Telecommunication
Churn Analysis

Final Project

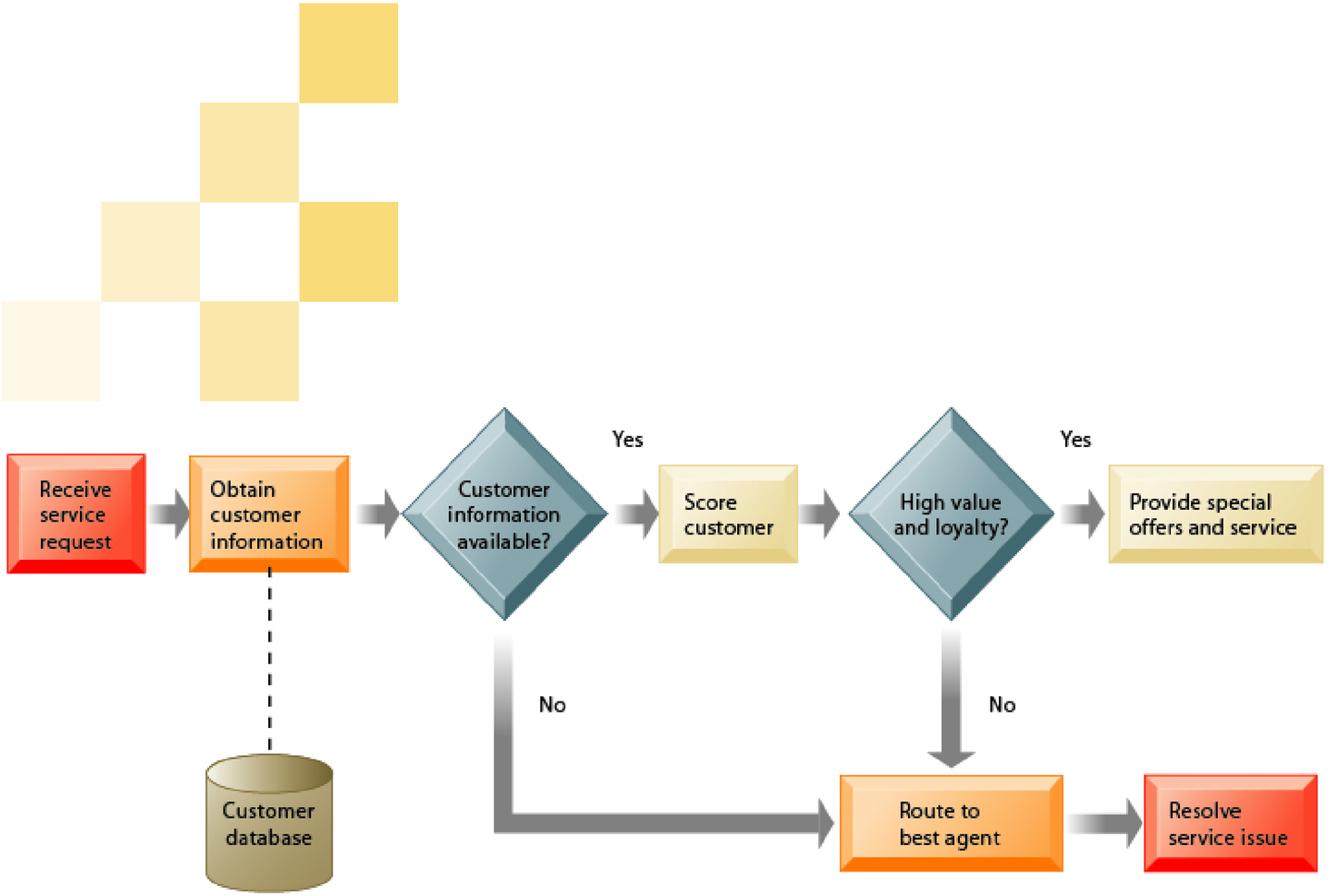
- Ghifari Umar Noorman
- Rully Hilman Simeon
- Yudo Witnu Prasetyo

Final Project Contents

- Context
- Problem Statement
- Business Problem
- Project Goals
- Data Description
- Business Question
- Metric Evaluation
- Explanatory Data Analysis
- Machine Learning
- Classification Report & Confusion Matrix
- Conclusion & Recommendation

Context

- Connectivity
- Telecommunication Infrastructures
- Company competitiveness
- Marketing & CRM



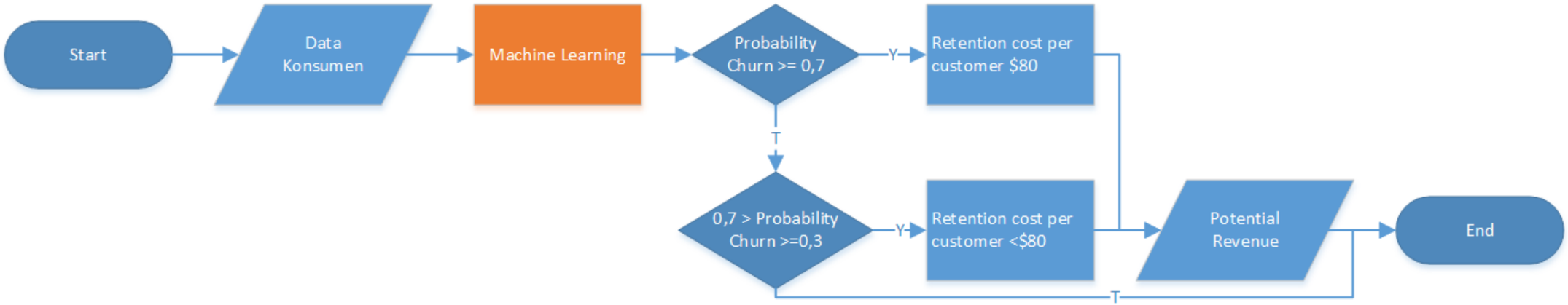
Problem Statement

- new entrants and competitors
- telecommunication company pays attention more on consumer loyalty
- does not take into account customer with lower transaction frequency but with high transaction value

Before


Start
Data Konsumen
Menentukan faktor-faktor yang mempengauhi kepuasan pelanggan & loyalitas konsumen
Melakukan perhitungan untuk menentukan score loyalitas konsumen
Konsumen memiliki score tinggi
Y
Memberikan special offer and service
T
Potensi kehilangan konsumen
Hilangnya potential revenue
End

Project Goals



- Utilizing Machine Learning in the Business process
- Predicting churn customer for cost efficiency

Data Description

Attribute	Data Type	Description
customerID	Object	Customer ID
gender	Object	Whether the customer is a male or a female
SeniorCitizen	Integer	Whether the customer is a senuior citizen or not (1,0)
Partner	Object	Whether the customer has a partner or not (Yes, No)
Dependents	Object	Whether the csutomer has dependendents or not (Yes, No)
tenure	Integer	Number of months customer has stayed with the company
PhoneService	Object	Whether the customer has a phone service or not (Yes, No)
MultipleLines	Object	Whether the cutomer has multiple lines or not (Yes, No)
InternetService	Object	Customer's internet service provider (DSL, Fiber optic, No)
OnlineSecurity	Object	Whether the customer has online security or not (Yes, No, No internet service)
OnlineBackup	Object	Whether the customer has online backup or not (Yes, No, No internet service)
DeviceProtection	Object	Whether the customer has device protection or not (Yes, No, No internet service)
TechSupport	Object	Whether the customer has tech support or not (Yes, No, No internet service)
StreamingTV	Object	Whether the customer has streaming TV or not (Yes, No, No internet service)
StreamingMovies	Object	Whether the customer has streaming movies or not (Yes, No, No internet service)
Contract	Object	The contract term of the customer (Month-to-month, One year, Two year)
PaperlessBilling	Object	Whether the customer has paperless billing or not (Yes, No)
PaymentMethod	Object	The customer's payment method (Electronic check, Mailed check, Bank transfer (automatic), Credit card (automatic))
MonthlyCharges	Float	The amount charged to the customer monthly
TotalCharges	Object	The total amount charged to the customer
Churn	Object	Whether the customer churned or not (Yes or No)

Source: https://www.kaggle.com/datasets/blastchar/telco-customer-churn

Business Question

- How does type of contract affect churn percentage?
- How does monthly charges paid by subscriber played a role in customer churn?
- Which type of internet service affects customer churn most and how to improve it?
- Which type of payment method affects churn rate the most?
- Does Tenure affects churn rate and how?
- Which type of additional features affects churn and customer amount?

Metric Evaluation

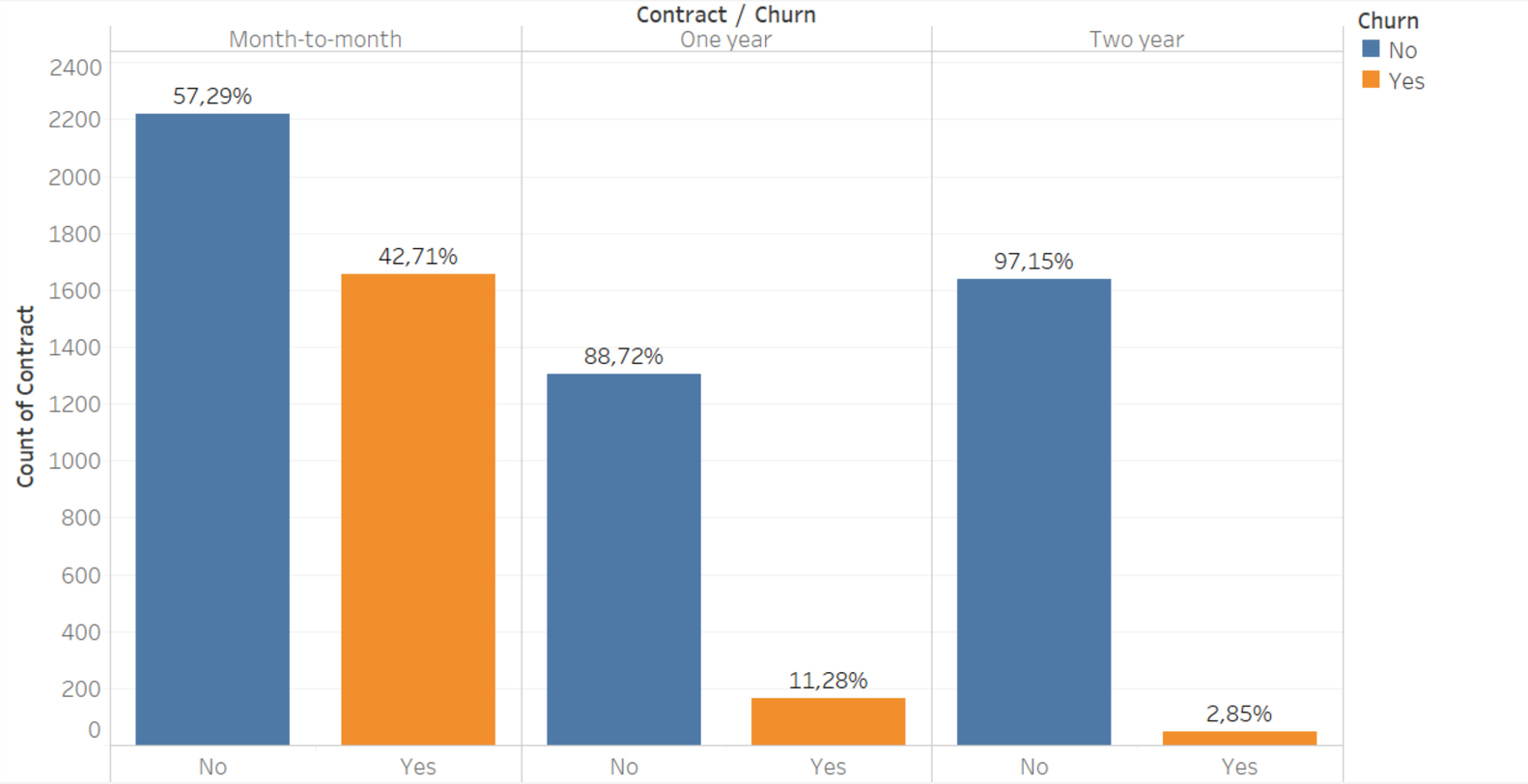
		Predicted		
		Customer Churn (Positive)	Cusomter Doesn't Churn (Negative)	
Actual	Customer Churn (Positive)	True Positive (TP) Model predicts customer churn while it does churn	False Negative (FN) Model predict customer does not churn while it does churn	Recall (+)
	Cusomter Doesn't Churn (Negative)	False Positive (FP) Model predicts customer churn while it does not churn	True Negative (TN) Model predict customer does not churn whilte it doesn't churn	Recall (-)
		Precision (+)	Precision (-)	

- Type 1 error : False Positive (Model predicts customer churn while it does not churn)
- Consequence: retention cost (promotion) spent in vain

- Type 2 error : False Negative (Model predict customer does not churn while it does churn)
- Consequence: losing potention revenue

- **RECALL(+) CHURN**
- **minimize potential loss of revenue by having the least value of recall**

Churn Percentage based on the Contract

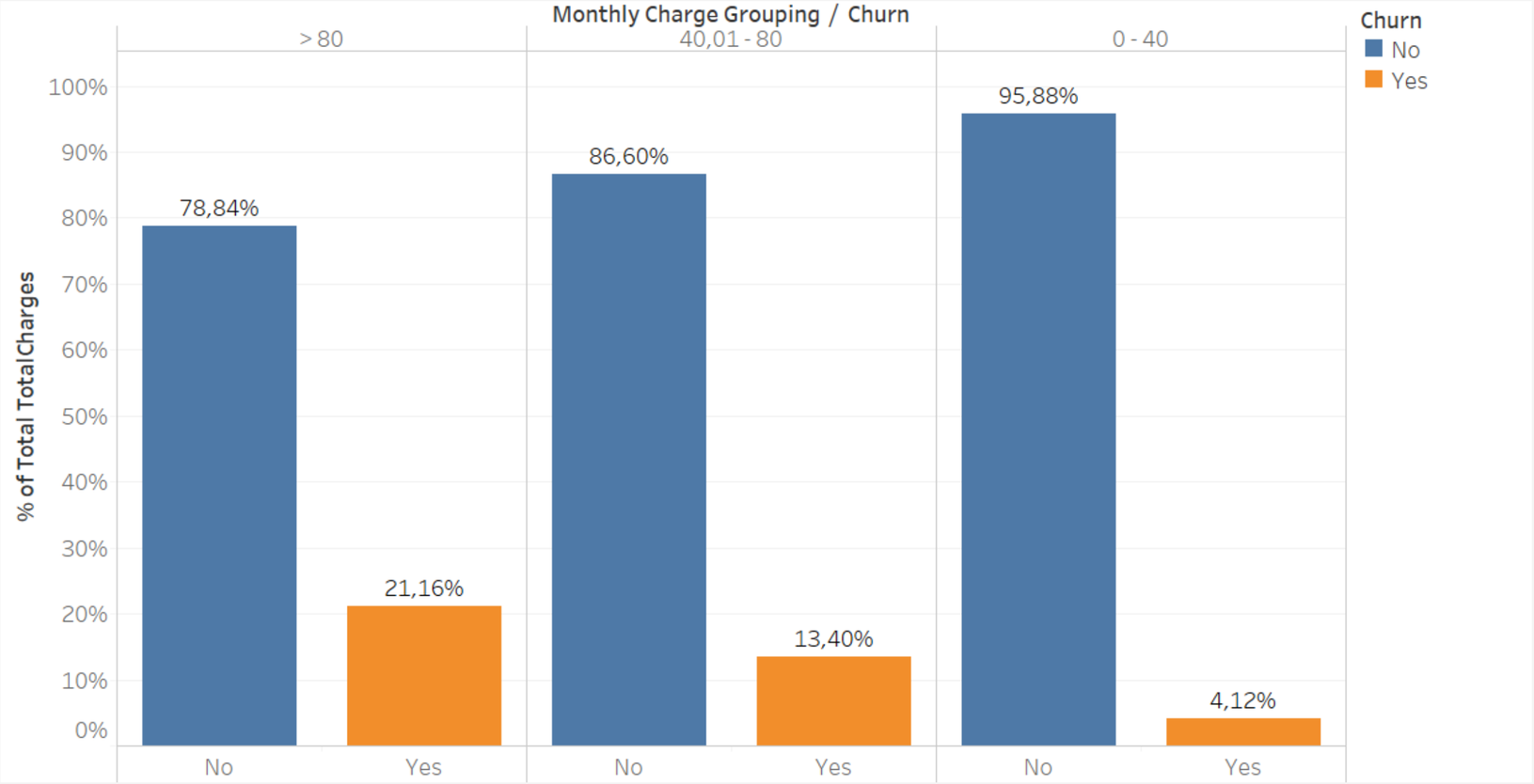


- A two year contract has the lowest churn rate
- A month to month contract has the highest churn rate

Insight:

- Encouraging customer to subscribe a longer contract
- Giving more privilege and incentives for longer contract

Churn Percentage based on Consumer Monthly Charges

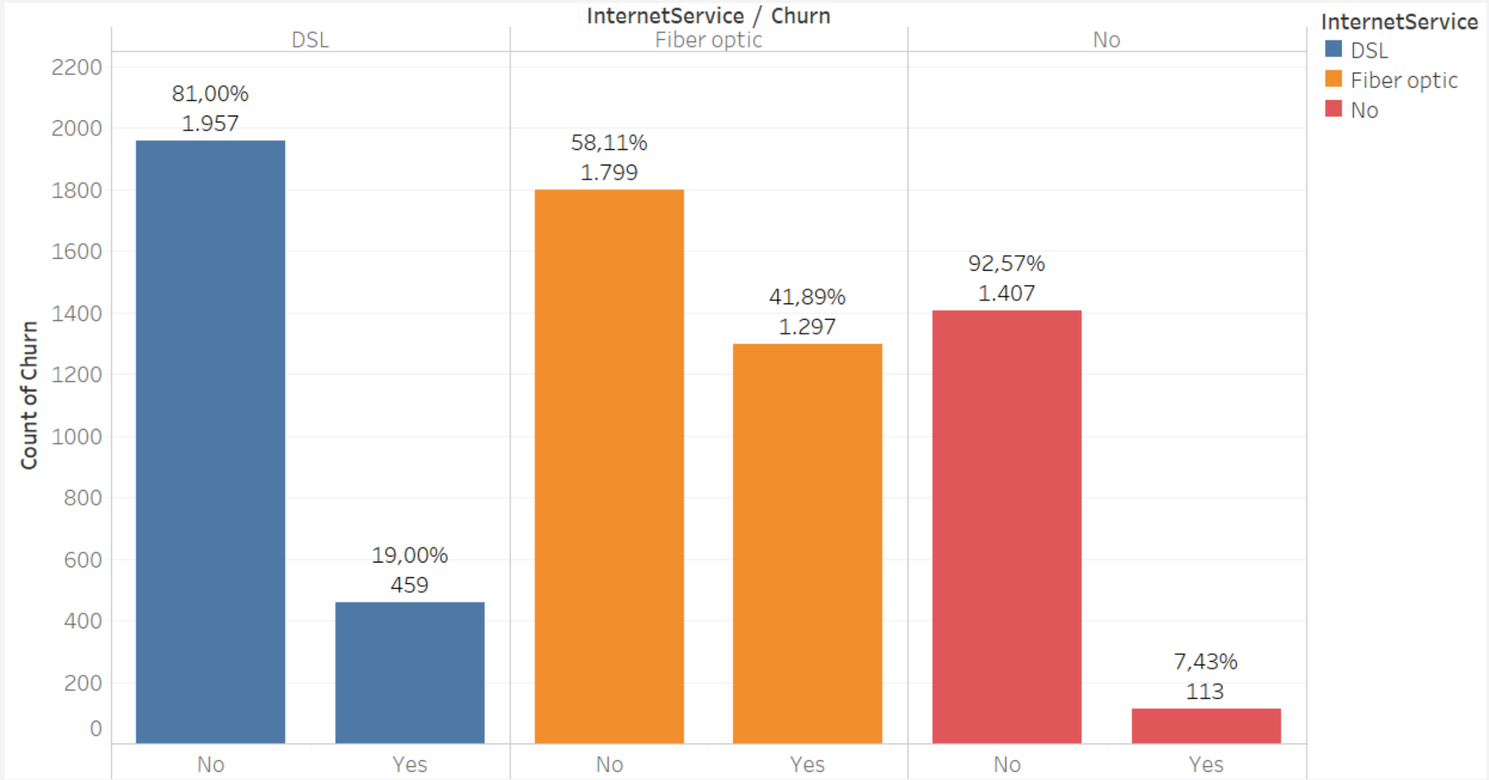


- The lower the charge that customer pay each month the less the churn rate is

Insight:

- Give low cost incentives
- Promotion program such as discount voucher

Churn Percentage based on Internet Service



- DSL has the highest subscriber, followed by fiber optic
- Although Fiber Optic is the best in terms of speed, it has the highest churn rate

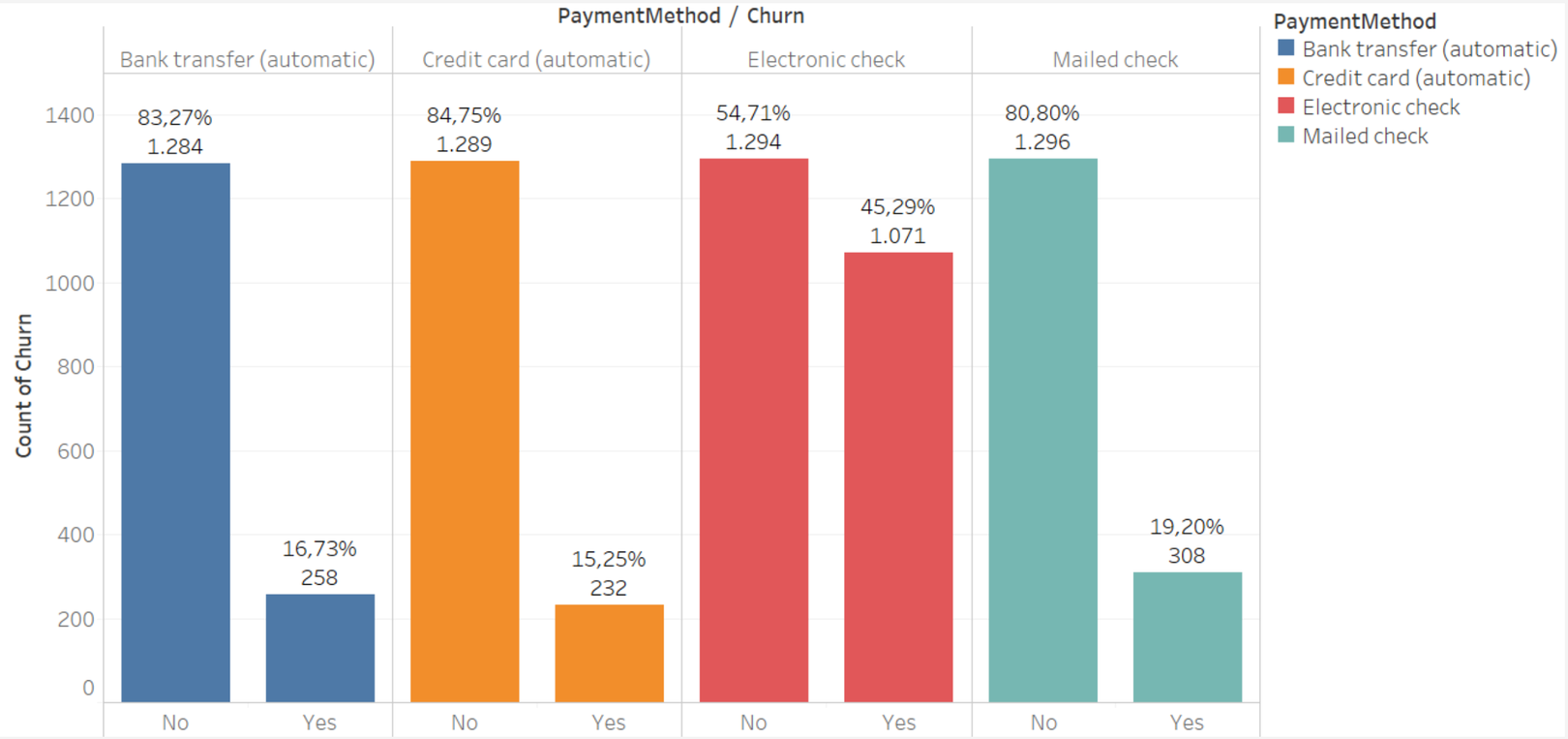
Insight:

- investigation needs to be done in this area

Possible fiber optic problem

- Quality assurance constraint (fragile fiber optic cable)
- Constant signal loss
- High Traffic results in slow speed

Churn Percentage based on Payment method

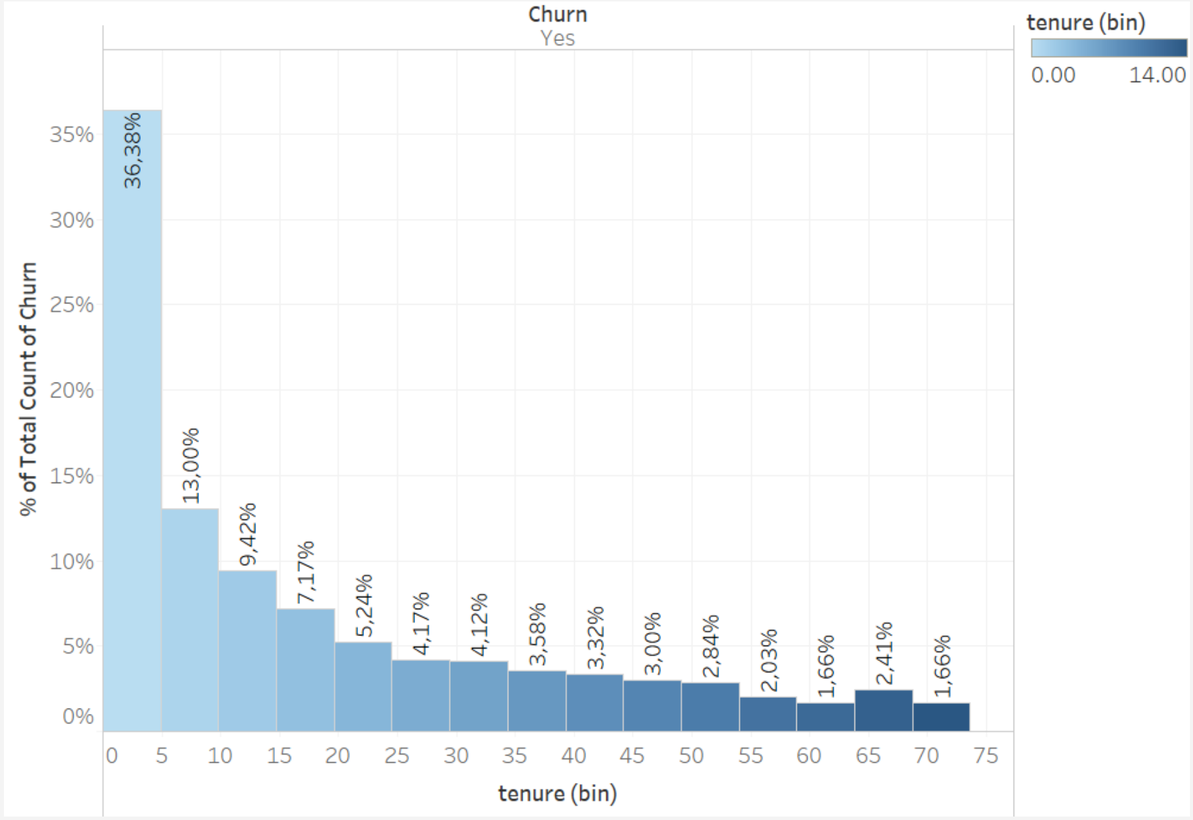


- Most of them has the similar amount
- Electronic checks method has the highest number overall with also the higest number of churn

Insight:

- Encouraging customers towards a simpler payment method

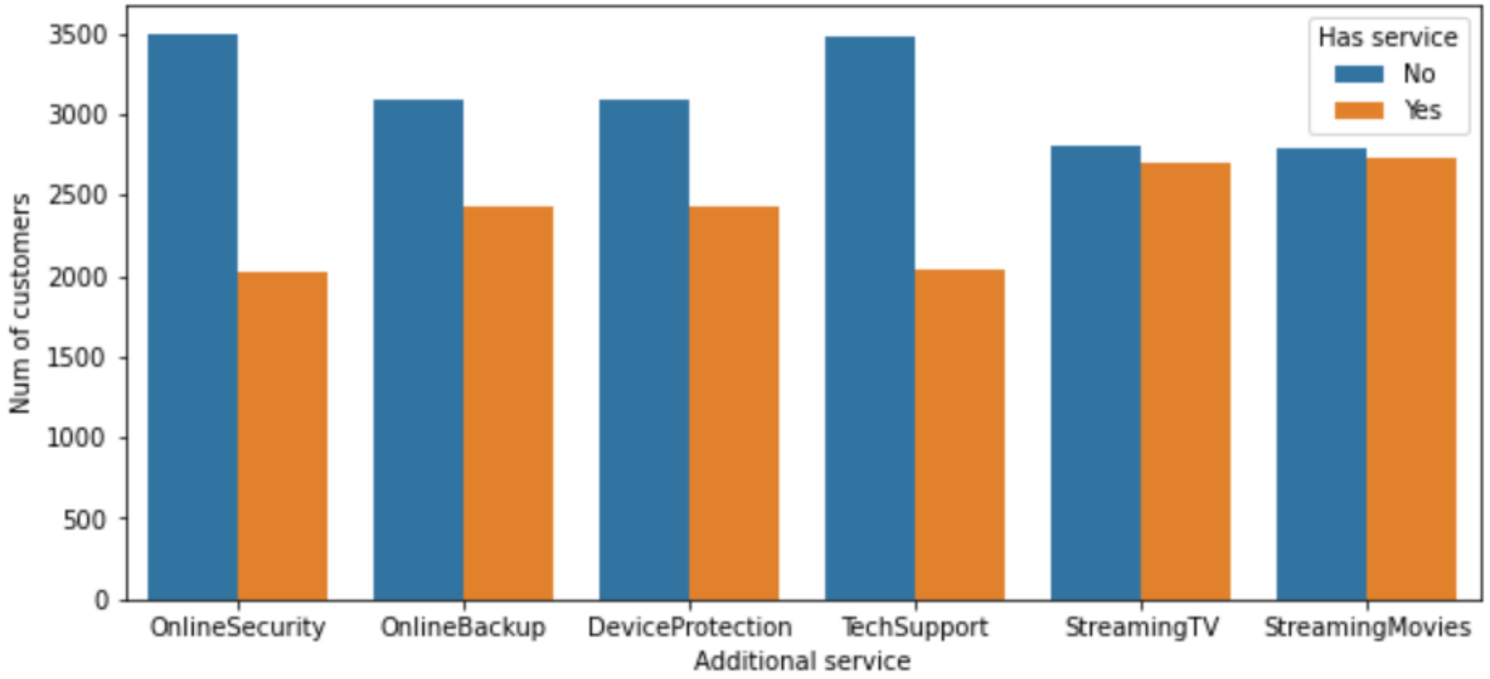
Churn Percentage based on Tenure

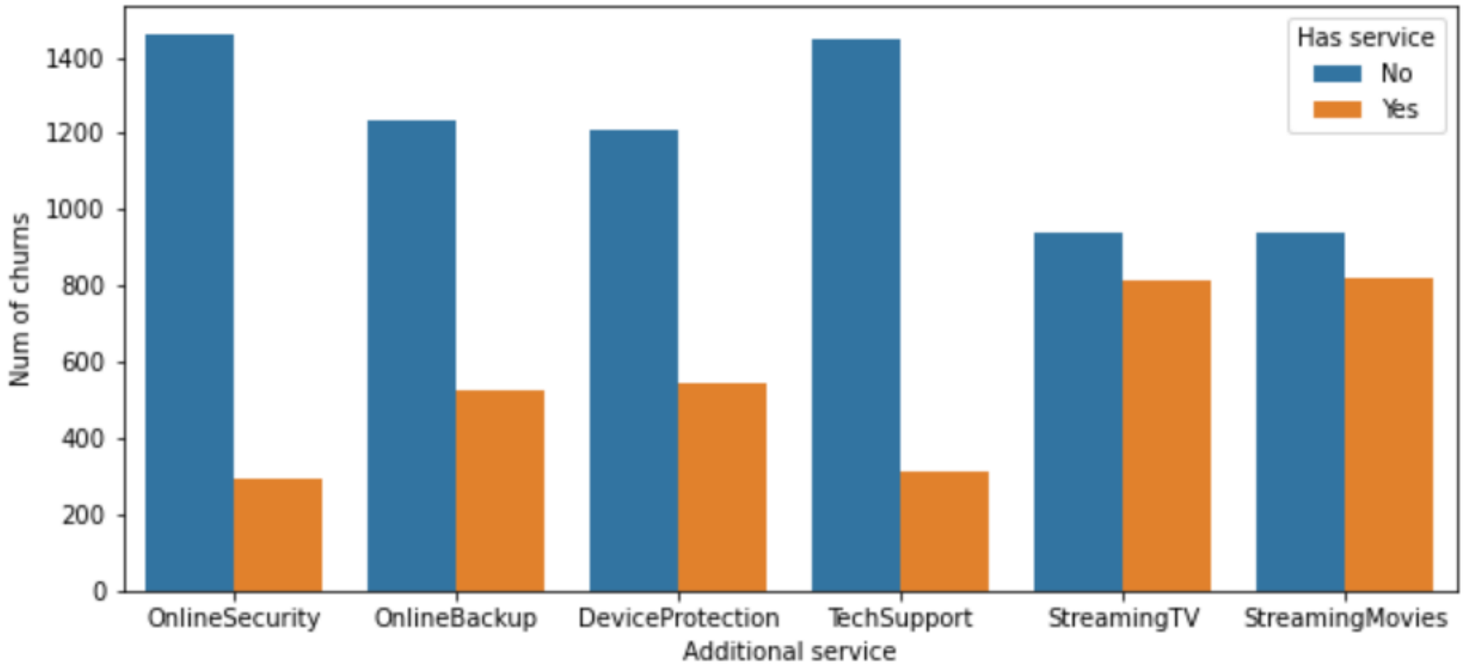


- Consumer churn happens on the first half year
- The longer a customer has subscribed, the less the probability they churn

Insight:

- More active and engaging to customer on the first year
- Developing programs and promotions might be an alternative





Additional Services in Relation with Customer and Churn Amount

- The first plot shows the total number of customers for each additional service
- The second shows the number of clients that churn
- Streaming services has high rate of churn

Classification Report Before Tuning

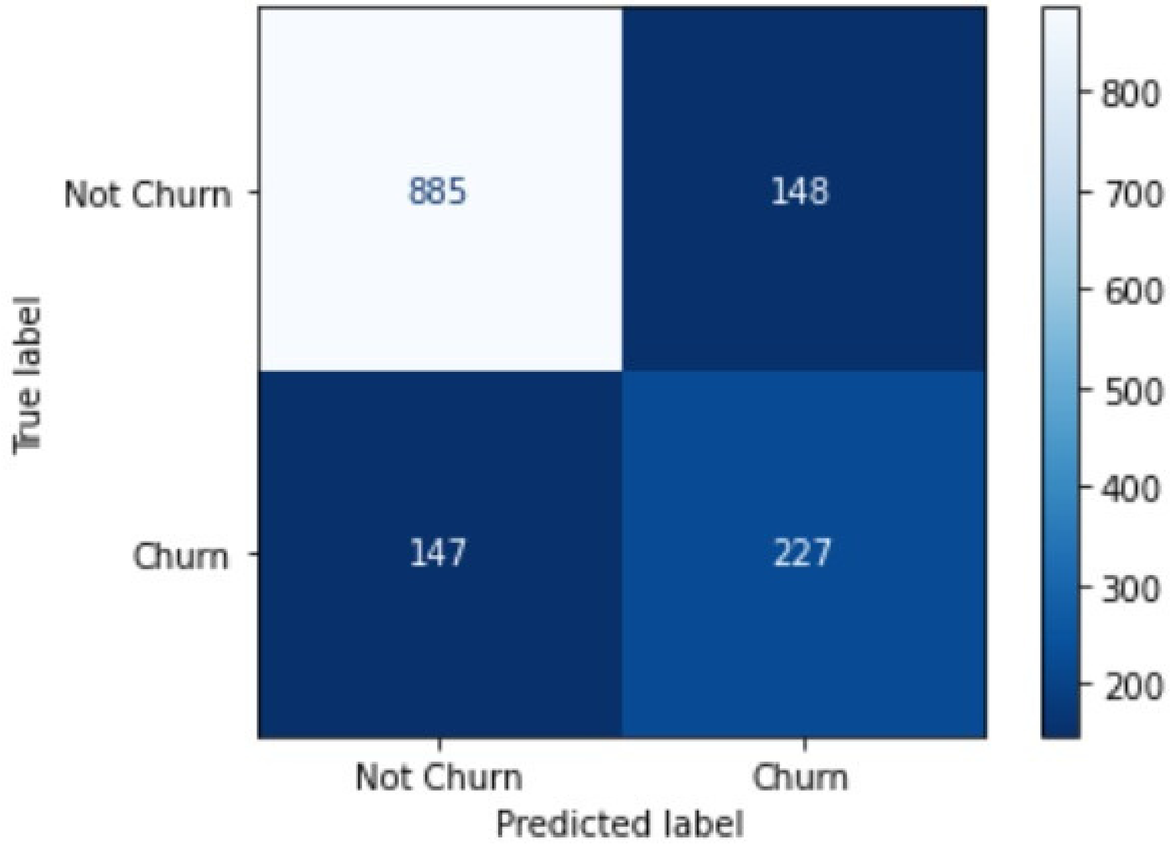
[40]:

	Accuracy	Recall	Precision	F1
Train RF	0.998756	0.996656	0.998660	0.997657
Train XGB	0.937600	0.853512	0.906250	0.879091
Train KNN	0.840889	0.652174	0.722222	0.685413
Test LogReg	0.802416	0.572193	0.644578	0.606232
Test KNN	0.761905	0.553476	0.552000	0.552737
Train LogReg	0.803200	0.547826	0.655200	0.596721
Test XGB	0.775409	0.532086	0.585294	0.557423
Test RF	0.789623	0.508021	0.629139	0.562130

Classification Report After Tuning

	Accuracy	Recall	Precision	F1
Test KNN	0.790334	0.606952	0.605333	0.606142
Train KNN	0.803022	0.589298	0.640727	0.613937
Test LogReg	0.801706	0.577540	0.640950	0.607595
Train LogReg	0.803378	0.551171	0.654488	0.598402

Confusion Matrix



- False Negative and False Positive more or less the same.
- Recall (0.606), Precision (0.605)
- Retention Cost Wasted
  $80 * 147 = $11,760
- Acquiring Cost needed to Recover lost customer
  $300 * 148 = $44,400

Conclusion

- Best Model - KNN (Recall 0.606)
- Month to month contract has the highest churn rate.
- Monthly charge 80 dollars or more has the highest churn rate and highest revenue contributor
- Fiber optic proves to be the highest churn contributor
- Checks method has higher rate of churning
- Longer tenure the lower the probability of churning

Recommendation

Recommendation 1 Based on EDA

- Encourage customer using longer contract
- Formulate new product with 6 month contract
- Offer moreincentives for high spender
- Improve fiber optic serivces
- Encouraging user to migrate to automated payment system

Recommendation 2

- Spending Budget on all customer:
  - Cost per Customer spent: $80
  - Number of Customer: 7043
  - Total Money spend for retaining: $563,440

- Spending Budget on targeted customer (From 7043 customer in the data around 27% are predicted to churn)
  - Cost per Customer spent: $ 80
  - Targeted Customer: 7043 * 27%(churn percentage) = 1902 customer
  - Total Money spend for retaining: $80 * 1902 = $152,129

Money Spend Difference: $563,440 - $152,129 = $411,311

Recommendation 3

Setting a threshold for targeted action.
0 - 0.3 = no action needed
0.3 - 0.7 = half retention cost per customer = $40
0.7 - 1 = full retention cost per customer = $80

Thank You!