Telecom Customer Churn Analysis

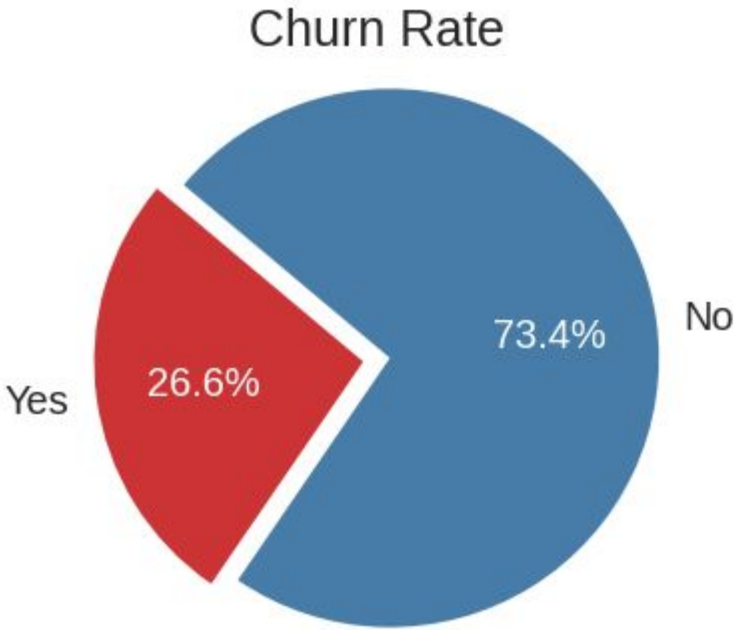
Amalia Temneanu, Dae-Jin Rhee, Darya Bomberger, Marcela Ulloa, Solomon Araya

Churn Rate

In the marketing, cost spend by the company to acquire a new customer is greater than retaining existing customers.

The goal of this Analysis is to create a model that can help the company to reduce the Churn Rate.

The model should predict customer churn, giving the company the opportunity to act beforehand and increase customer retention.



Methodology

Problem Definition

- Determine customers who are likely to churn
- Establish features correlated with churn rate

Data Preparation

- Import Telecom users dataset
- Clean the data
- Explore data for important information
- Analyse target column and features
- Visualize the data

Machine Learning Model

- Split data into train/ test sets
- Define pipeline for preprocessing
- check : Pycaret to determine best model
- Tune the model

Evaluation and Feature Importance

- Create pipeline to compare models
- Use SHAP to establish feature importance
- Add custom metric to determine customer value

Dataset – Customer Churn

Demographic Information

- Gender
- Marital Status
- Seniority
- Dependents

+

Contract Information

- Contract Type
- Tenure
- Payment Method
- Monthly Charges
- Total Charges

+

Services Information

- Phone Service
- Number of Lines
- Internet Services
- Online Security/Backup
- Tech Support
- Streaming

+

Churn

No Churn

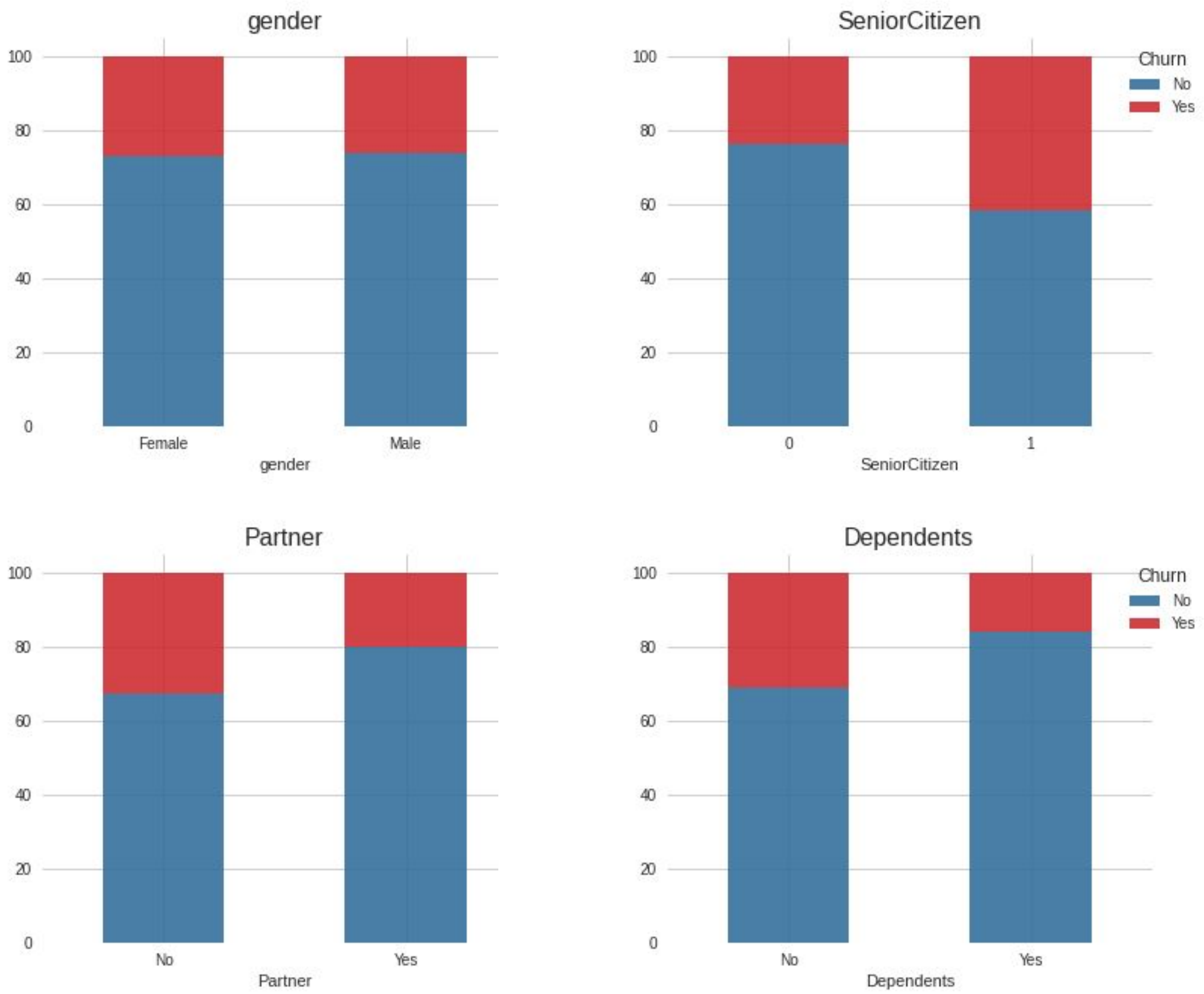
→

Dataset

5976 Samples

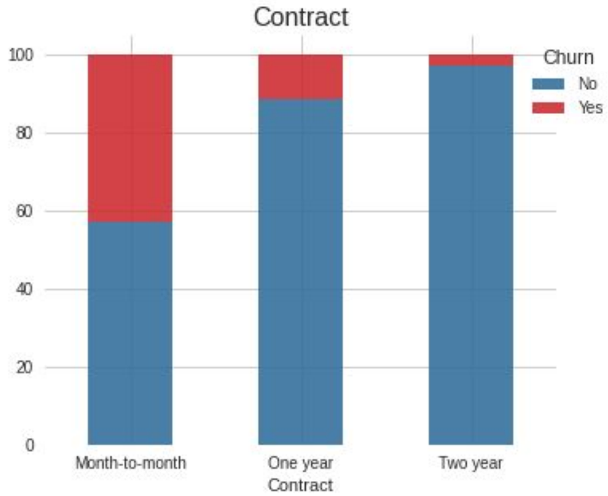
Churn Demographics

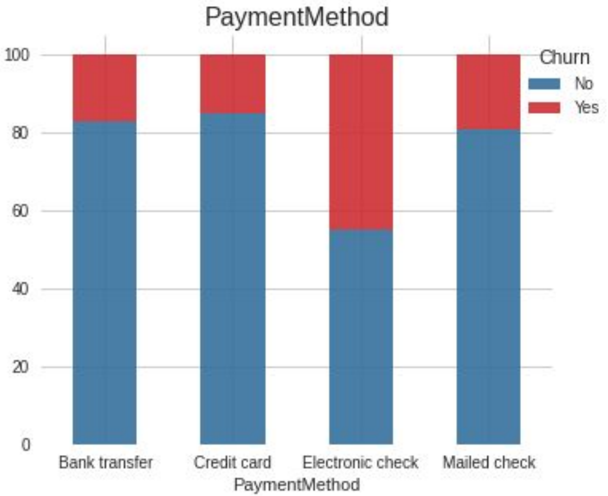
- No difference in churn rate regarding gender.
- Churn rate of **senior citizens** is almost double that of young citizens.
- Customers with a partner churn less than customers with **no partner**.

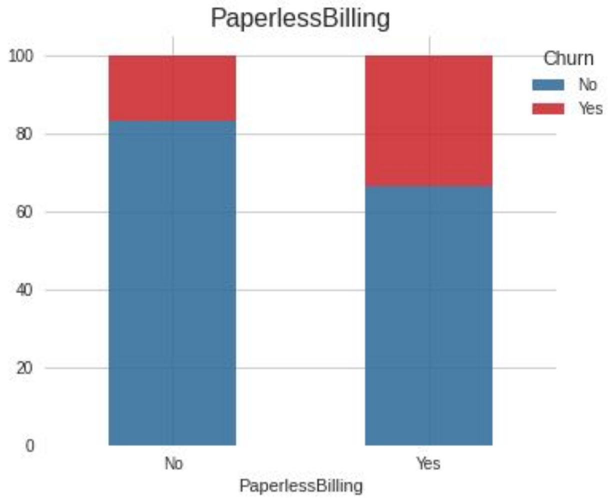


Churn and Customer Account

- Customers with **month-to-month** contracts have a higher churn rate.
- Customers who opted for an **electronic check** as paying method are more likely to leave the company.
- Customers subscribed to **paperless billing** churn more than those who are not subscribed.

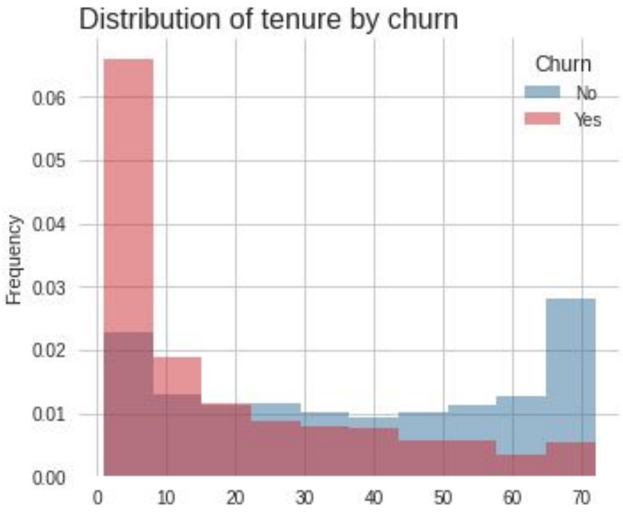


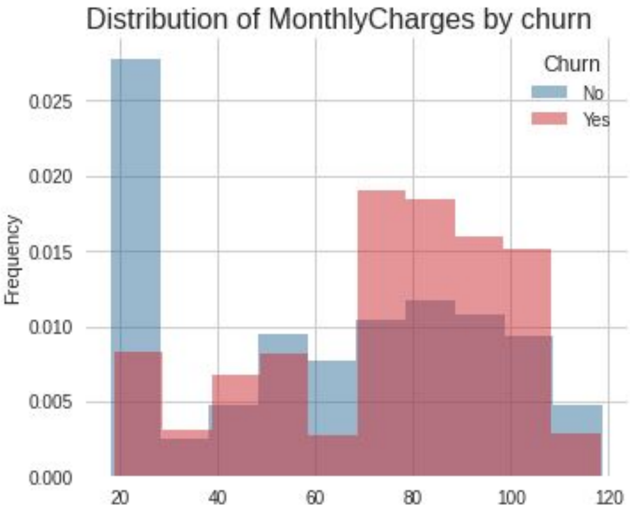


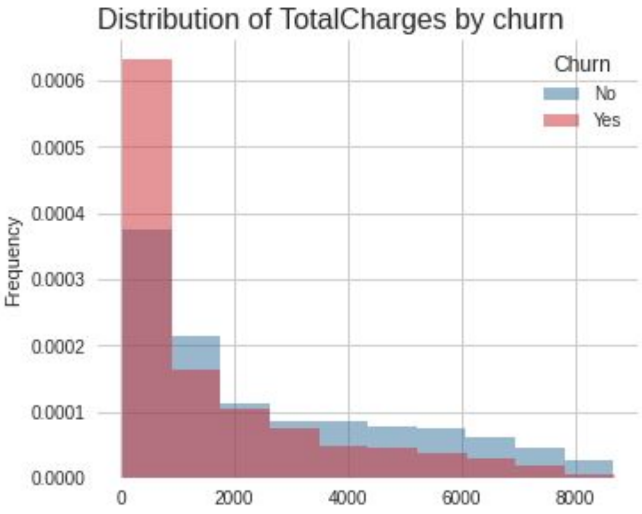


Churn and Customer Account

- The churn rate tends to be larger when **monthly charges are high.**
- **New customers** (low tenure) are more likely to churn.
- Clients with **high total charges are less likely** to leave the company.

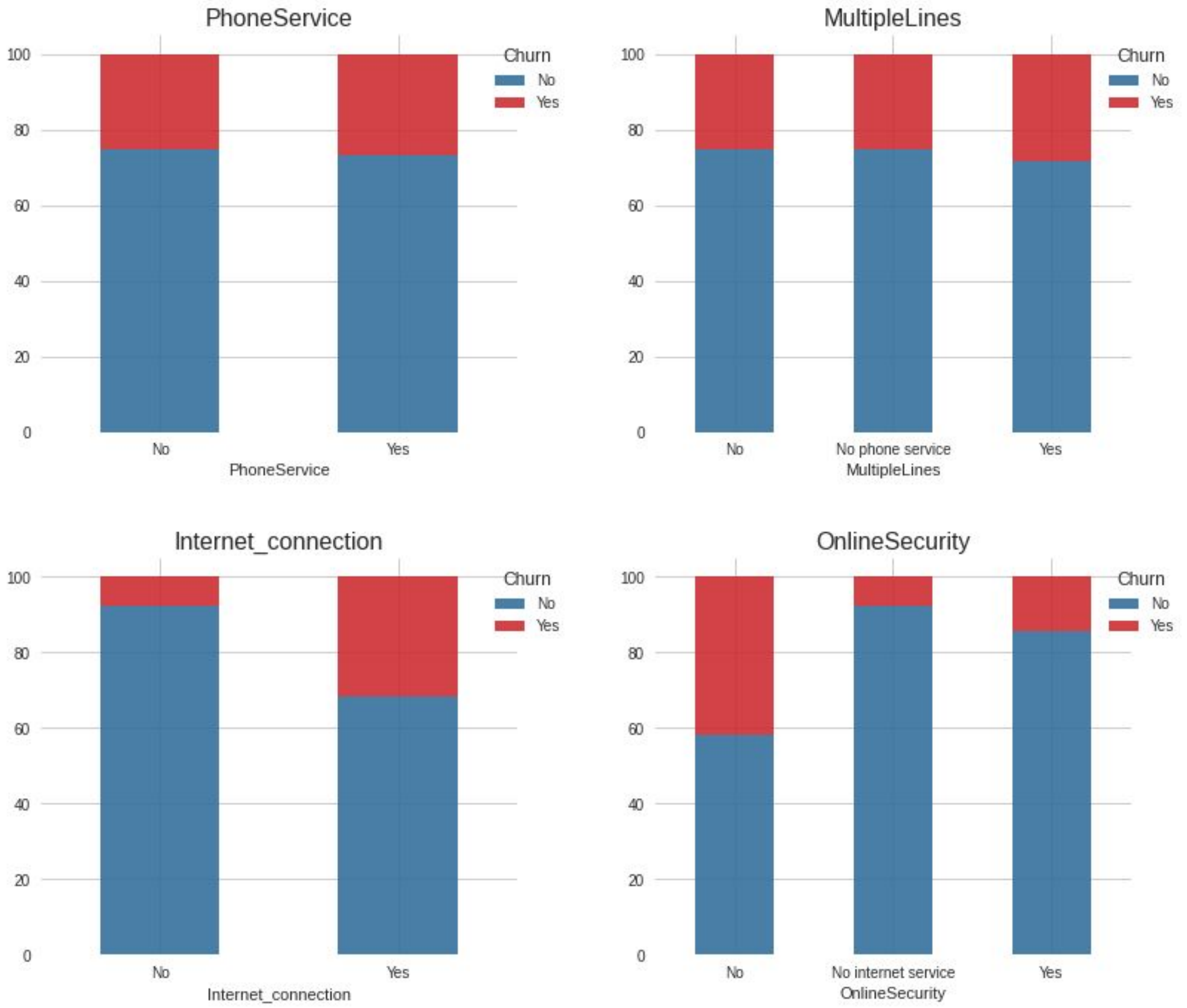






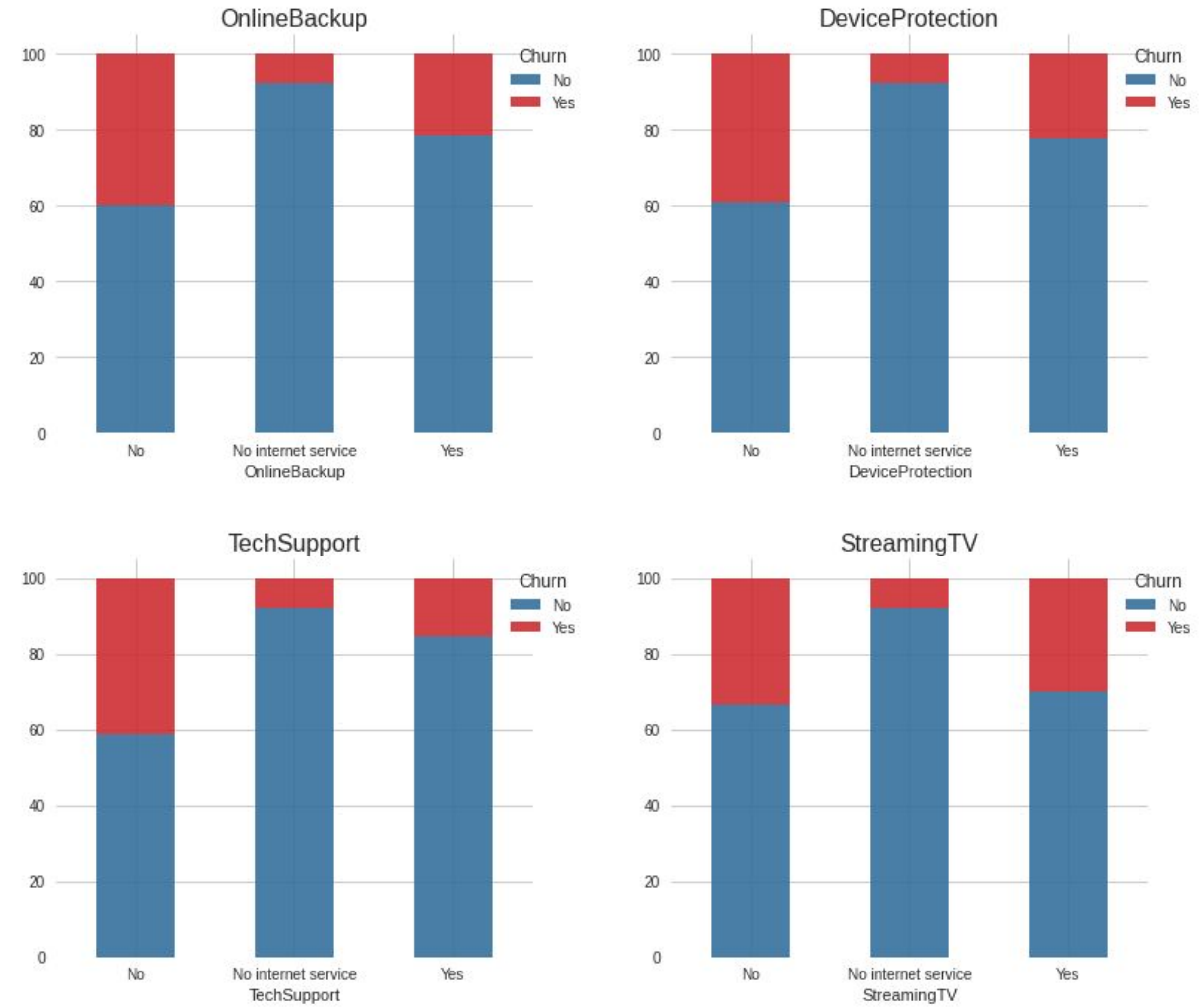
Churn vs Services

- We do not expect phone attributes (PhoneService and MultipleLines) to have significant predictive power.
- The percentage of churn for all classes in both independent variables is nearly the same.
- Clients with **online security churn less** than those without it.



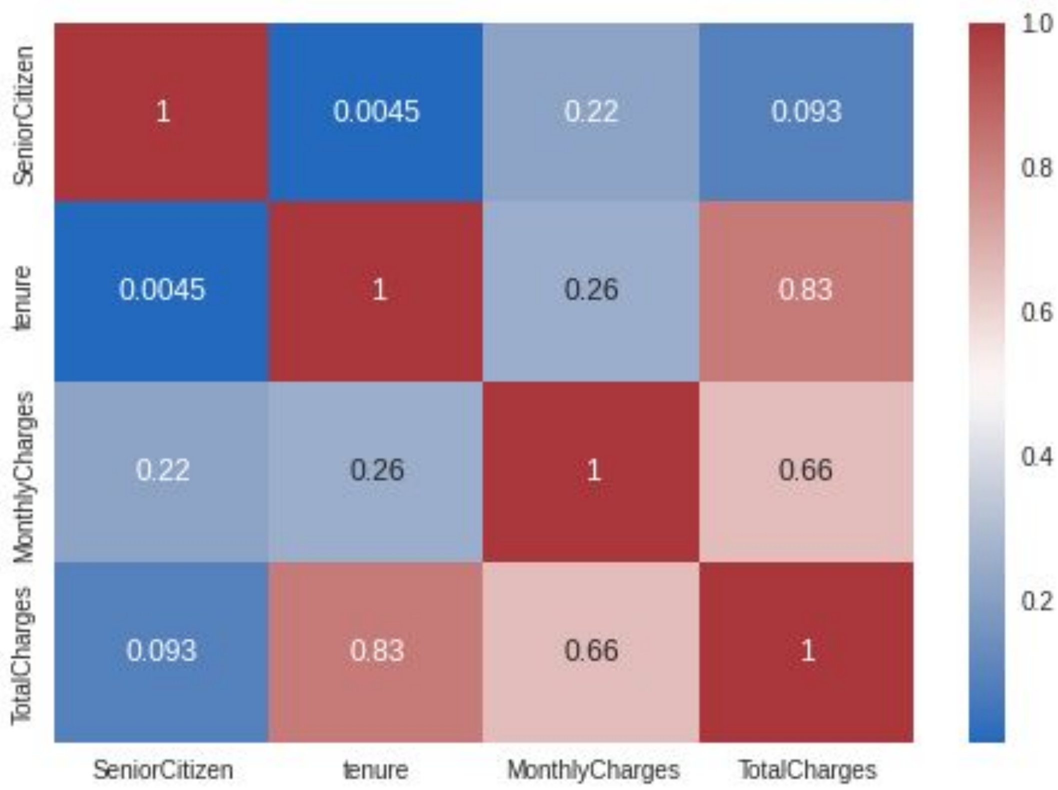
Churn vs Services

- Customers with **no tech support** tend to churn more often than those with tech support.
- Customer with **less extra packages** tend to churn more.



Correlation Map

- **Total Charges** is closely correlated with **Monthly Charges** and **tenure**



Cost of Churn

On average the monthly cost for customers that churn is higher around $15 per month.

If we offer a **$180 annual voucher** to all the customers flagged by the model as potential churn, **we gain $720 per customer in 1 year.**

Churn	Contract	Monthly Charges	Total Charges
No	Month-to-month	$62	$1,531
No	One year	$62	$2,881
No	Two year	$60	$3,711
Yes	Month-to-month	$73	$1,159
Yes	One year	$86	$4,189
Yes	Two year	$87	$5,379

We created a **Custom Metric** in PyCart - Profit - to select the model that maximizes the business value.

If we predict churn and the real value is churn, we gain ($720 - $180) per customer

If we predict churn and the real value is not churn, we loose (- $180) per customer

Model Recommendation

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	Profit	TT (Sec)
9	Naive Bayes	0.7350	0.8202	0.7553	0.5051	0.6049	0.4175	0.4369	21510.0000	0.0190
0	Logistic Regression	0.7951	0.8366	0.5248	0.6468	0.5789	0.4454	0.4501	18162.0000	0.8120
5	Ada Boost Classifier	0.7920	0.8379	0.5273	0.6363	0.5761	0.4400	0.4437	18072.0000	0.0960
6	CatBoost Classifier	0.7859	0.8290	0.4955	0.6285	0.5533	0.4152	0.4208	16848.0000	2.2030
4	Extreme Gradient Boosting	0.7746	0.8071	0.5123	0.5955	0.5493	0.4004	0.4034	16722.0000	0.2770
1	K Neighbors Classifier	0.7575	0.7631	0.5033	0.5543	0.5265	0.3643	0.3658	15570.0000	0.0780
3	Random Forest Classifier	0.7777	0.8097	0.4598	0.6155	0.5260	0.3847	0.3920	15408.0000	0.2580
8	SVM - Radial Kernel	0.7859	0.7613	0.4216	0.6585	0.5133	0.3845	0.4006	14760.0000	0.5760
2	Decision Tree Classifier	0.7254	0.6570	0.5021	0.4894	0.4949	0.3067	0.3072	13842.0000	0.0150
7	SVM - Linear Kernel	0.7429	0.0000	0.4606	0.6684	0.4234	0.2947	0.3490	13608.0000	0.0330

Confusion Matrix

For a Test Sample of 1793 we would gain in this exemple $18,162 with the Proposed Logistic Regression Model.

```
Classification Report:
              precision    recall  f1-score   support

           0       0.84      0.90      0.87      1317
           1       0.66      0.51      0.57       476

    accuracy                           0.80      1793
   macro avg       0.75      0.71      0.72      1793
weighted avg       0.79      0.80      0.79      1793

----------------------------------------------------------
Standard Confussion Matrix (error matrix):
 [[1190  127]
 [ 233  243]]
Accuracy Score obtained is: 79.92%
----------------------------------------------------------
f1_macro Score obtained is: 72.15%
----------------------------------------------------------
f1_micro Score obtained is: 79.92%
----------------------------------------------------------
f1_weighted Score obtained is: 79.05%
----------------------------------------------------------
f1 Score obtained is: 57.45%
```

SHAP Analysis

Interpretation

Likelihood to churn is 1.0 because

MonthlyCharges is **20.55**

Tenure is **2** months

Contract is **Month-to-month**

Interpretation

Likelihood to churn is 0.0 because

MonthlyCharges is **35.4**

Tenure is **50**

Contract is **One year**

Recommendations / Future Scope

Provide vouchers to clients

- Retain customers by offering them vouchers of $180
- Extend the customers value by $720 if kept for an additional year

Offer more personalized discounts

- Engage with the product managers, marketing and financial department to offer potential churn customers personalized discounts and packages to deter churning