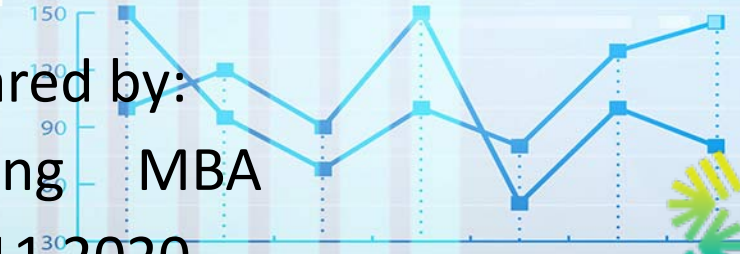
Survival Analysis on Telco Customer Churn



Prepared by:

Steve Song MBA

May 11,2020

Springboard

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 dataset

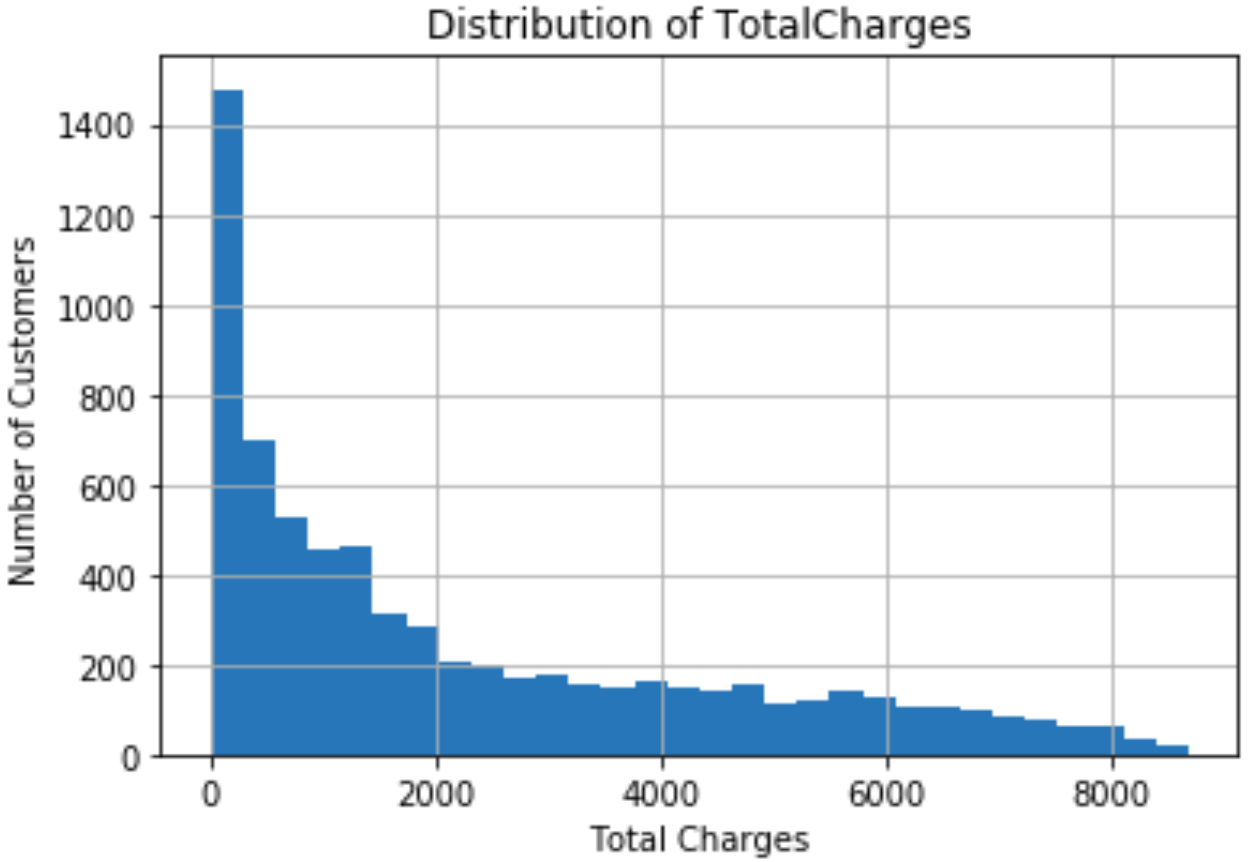
- size of data : 7,043 unique customers

- 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

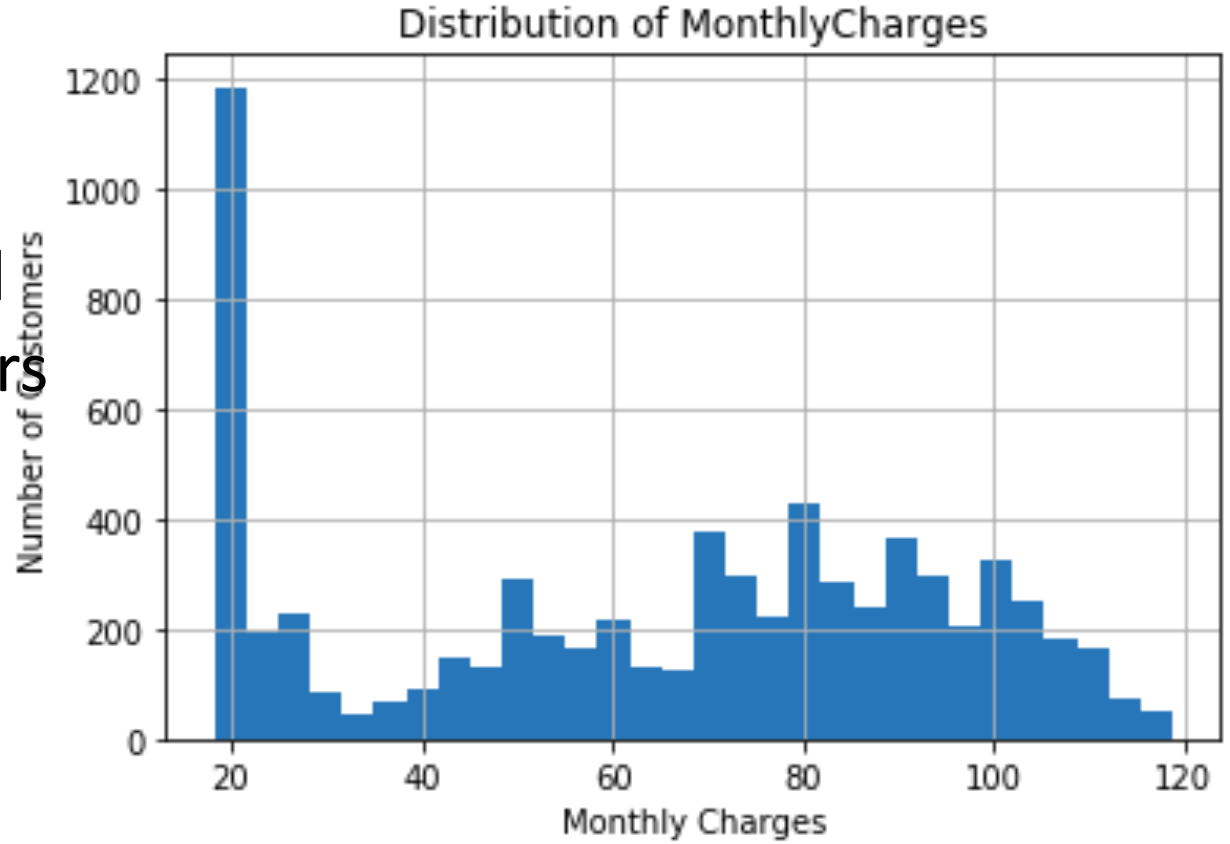
Distribution of Total Charges

- 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



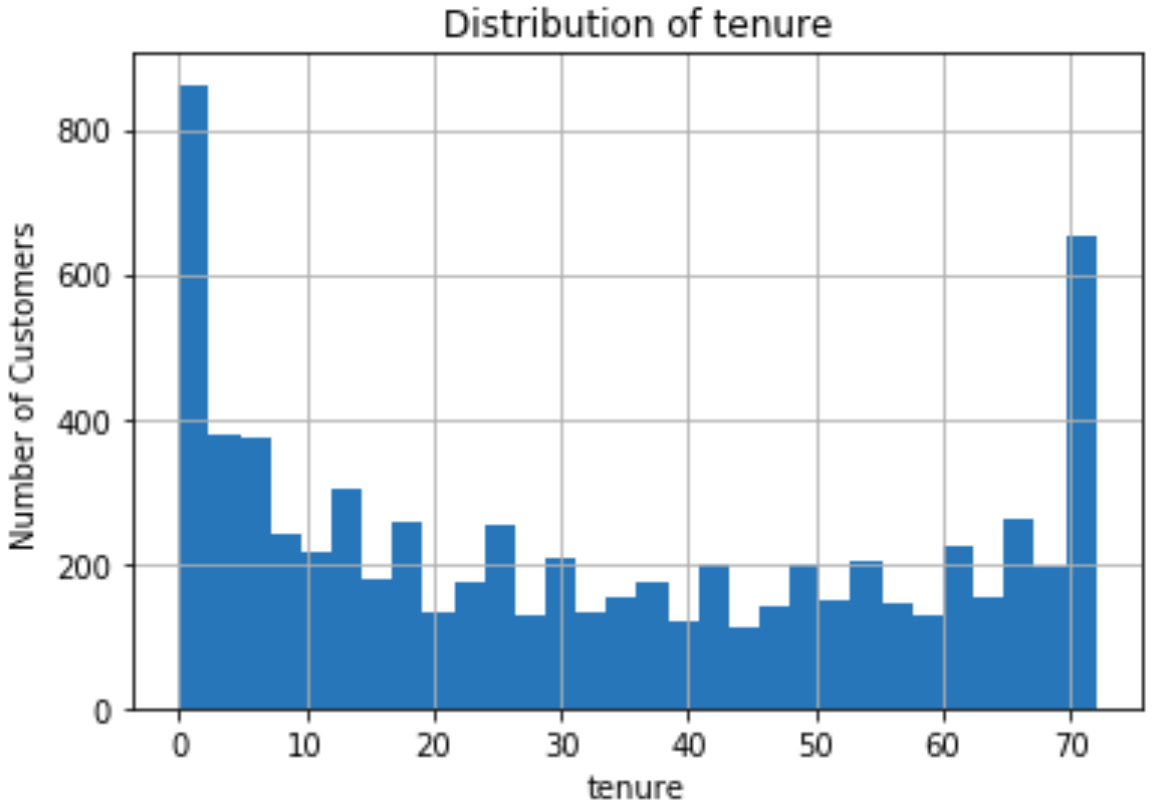
Distribution of Monthly Charges

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

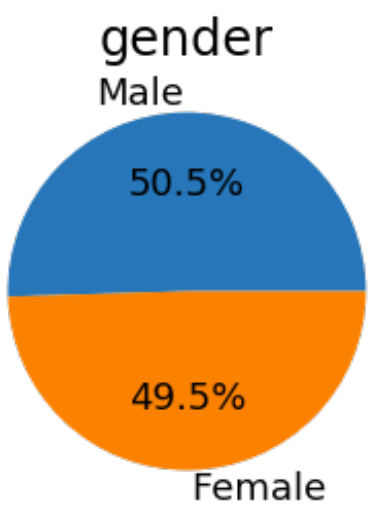


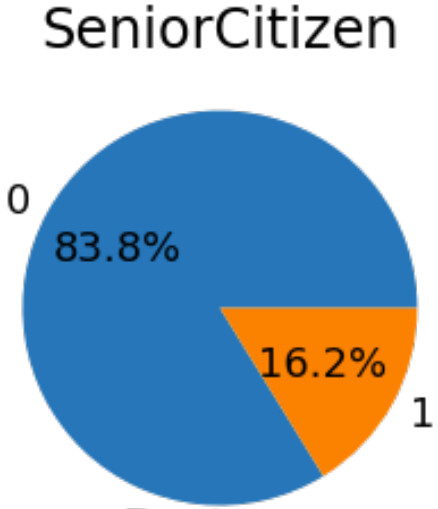
Distribution of tenure

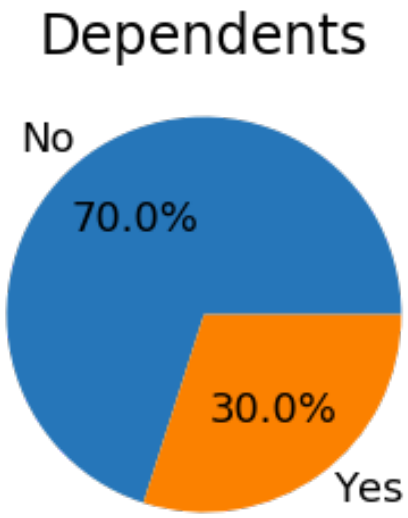
- relatively stable trend between 10 months and 60 months.
- significant number of customers stayed with business for 10 months (new customers) or above 60 months (loyal customers).

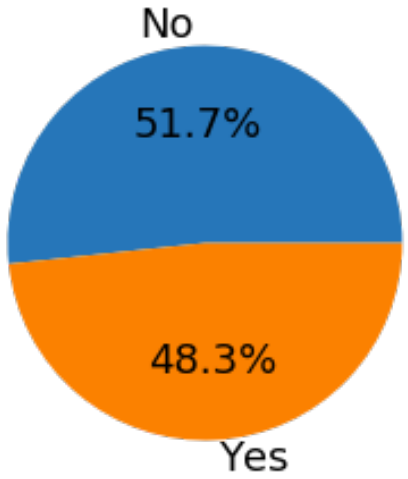


Customer attributes analysis

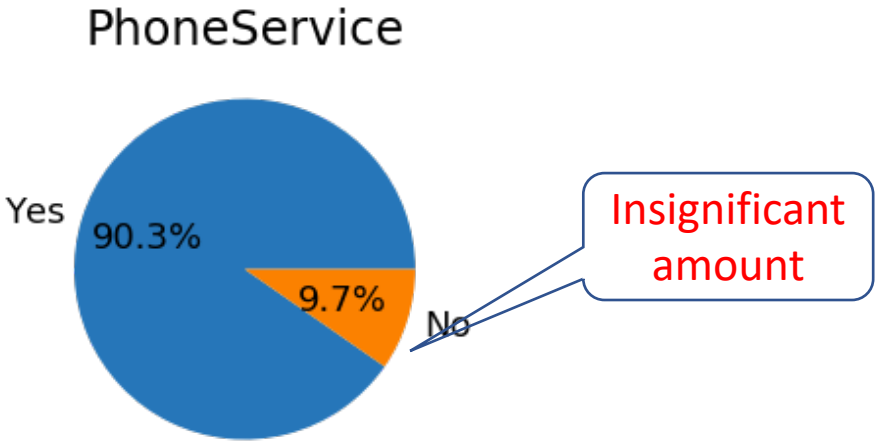


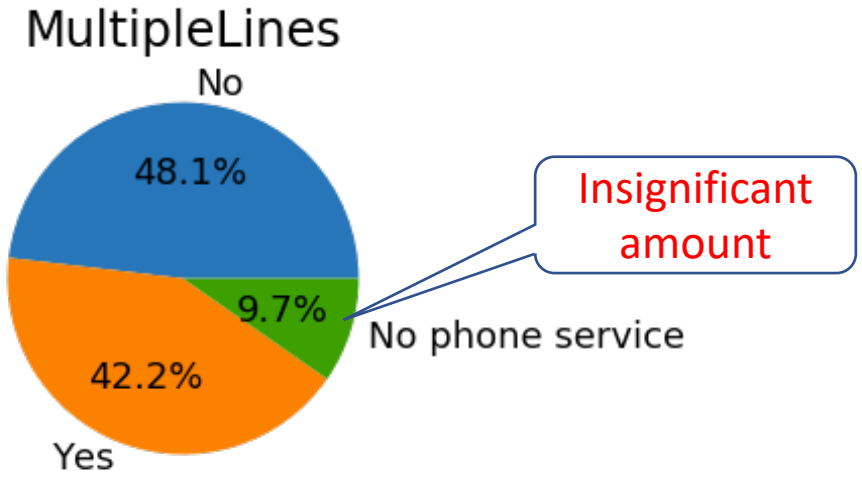


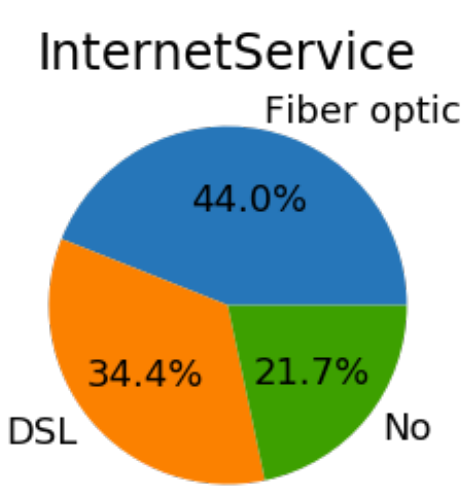


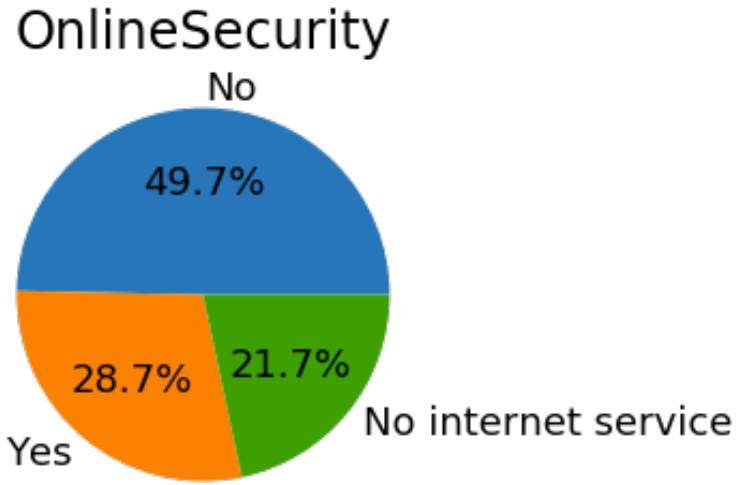


consumer behavior analysis (1)

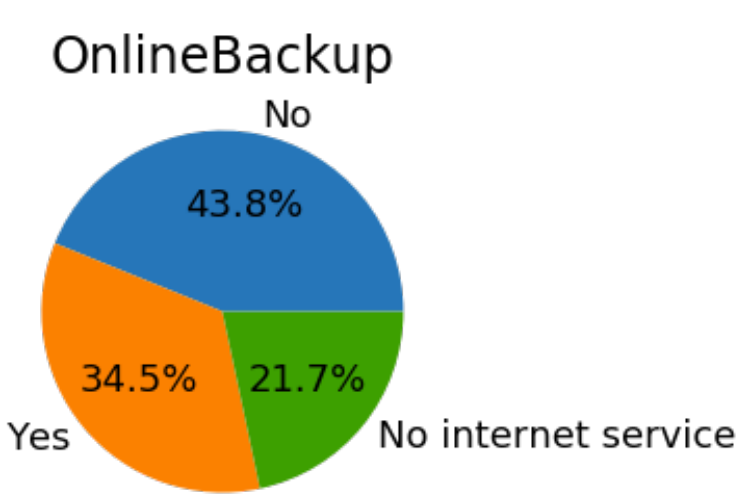


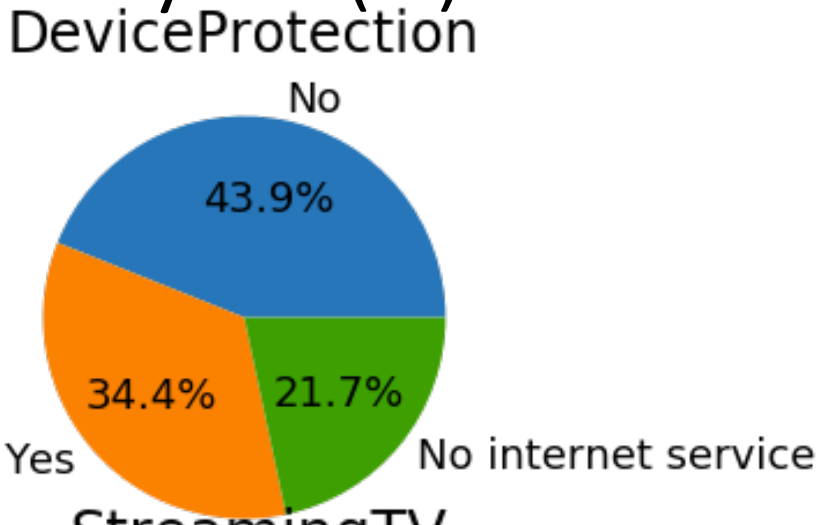


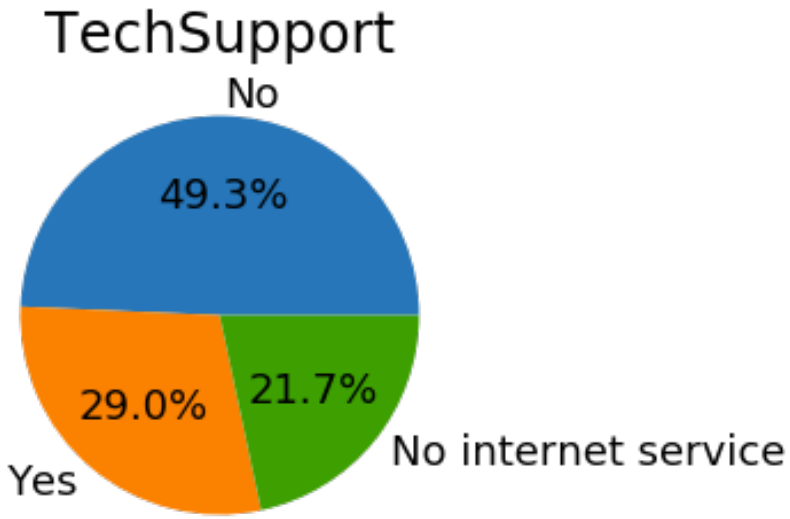


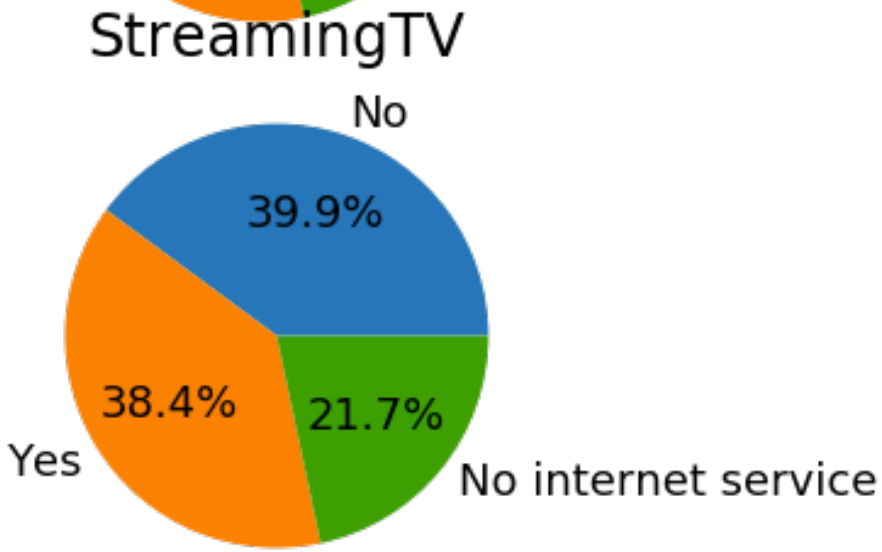


consumer behavior analysis (2)

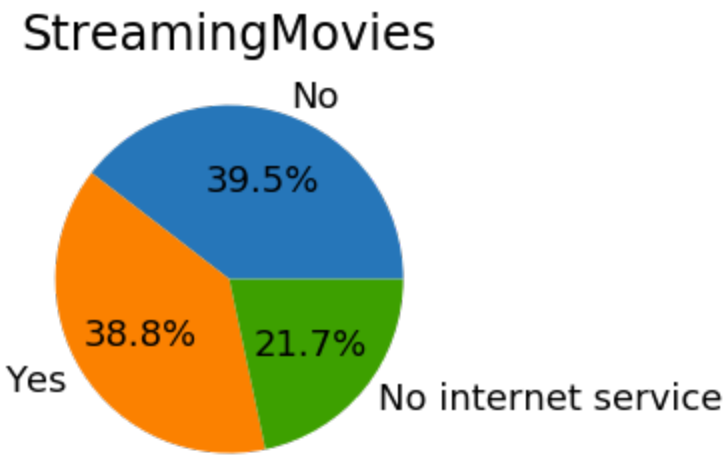


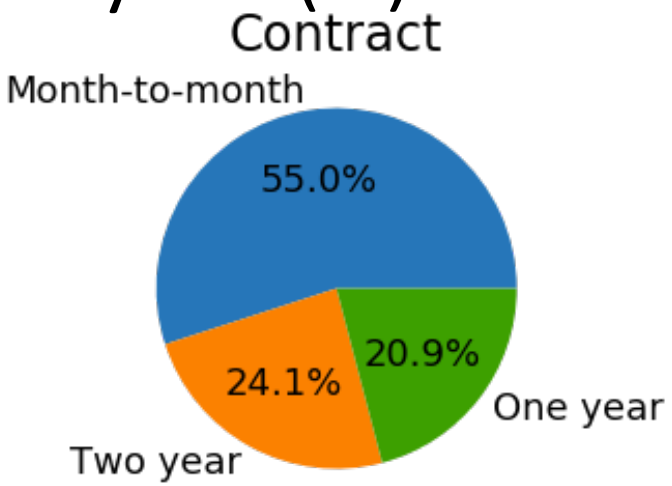


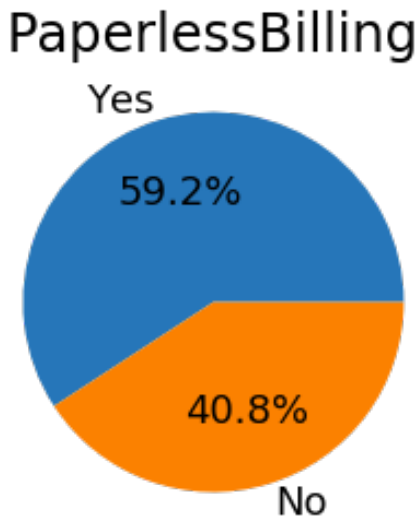


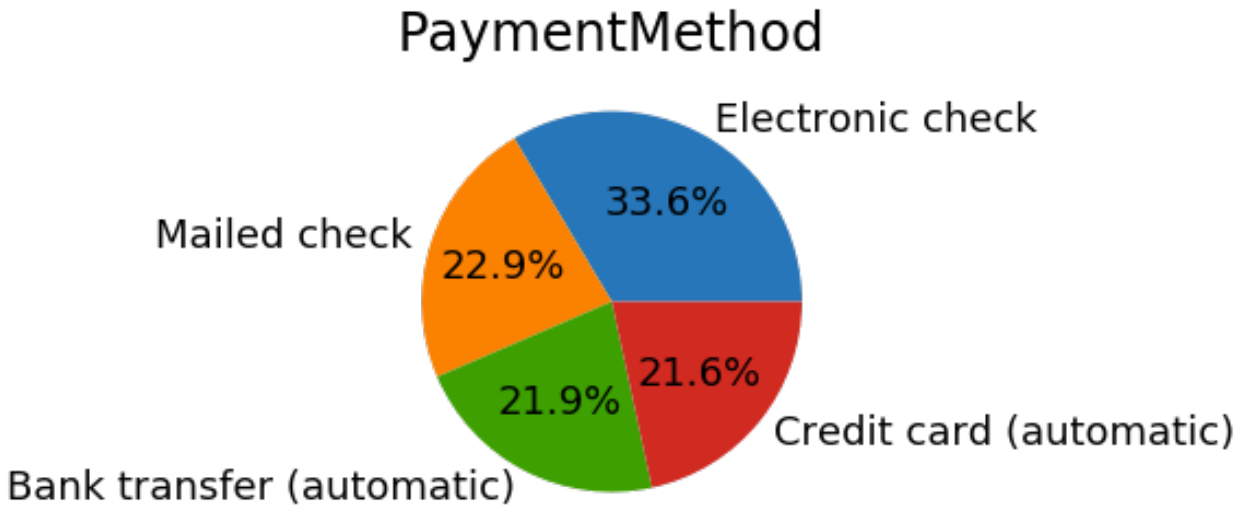


consumer behavior analysis (3)

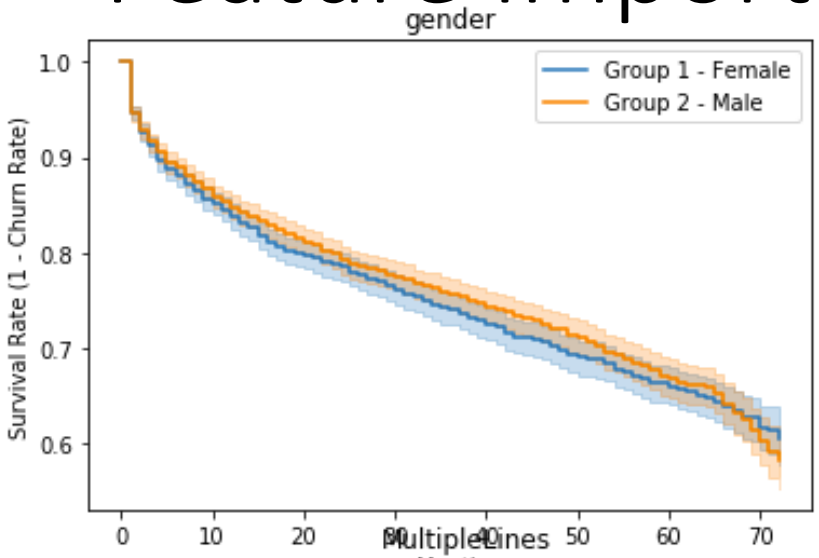


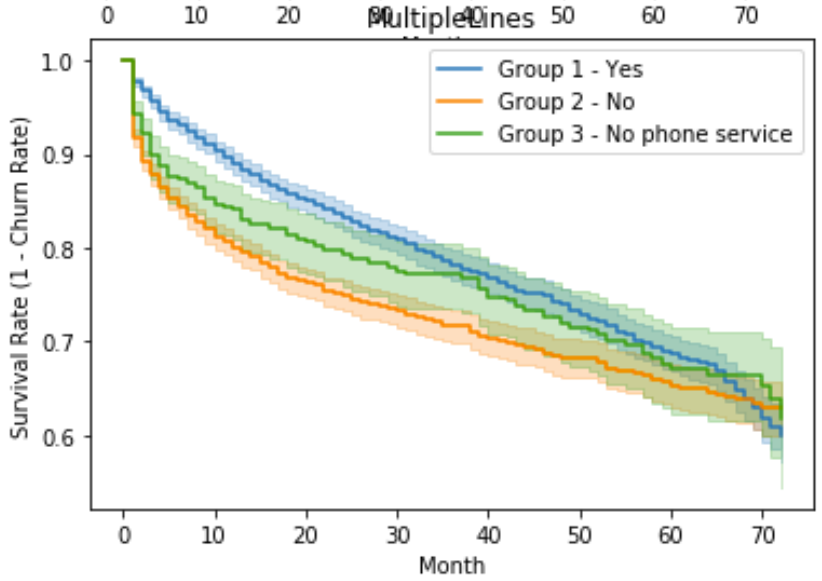




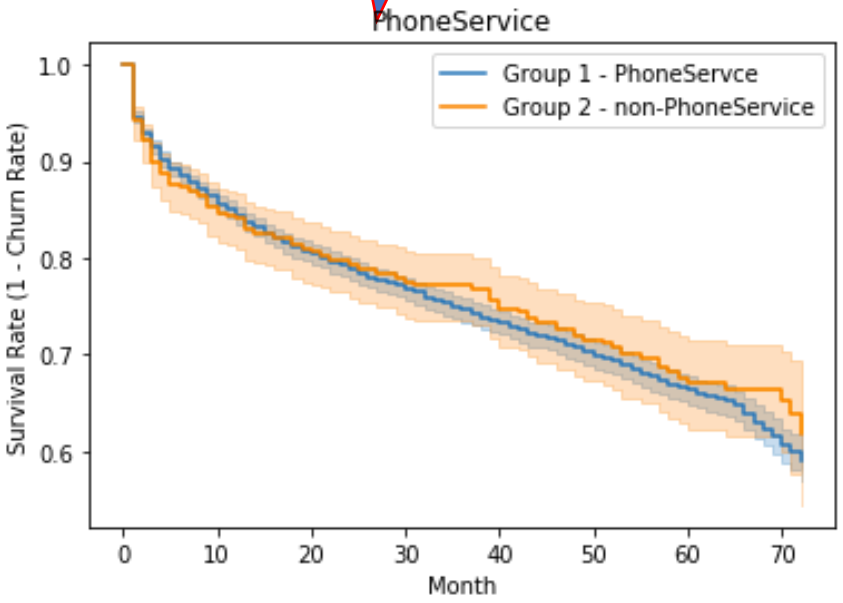


Feature Importance to Predict Churn (1)

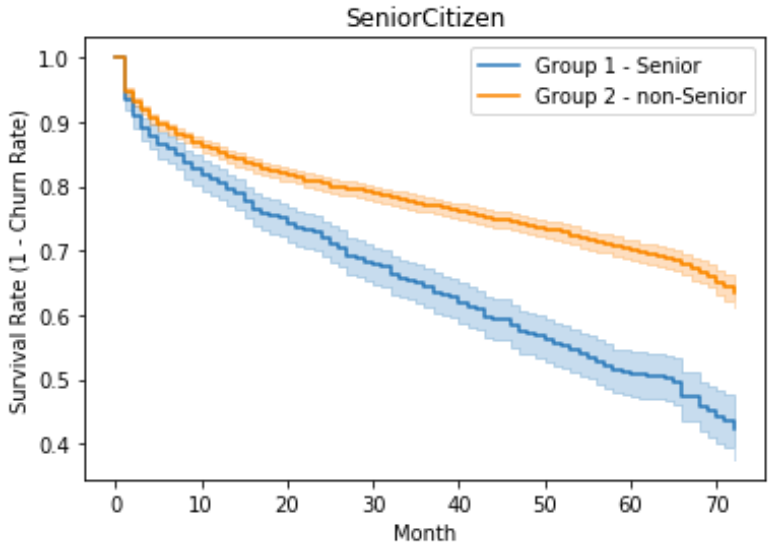




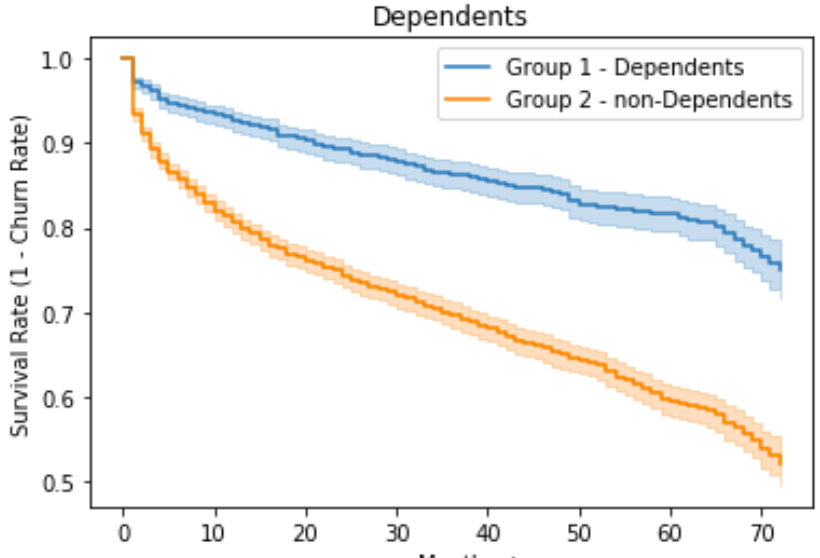
All these predictors are not good predictors

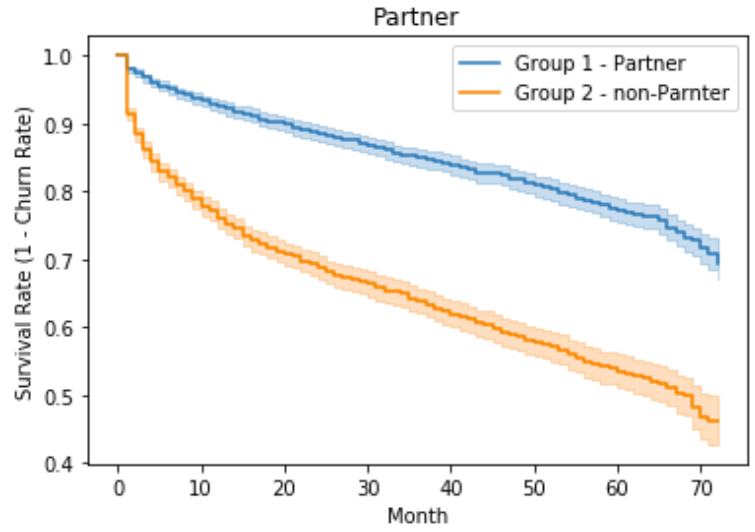


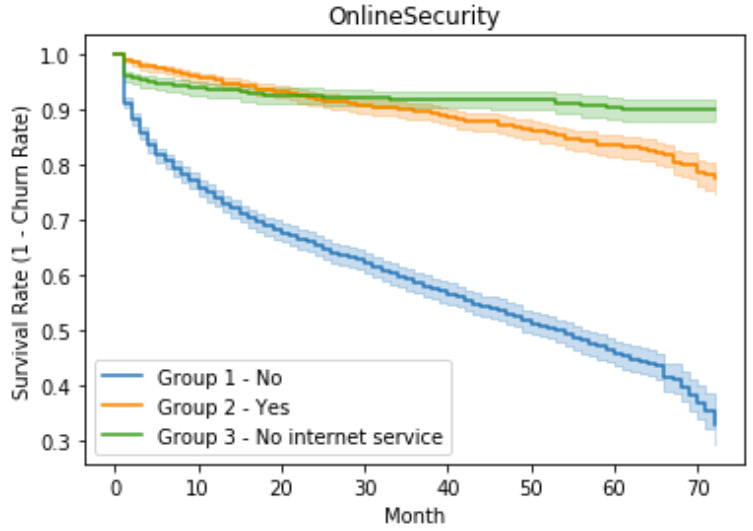
Feature Importance to Predict Churn (2)

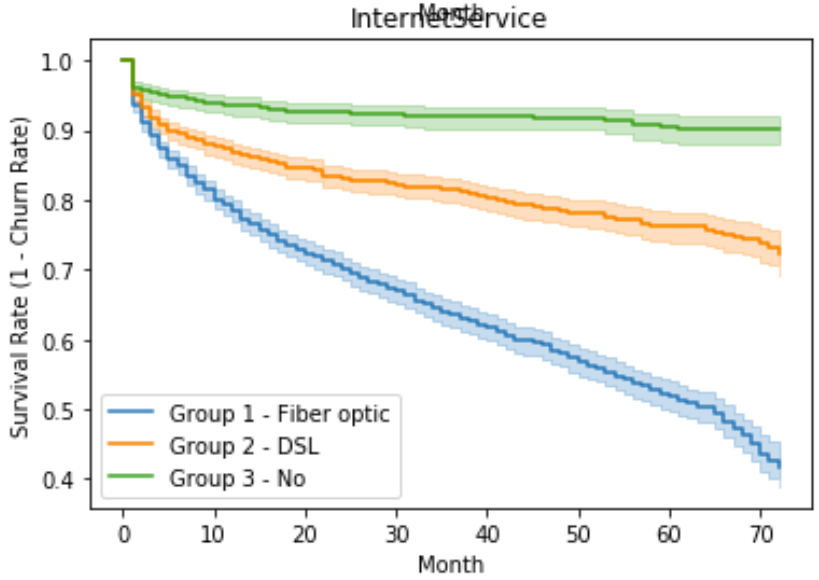


All these predictors are good predictors

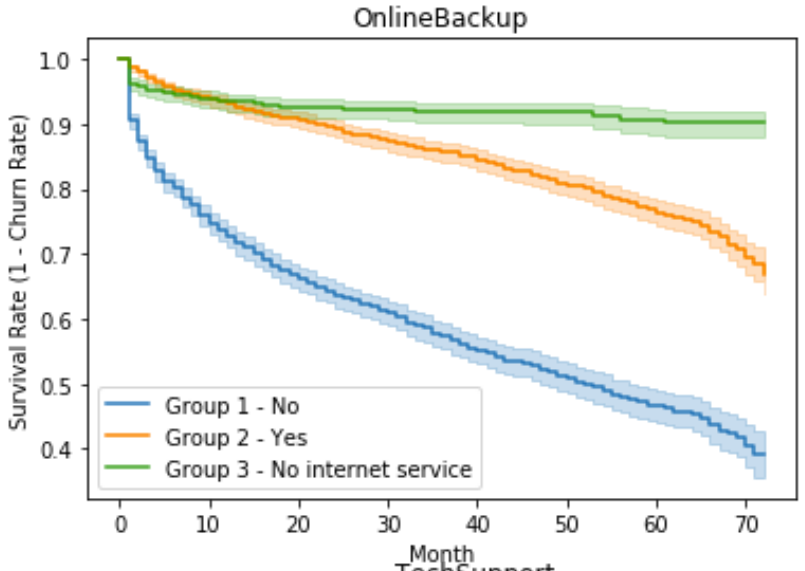




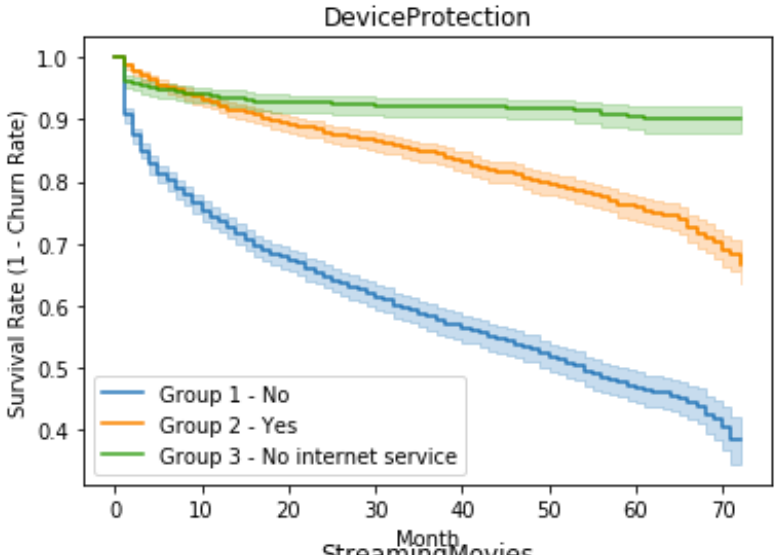


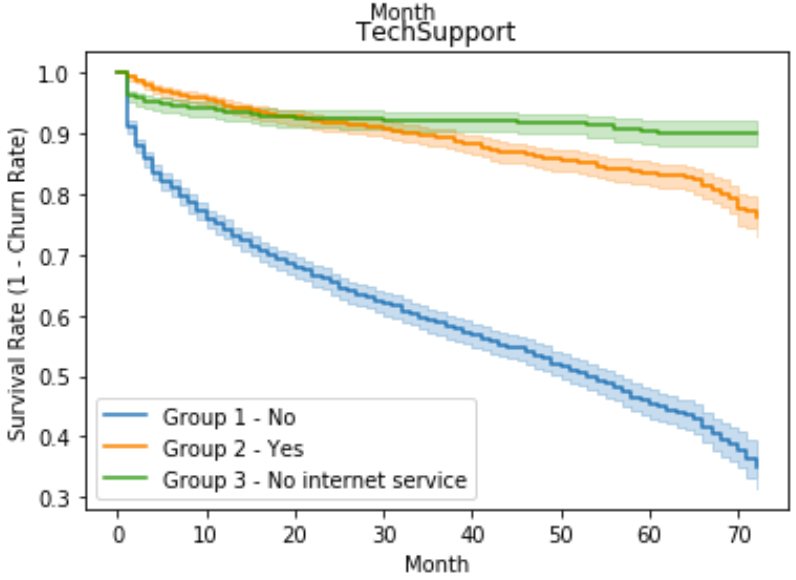


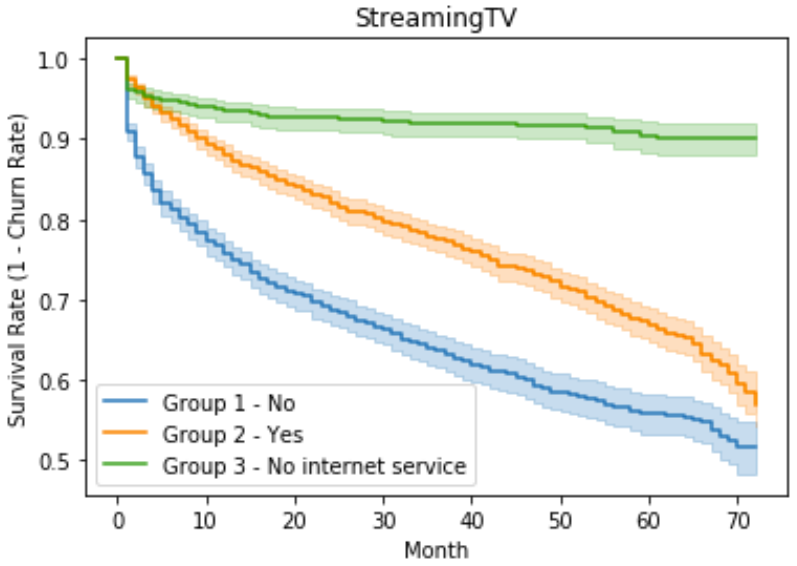
Feature Importance to Predict Churn (3)

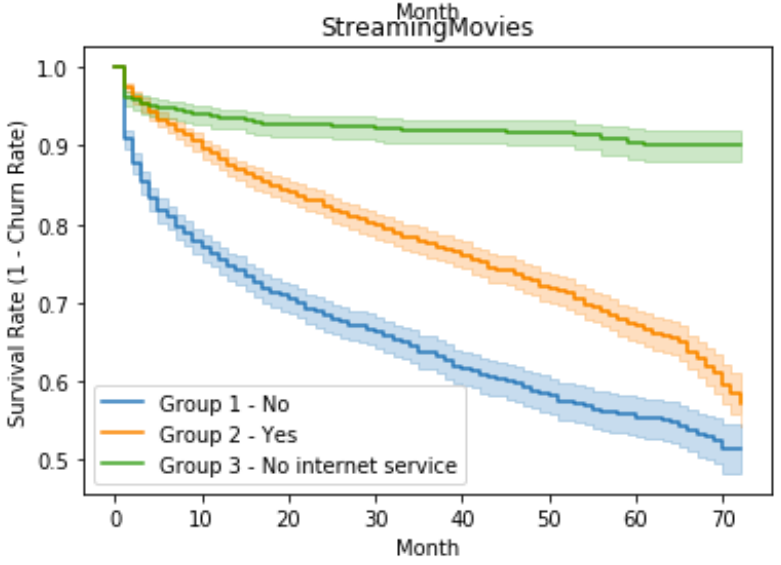


All these predictors are good predictors

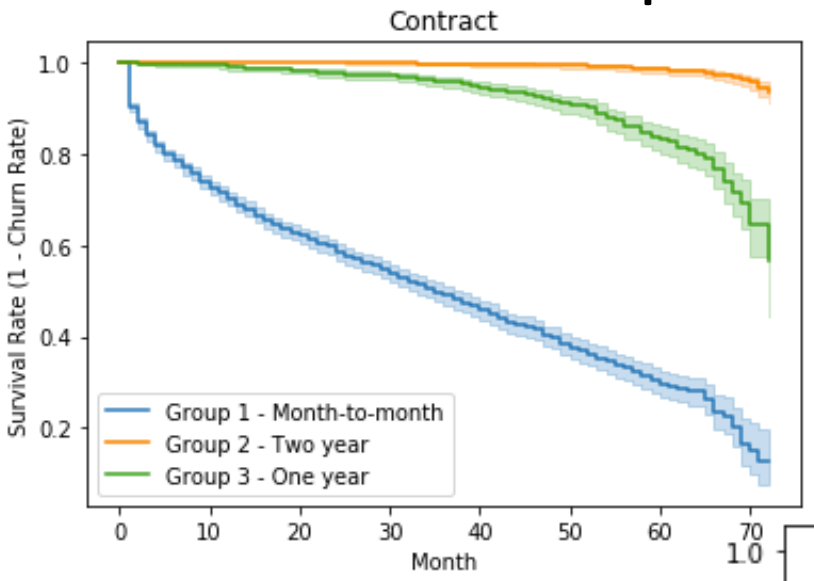




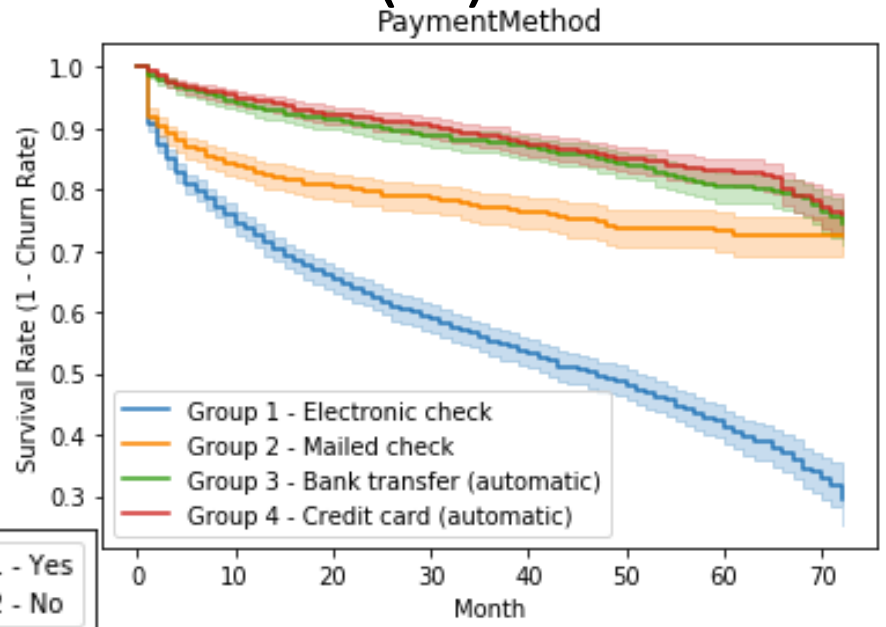


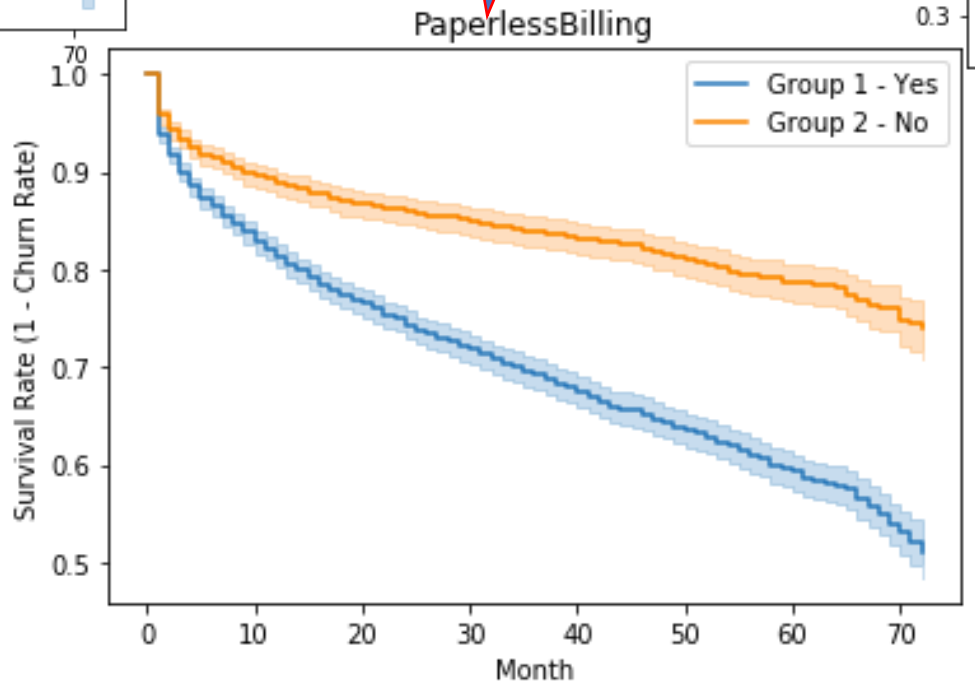


Feature Importance to Predict Churn (4)

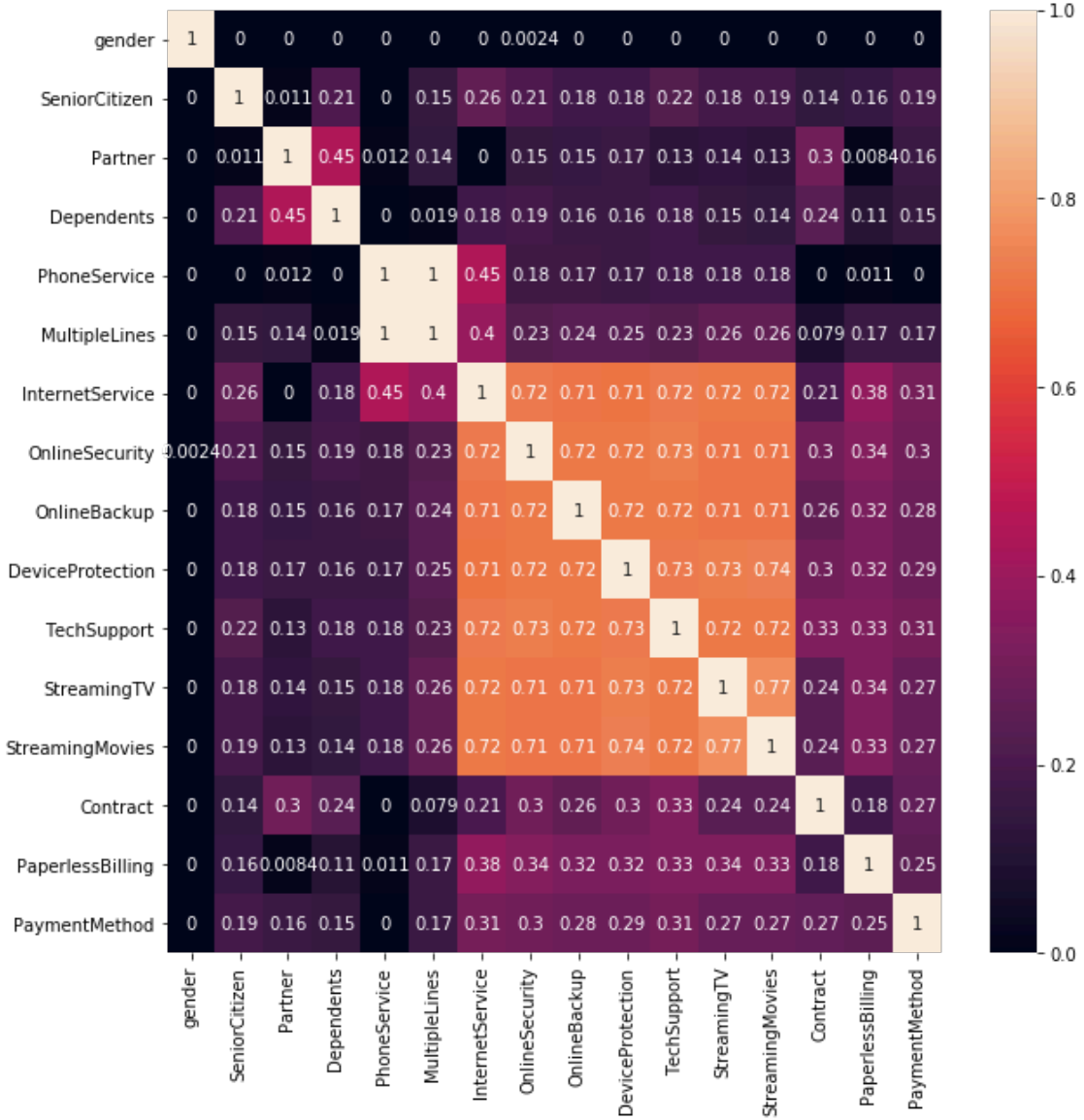


All these predictors are good predictors





Cramers'V to check categorical correlation



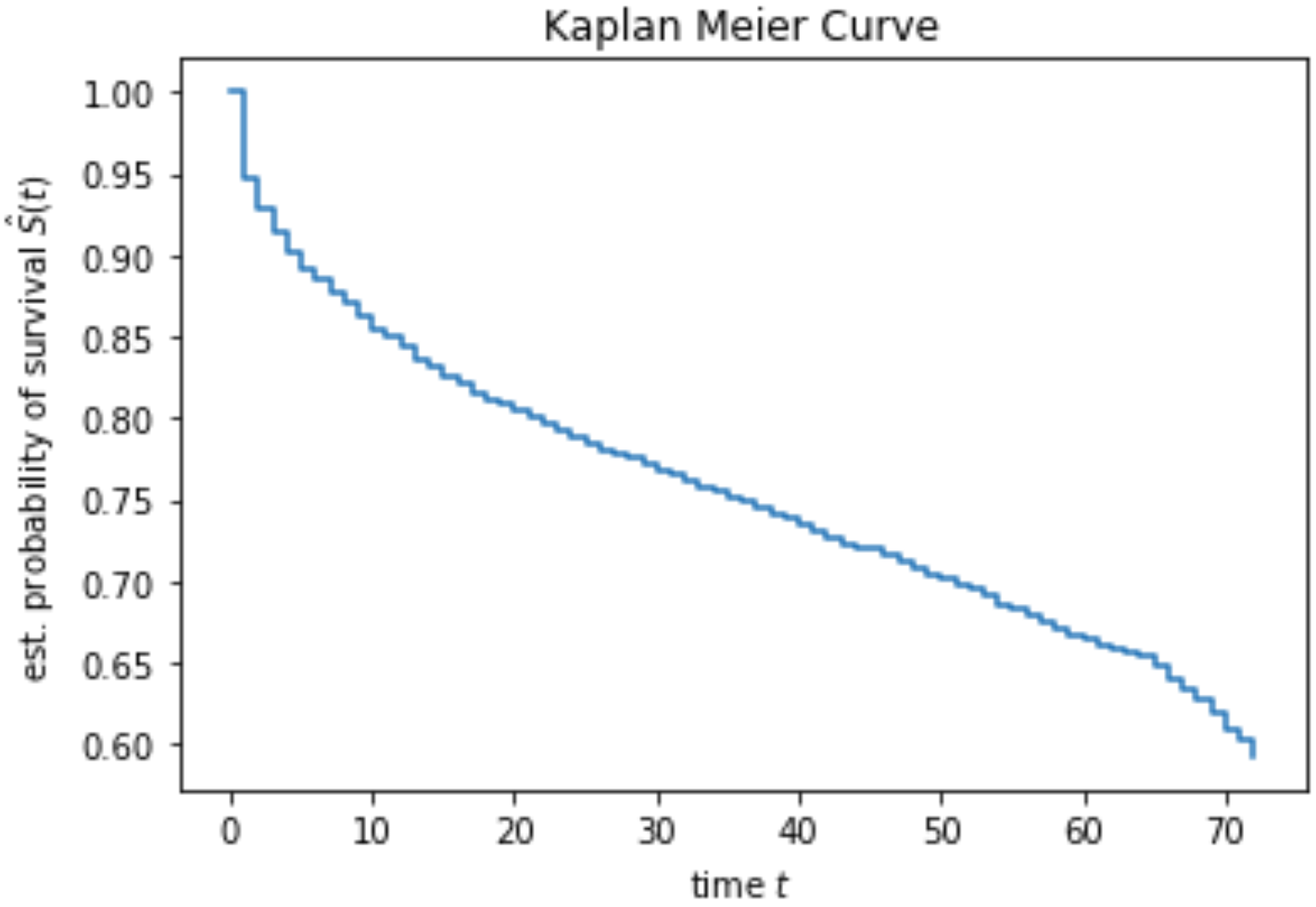
Notes:

1. The majority of predictors is categorical;

2. Use Cramer’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.

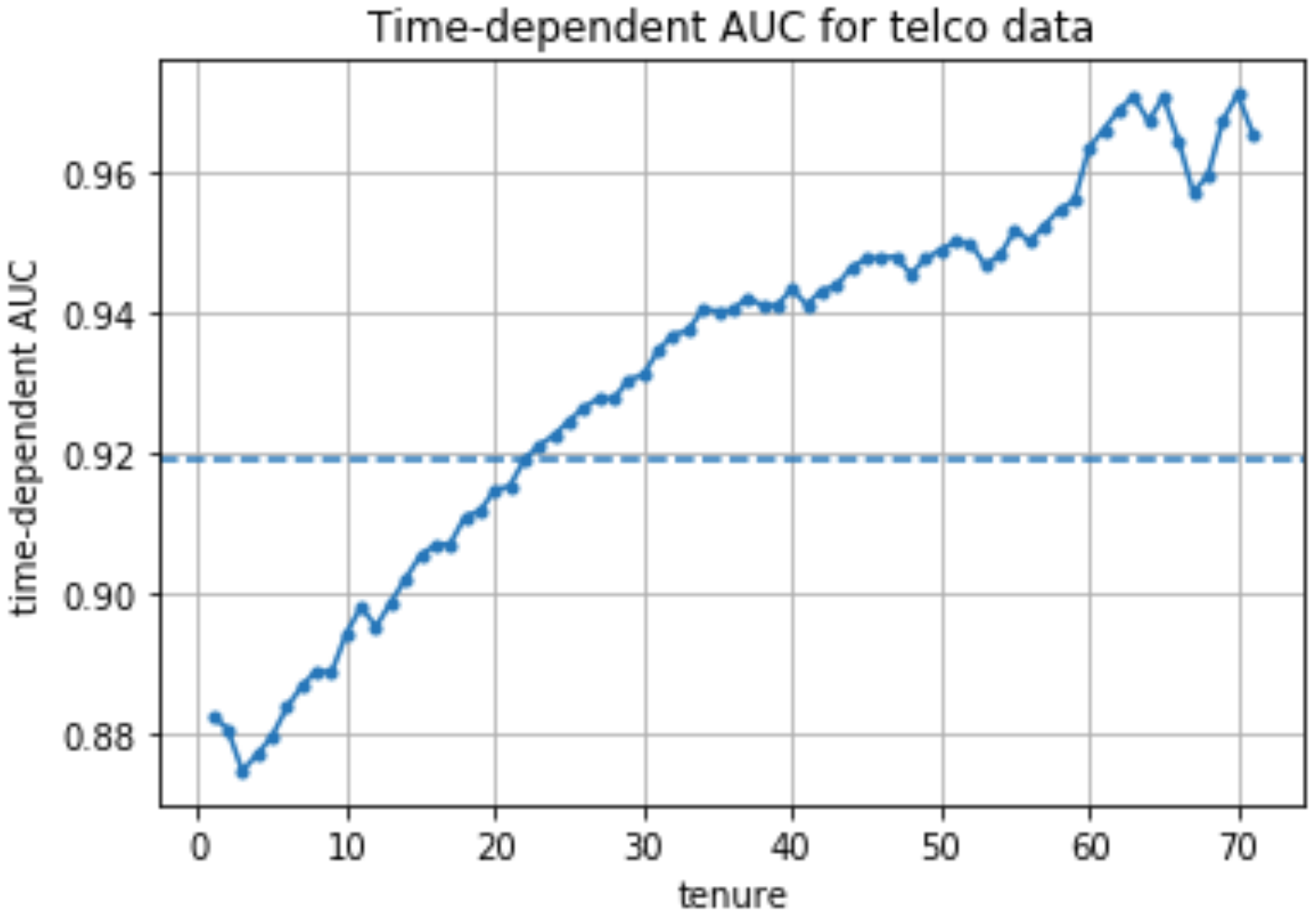
Tenure-Churn Kaplan Meier Curve



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 Hazards model



Performance:

concordance index censored: 0.8581

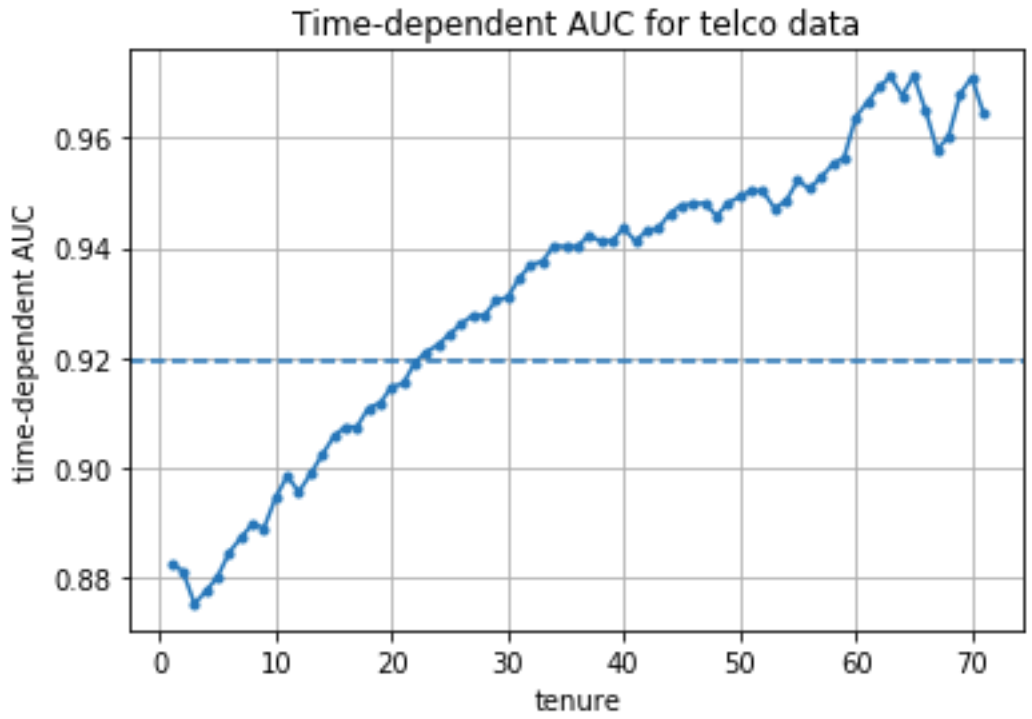
concordance index ipcw: 0.8593

time-dependent AUC: 0.9192

Feature Selection

L1-based feature selection using CoxnetSurvivalAnalysis

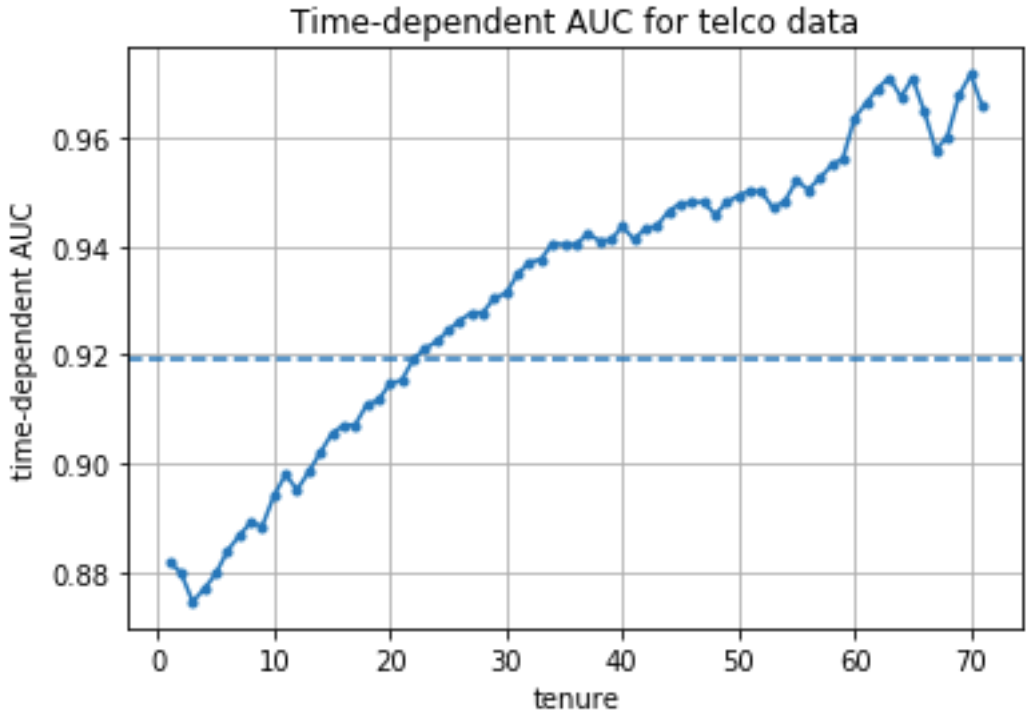
- we find out the there are 2 features that can be ignored:
- MultipleLines=No phone service; MonthlyCharges
- time-dependent AUC: 0.9194



Feature Selection

- Univariate feature selection using SelectKBest

- time-dependent AUC: 0.9190



Survival Analysis Algorithms Comparison

model	concordance_index_censored	concordance_index_ipcw	time-dependent AUC
Cox PH	0.8581	0.8593	0.9192
Cox PH with l1 feature selection	0.8674	0.8334	0.9194
Cox PH with SelectKBest feature selection	0.8581	0.8540	0.9190
survival tree	0.8350	0.8436	0.9026
random survival forest	0.8375	0.9159	0.8951
Survival SVM	0.8623	0.7715	0.9273
Survival Kernel SVM	0.8440	0.8332	0.9117
Gradient boosting Survival analysis	0.8474	0.6837	0.9106

Model Performance Summary

- All of these models work well except Gradient Boosting Survival Analysis if evaluated by concordance_index_ipcw.

- If the churn rate over specific time range is of primary interest, Survival SVM outperforms other models with the highest time-dependent AUC score.

Business Implication

- .predict() method of Survival SVM model can generate risk scores
- If samples are ordered according to predicted risk score, the sequence of events(churn) is obtained as predicted by the model
- Lower ranks indicate shorter survival, higher ranks longer survival
- Saving customers with higher risk score (shorter survival time) should be top priority for the business
- The model is able to correctly predict the order of churn occurrence for 86.23% of all comparable sample pairs, as compared to 50% of random guess