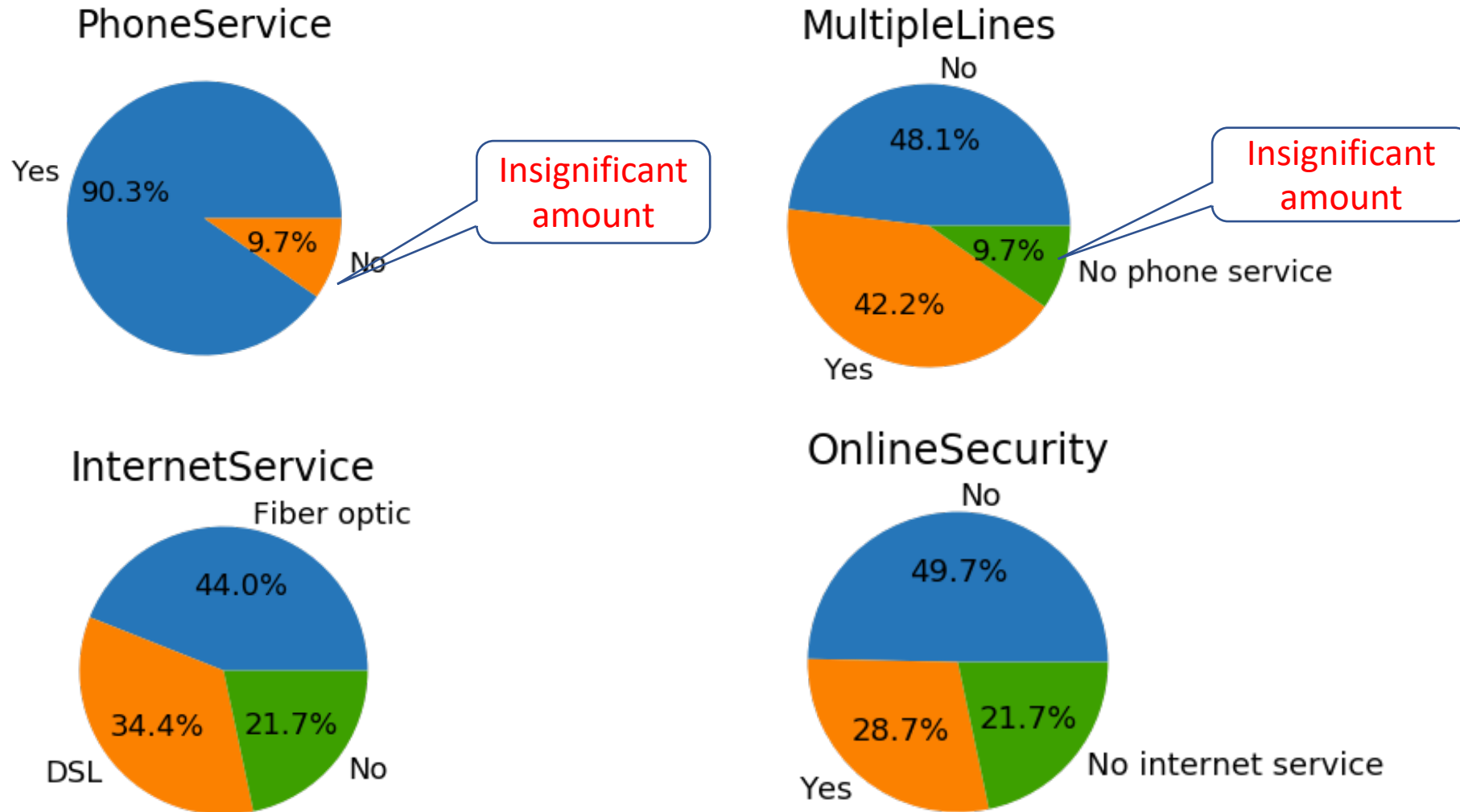
SeniorCitizen



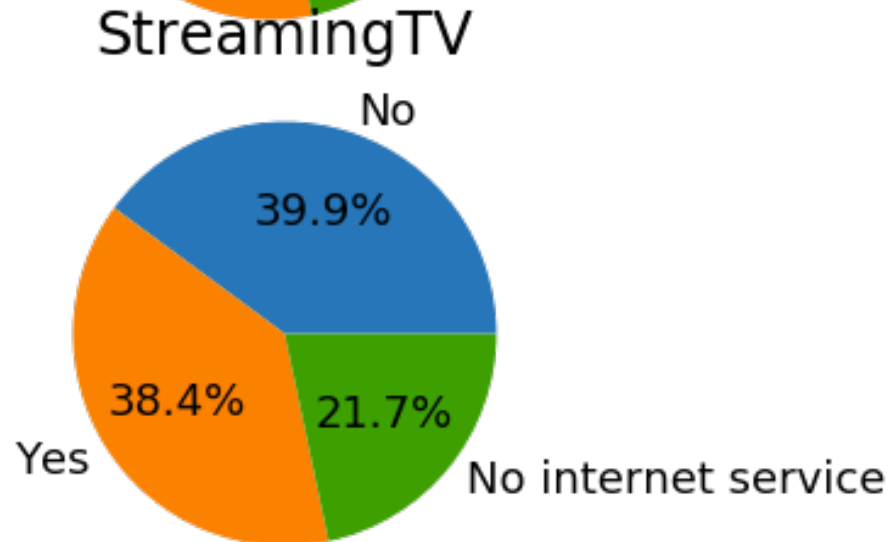
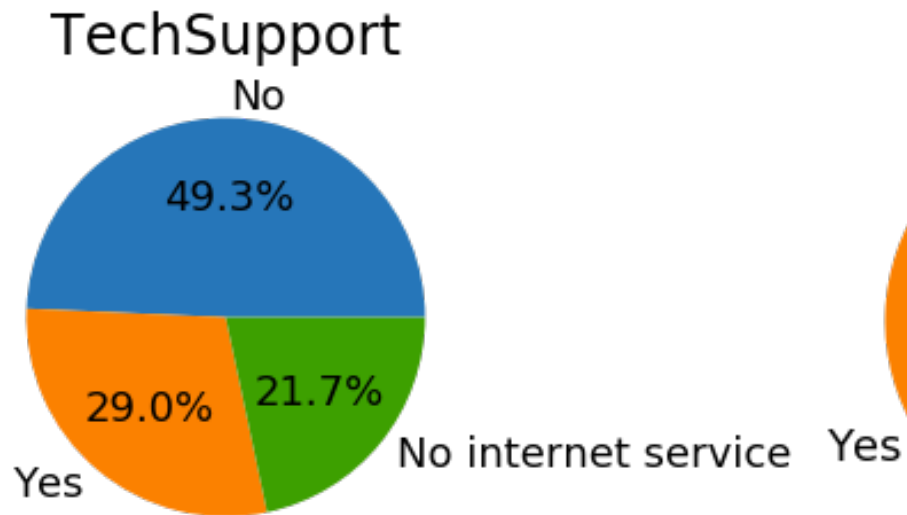
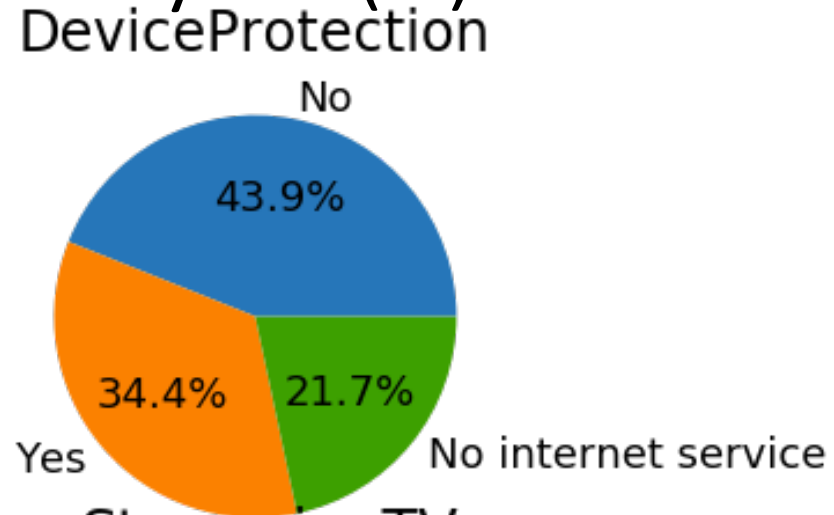
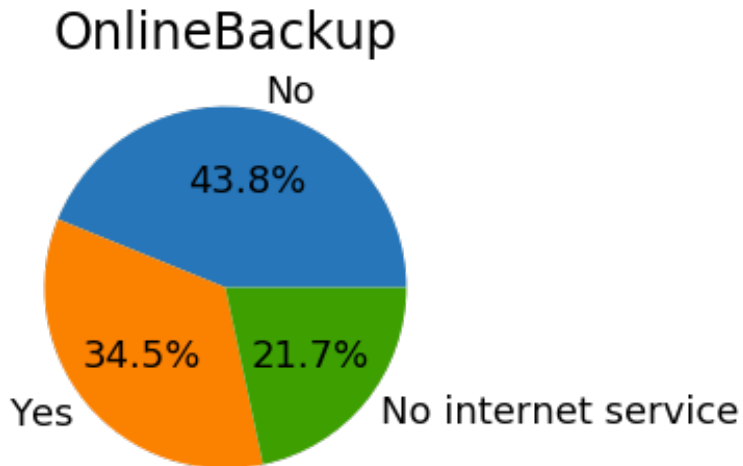
Partner



consumer behavior analysis (1)

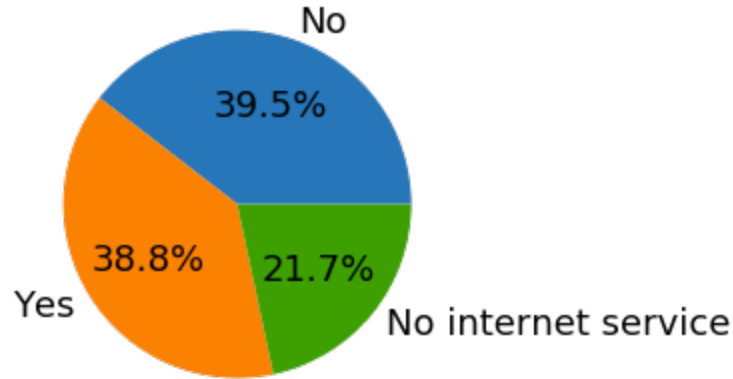


consumer behavior analysis (2)

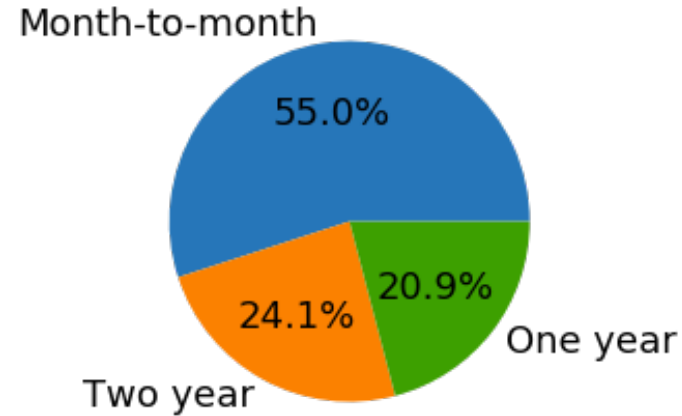


consumer behavior analysis (3)

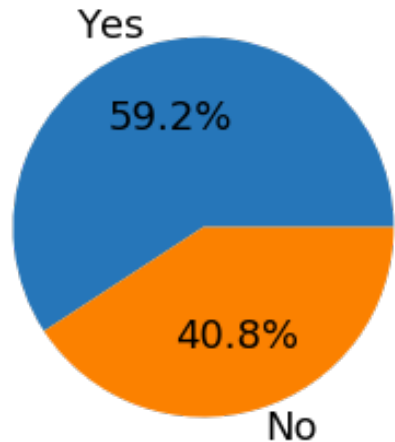
StreamingMovies



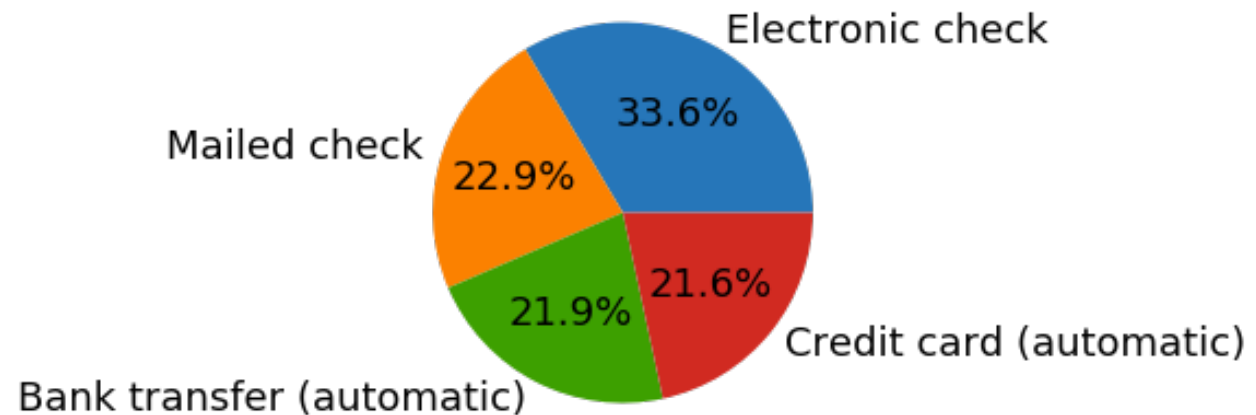
Contract



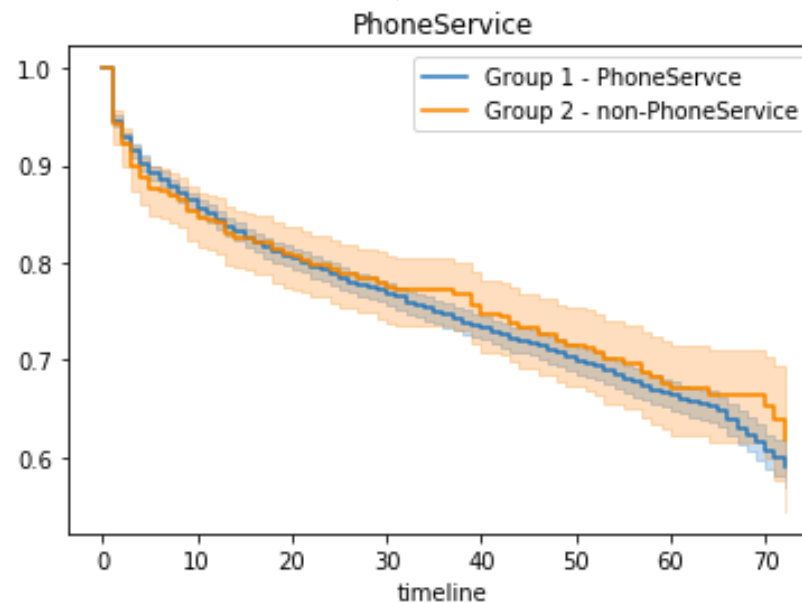
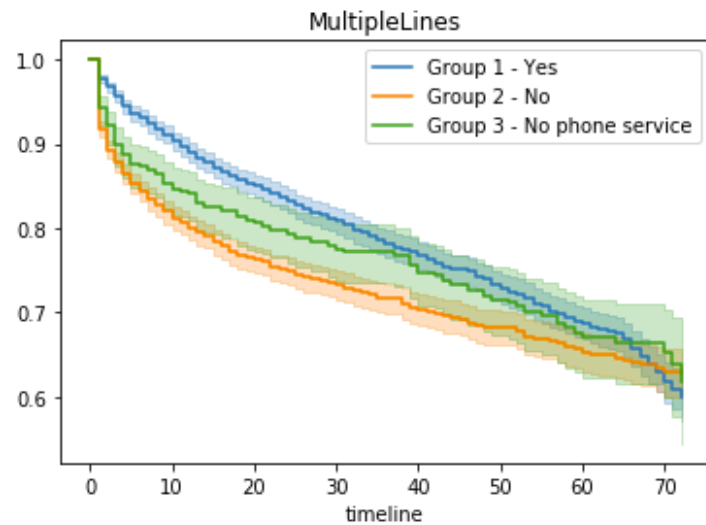
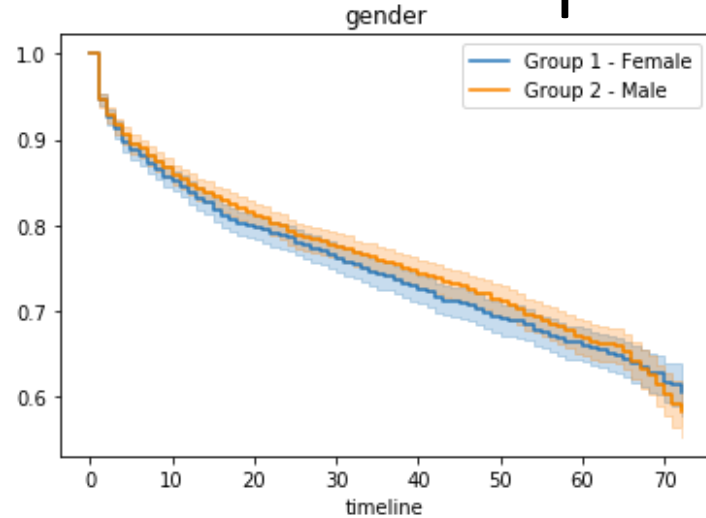
PaperlessBilling



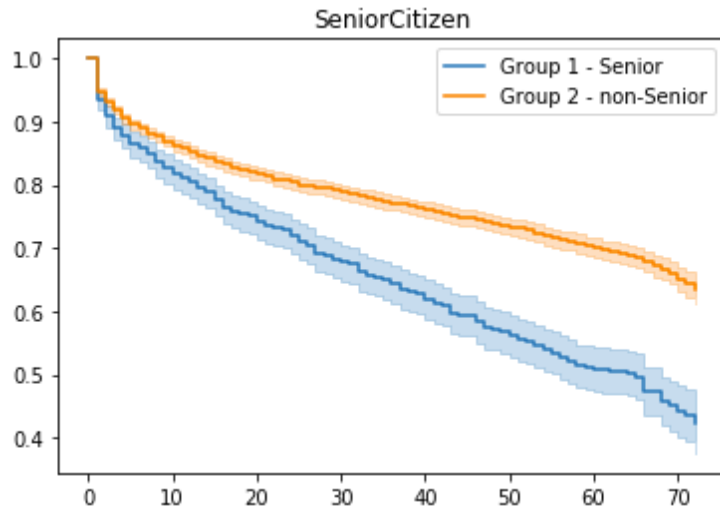
PaymentMethod



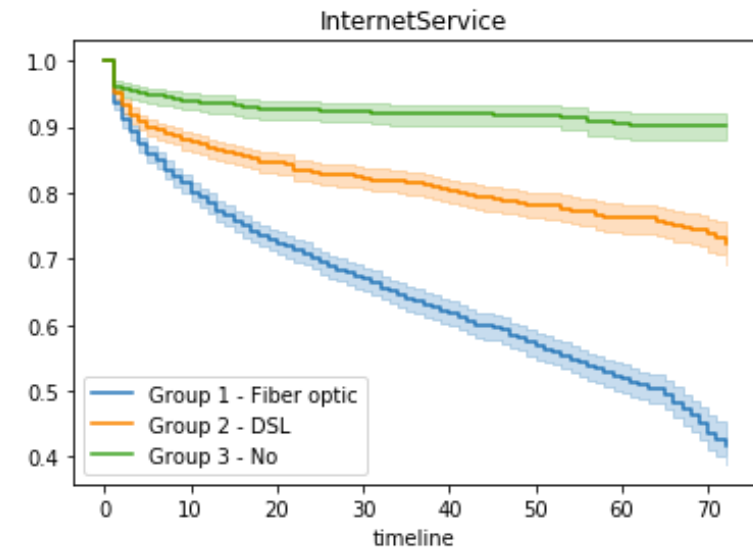
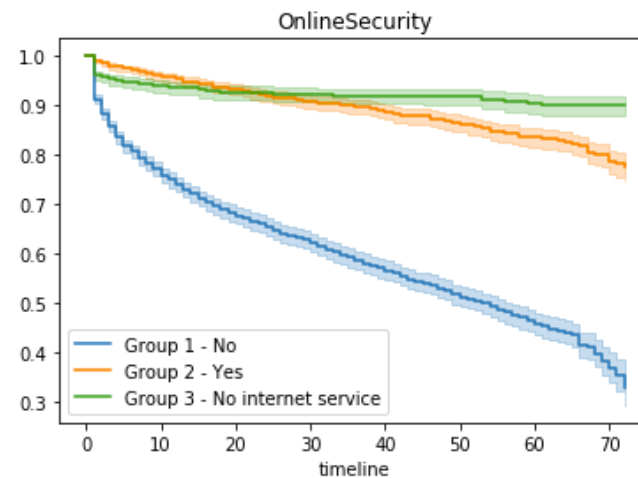
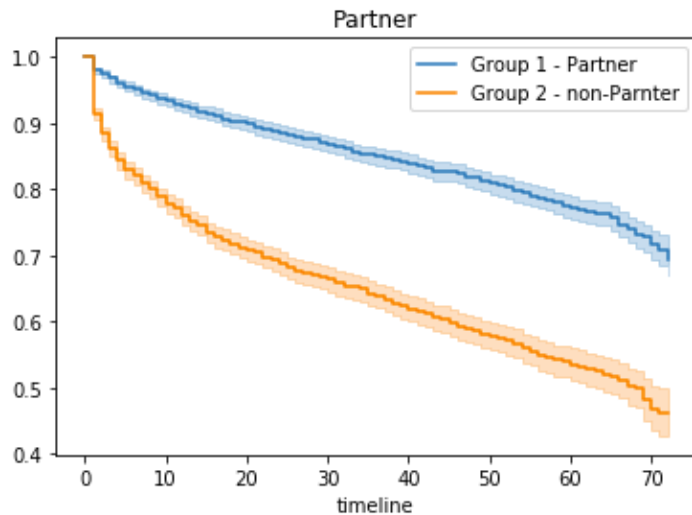
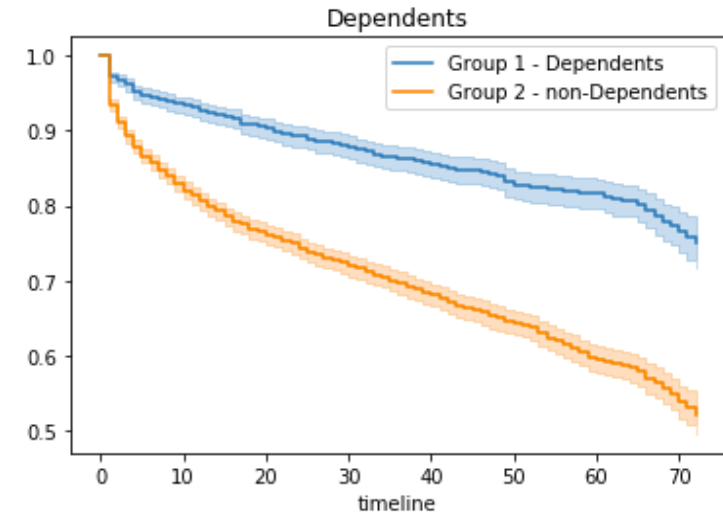
Feature Importance to Predict Churn (1)



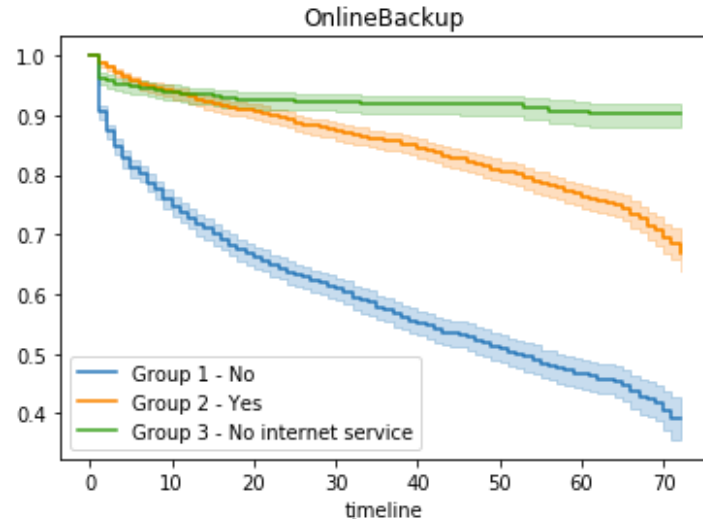
Feature Importance to Predict Churn (2)



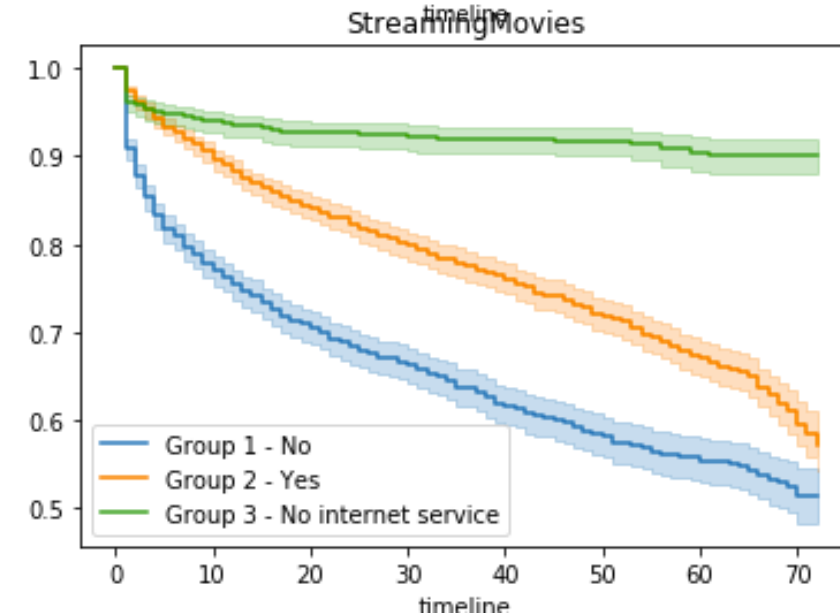
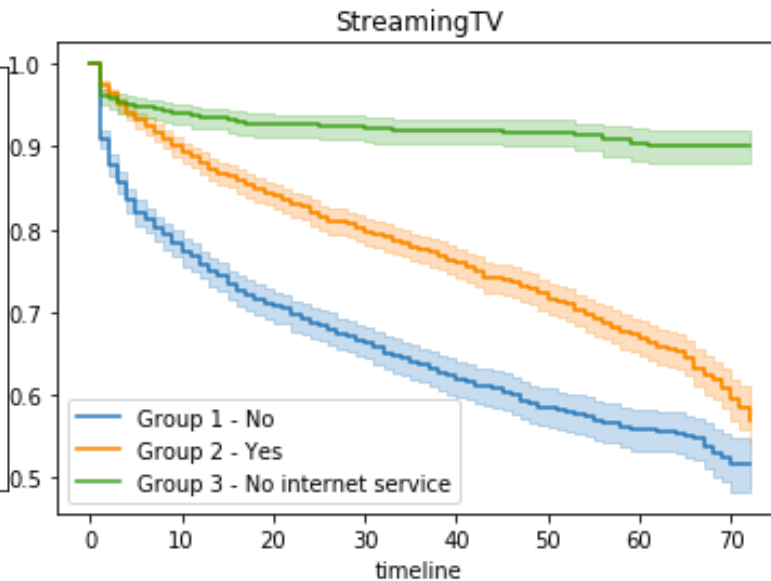
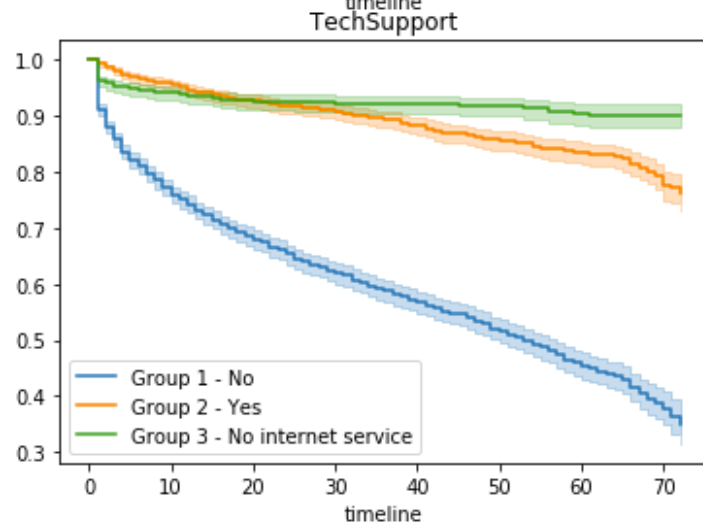
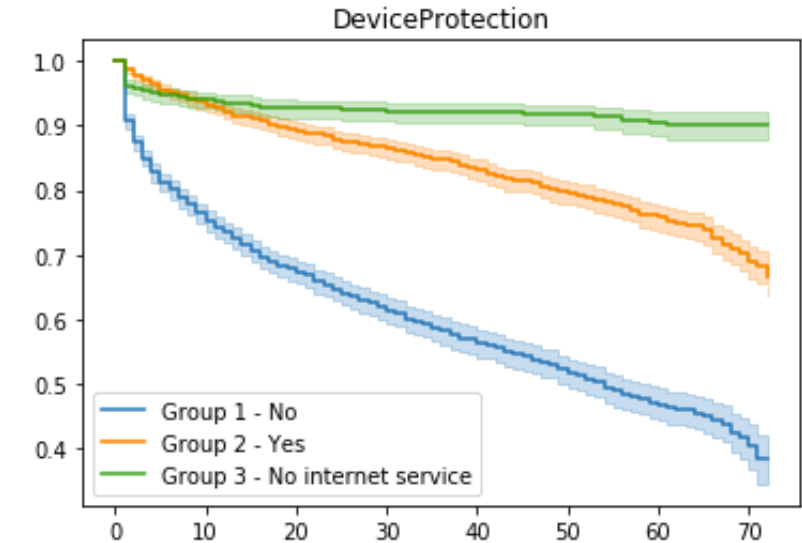
All these predictors are good predictors



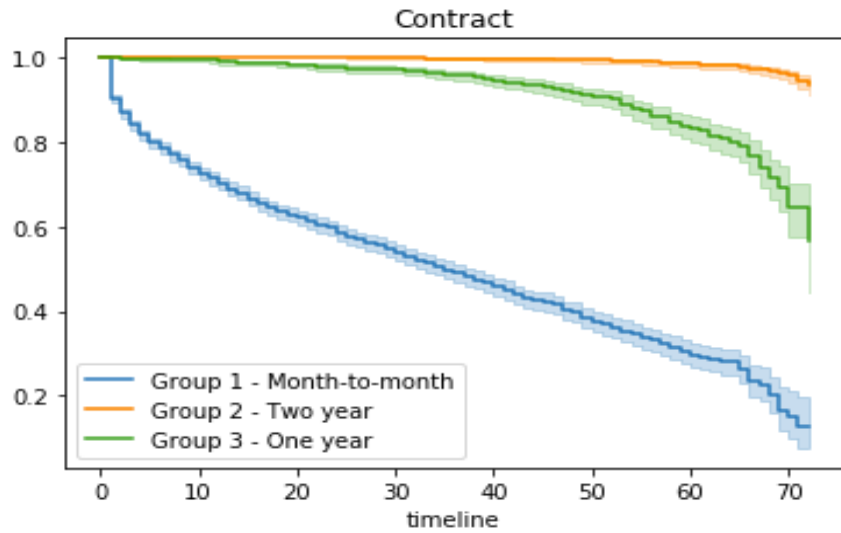
Feature Importance to Predict Churn (3)



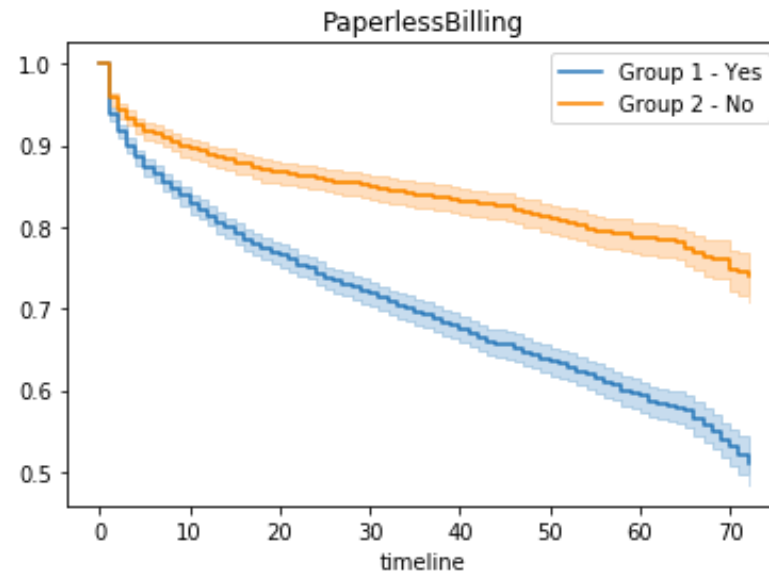
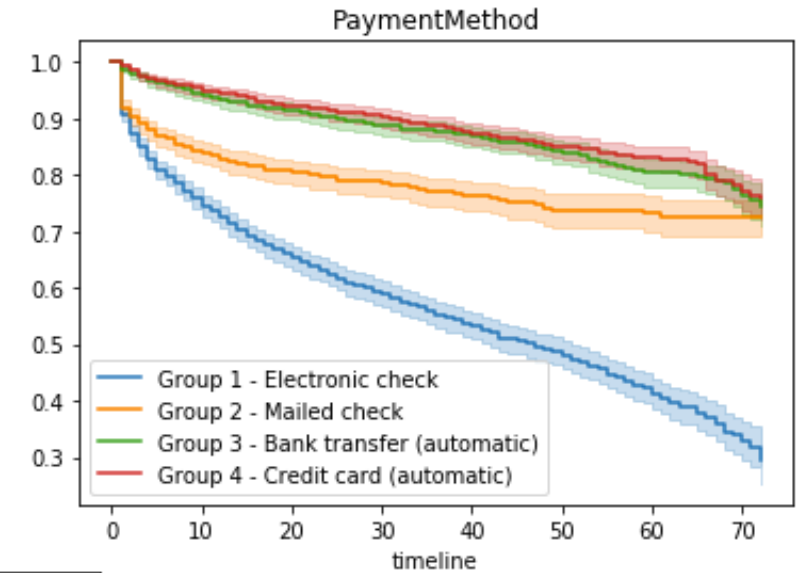
All these predictors are good predictors



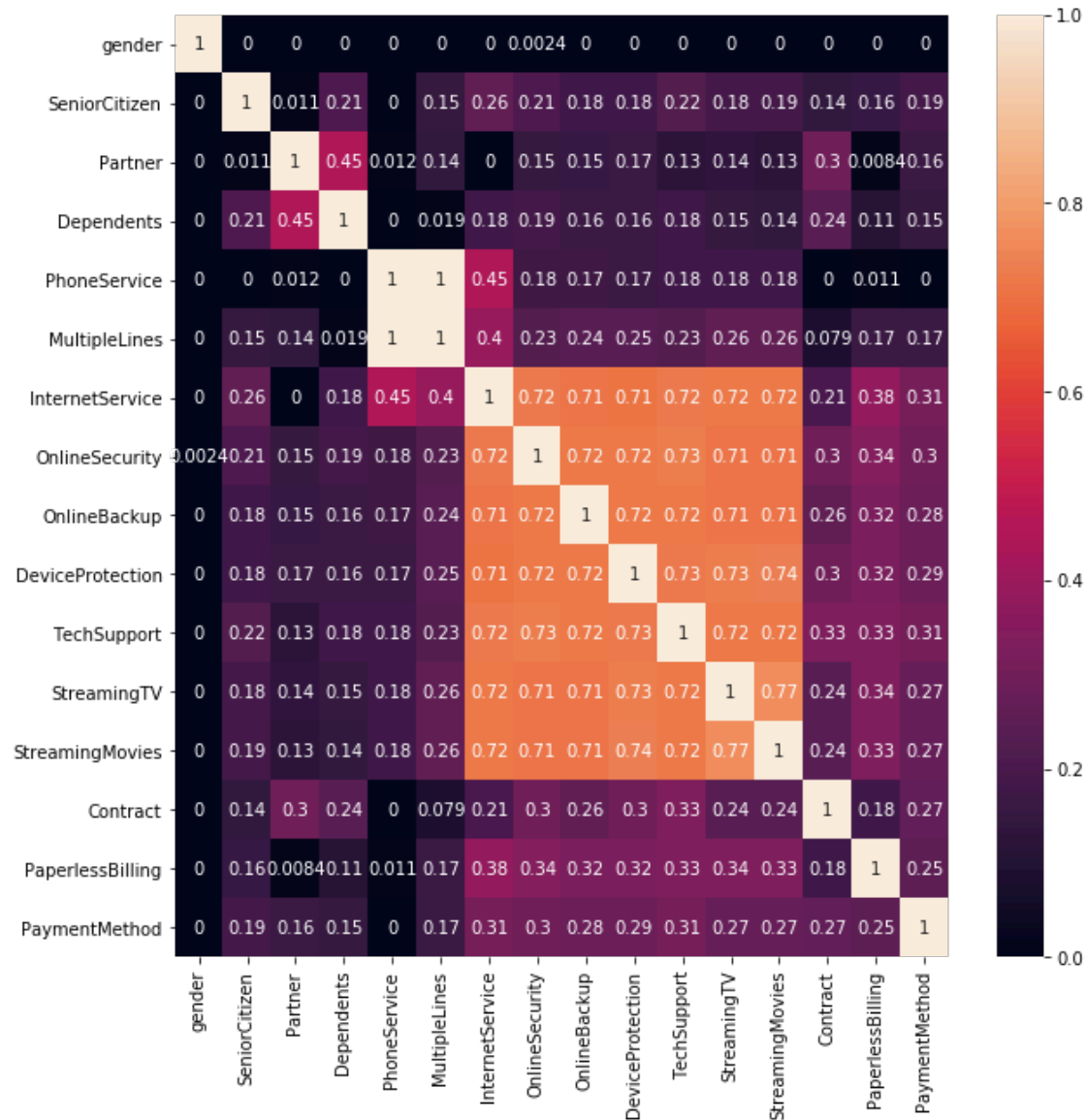
Feature Importance to Predict Churn (4)



All these predictors are good predictors



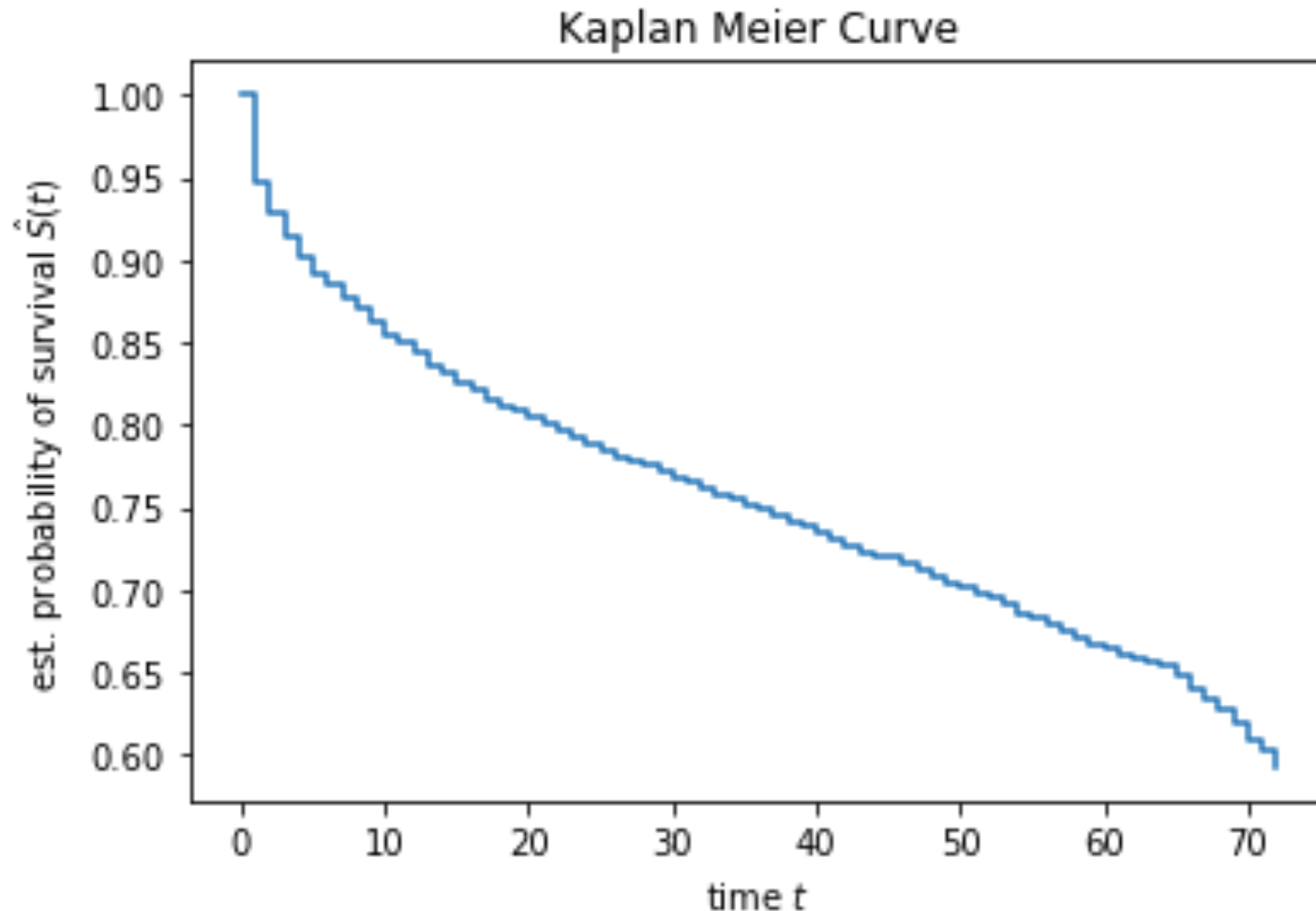
Cramers'V to check categorical correlation



Notes:

1. The majority of predictors is categorical;
2. Use Cramer's V method to convert the categorical correlations within the range [0.1]
3. InternetService is relatively highly related with other features: OnlineSecurity, OnlineBackup, DeviceProtection, TechSupport, StreamingTV and StreamingMovies.

Tenure-Churn Kaplan Meier Curve



Highlight:

- The trend of survival function shows that during the first three months of tenure, the event of churn occurred rapidly;
- Then the number of survivals declined gradually;
- In the late stage of the month 65 and month 72, a significant number of customers got churned.

Cox Proportional Hazard model performance

- concordance_index_censored

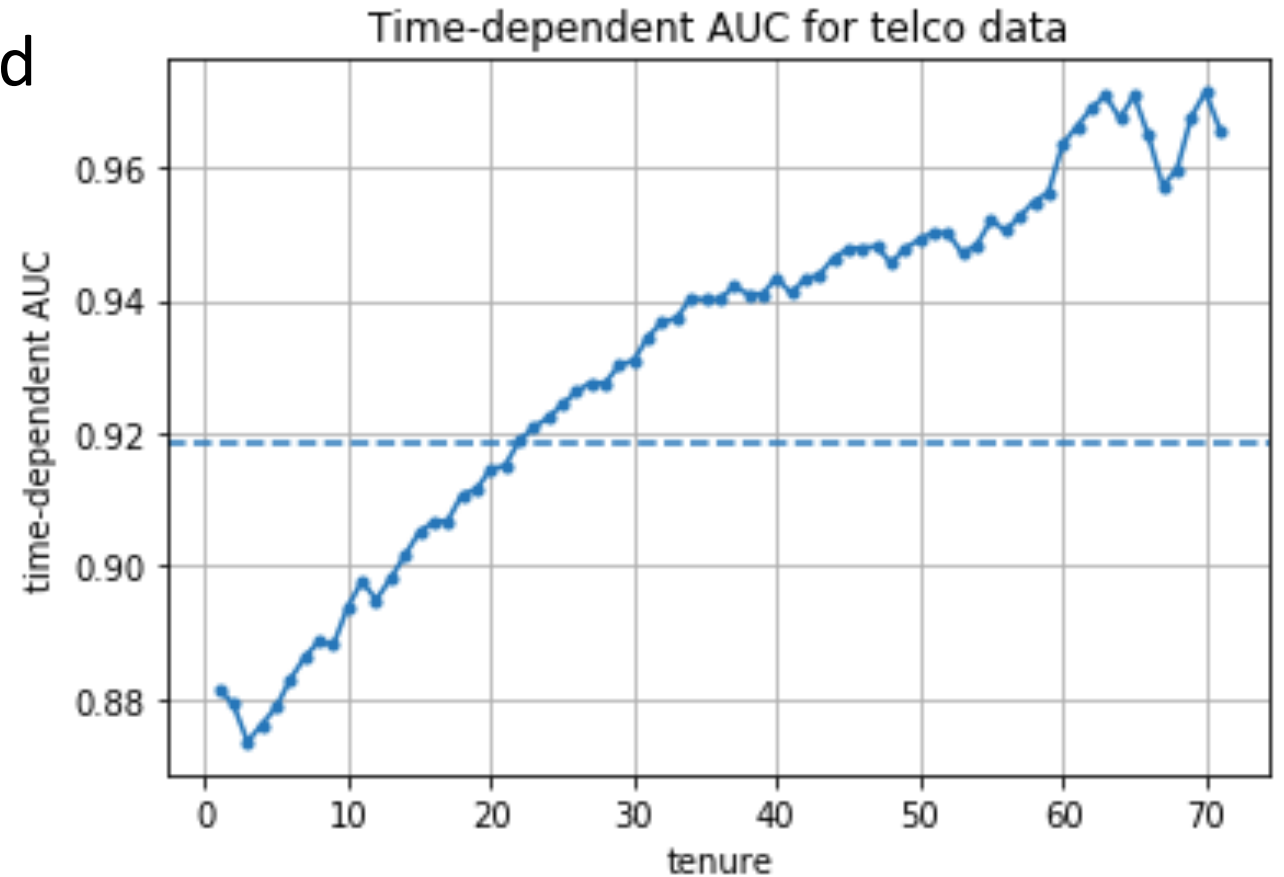
85.77%

- concordance_index_ipcw

85.92%

- time-dependent AUC

0.91



Summary

- Due to data leak, we discard the feature: TotalCharges, since there is no way to know the amount of total charges at the time of prediction.
- Six categorical variables are relatively highly related: OnlineSecurity, OnlineBackup, DeviceProtection, TechSupport, StreamingTV, StreamingMovies.
- For Gender, MultipleLine, and PhoneServices, no significant difference between survival curves for each variables' subgroups.
- Substantial drop of survival rate during the beginning and end period of customers' life cycle. More marketing effort should be made in these two phase to achieve high customer retention.