TELCO CHURN PREDICTION

A Model for Predicting Customer Retention in Telecom

Peerapat Tancharoen • Data scientist • github.com/peerapat-t

EXECUTIVE SUMMARY

Business problems

- Assume we are a telecommunications company dealing with customer churn issues.
- Retaining current customers is typically more cost-effective than acquiring new ones; studies show it can be 5 to 7 times more expensive to attract new customers.
- We focus on identifying signals and encouraging them to continue their relationship with our business.

Solution

- Develop a churn prediction model with an optimized threshold based on a cost-benefit analysis.

Impact

- The model improved revenue by 159.8% compared to the 'do nothing' program and by 35.7% compared to the 'retain all' program.

Challenges

- Imbalanced dataset
- Threshold selection

1. BUSINESS PROBLEM

1. High Acquisition Costs: The high costs associated with acquiring new customers significantly strain financial resources, emphasizing the need for efficient retention strategies.

2. Resource Allocation Inefficiency: Studies, such as one by Bain & Company, highlight the inefficiency in resource allocation when 5 to 7 times more resources are spent on acquiring new customers than on retaining existing ones.

3. Risk of Customer Churn: A lack of focus on nurturing relationships with familiar customers increases the risk of churn, as these customers might feel neglected compared to the efforts spent on attracting new patrons.

4. Revenue Stream Volatility: Failing to retain customers can lead to unpredictable revenue streams, affecting overall business stability as continuous business transactions are not secured.

5. Need for Predictive Modeling: Implementing a predictive model is essential for proactively identifying at-risk customers, enabling timely interventions that can enhance retention rates and stabilize revenue.

2. SOLUTION

Do nothing

- Business not take any action to retain customers.

Retain all customers

- Business take action to retain every customer.

Use a churn prediction model

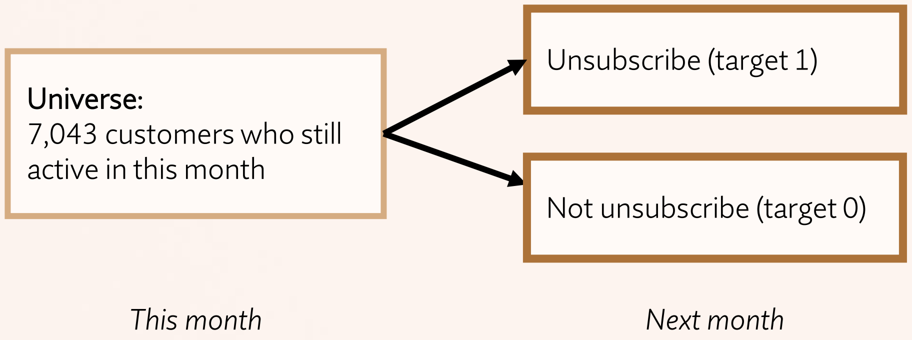
- Business use churn prediction model to predict the probability of churning and use this score to decide whether a customer will churn.
- Churn prediction can help businesses identify customers who are at risk of leaving.
- This allows businesses to take action to retain those customers, such as offering them discounts or special promotions.

Option	Cost	Effectiveness
Do nothing	Least expensive	Least effective
Retain all customers	Most expensive	Most effective
Use a churn prediction model	More cost-effective than retaining all customers	More effective than doing nothing

2. SOLUTION

Target

- There are 7,043 customers in the dataset, including churn and not-churn customers.
- Churn customers are those who have unsubscribed within the last month.



Features

- The dataset contains 22 features, which can be categorized into 3 groups:
  1. Services that each customer has signed up for: This includes phone, multiple lines, internet, online security, online backup, device protection, tech support, and streaming TV and movies.
  2. Customer account information: This includes how long they've been a customer, contract, payment method, paperless billing, monthly charges, and total charges.
  3. Demographic info about customers: This includes gender, age range, and if they have partners and dependents.

3. RESULT

Model result

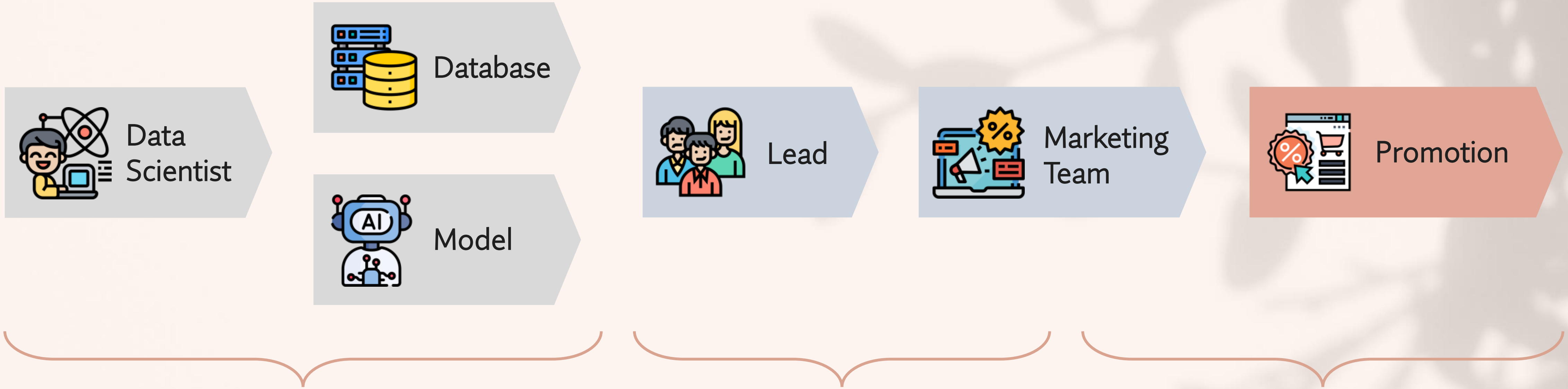
Method	Total (Test)	Actual churn	Overspend (FP)	Save (TP)	Gains (or Loss)
Do nothing	1,409	371	0	0	-185,500
Retain all	1,409	371	1038	371	81,700
Churn prediction model	1,409	371	113	258	110,900

** These calculations are based on the assumption that the cost of promotion is 100, the cost of loss is 500, and the savings per customer is 500.*

- Utilizing data from the test set involving approximately 1,409 customers, the churn prediction model yielded a gain of 159.8% when compared to the 'do nothing' program, and 35.7% when compared to the 'retain all' program.

4. APPLICATION

Collaboration framework



Step1: Data scientist team build churn prediction model by data from database.

Step2: Model predict probability to churn and automatically send the lead to marketing team.

Step3: Marketing team launch promotion only for customer who have higher probability to churn

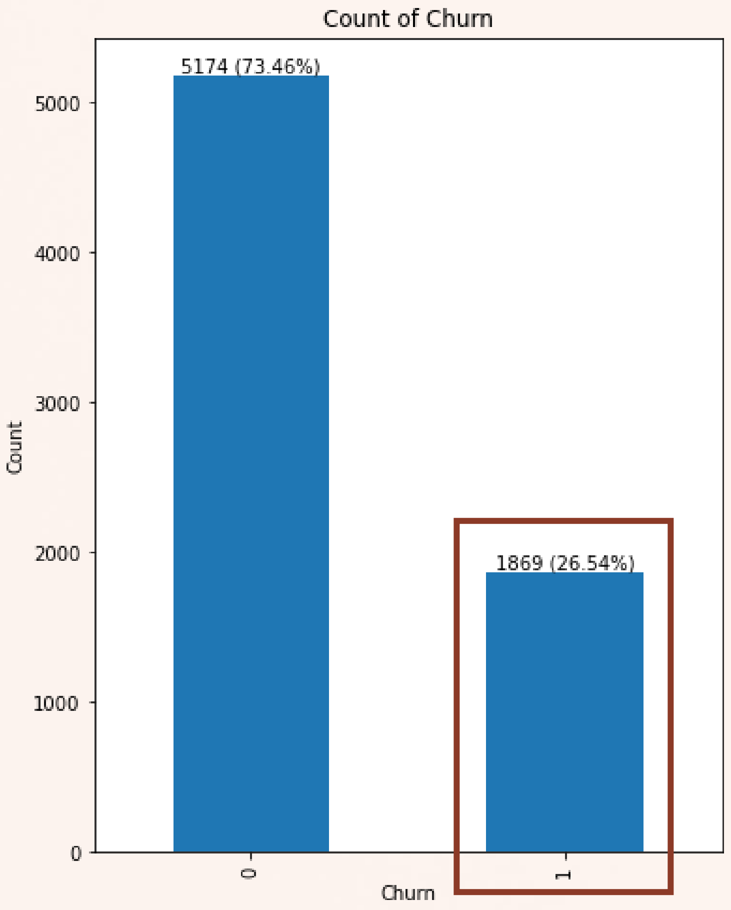
5. APPENDIX

Technique

Step	Topic	Cost
1	Dataset	• Split the data into train, test, and validation sets. • Use the train set to train the model, the validation set to tune the threshold, and the test set to evaluate the model.
2	Normalized	• Use Min-Max scaling to normalize the data.
3	Resampling	• Try oversampling (SMOTE, ANASYN) and under sampling.
4	Model	• Try random forest, LightGBM, and XGBoost.
5	Hyperparameter tuning	• Use RandomizedCV to find the best hyperparameters for each model.
6	Threshold tuning	• Use cost-sensitive learning to tune the threshold. • Assign a cost of 5 times more to acquiring new customers than to retaining existing customers.
7	Interpreting	• Use SHAP values.

5. APPENDIX

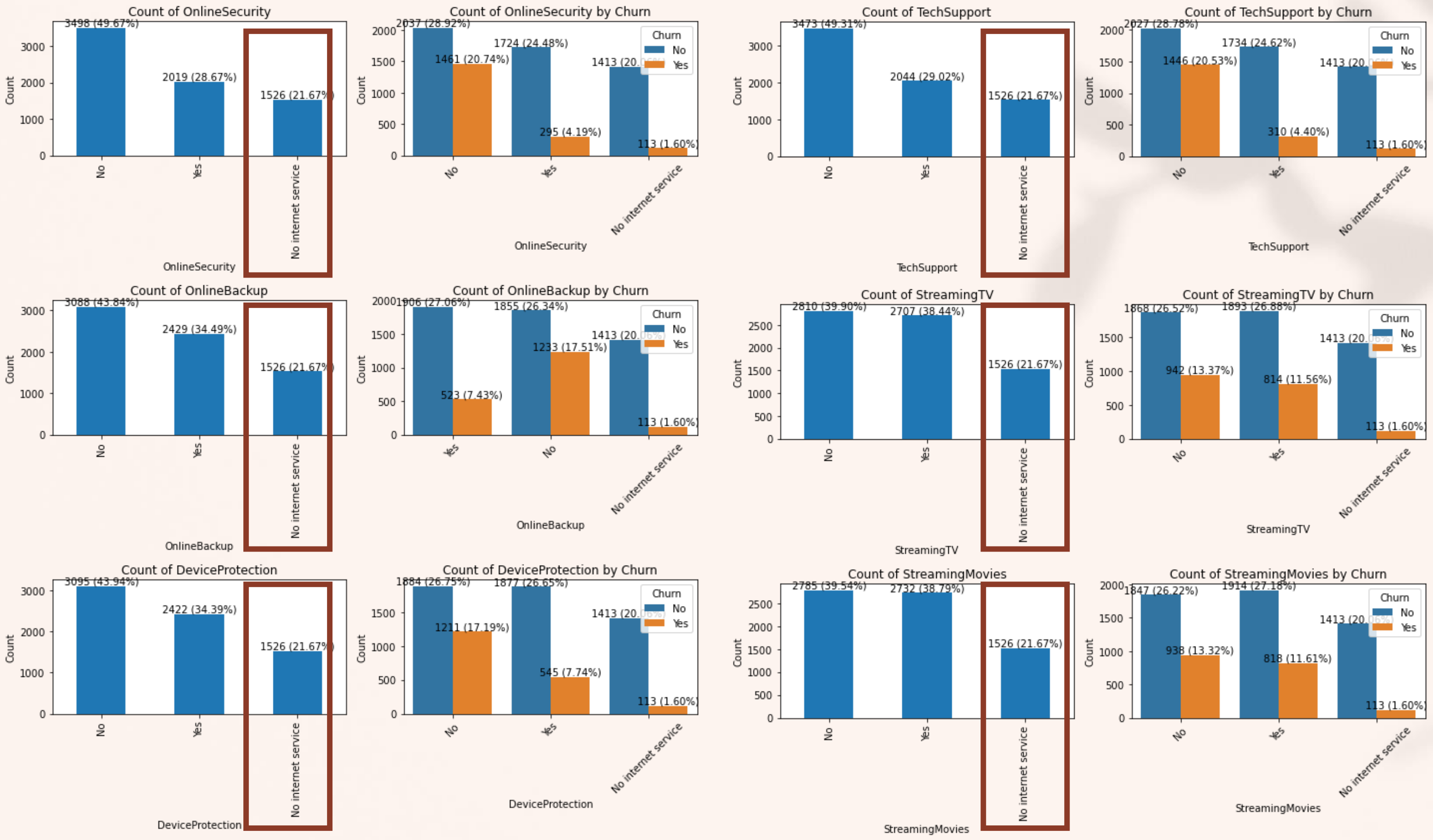
Exploratory data analysis (EDA)



- Only 27% of the customers have churned, which means the data is imbalanced.
- Resampling method or threshold tuning should be applied to handle this problem. Metrics like accuracy can be misleading and ineffective.
- Instead, it is essential to explore alternative evaluation measures, such as precision, recall, F1-score, or AUC-ROC.

5. APPENDIX

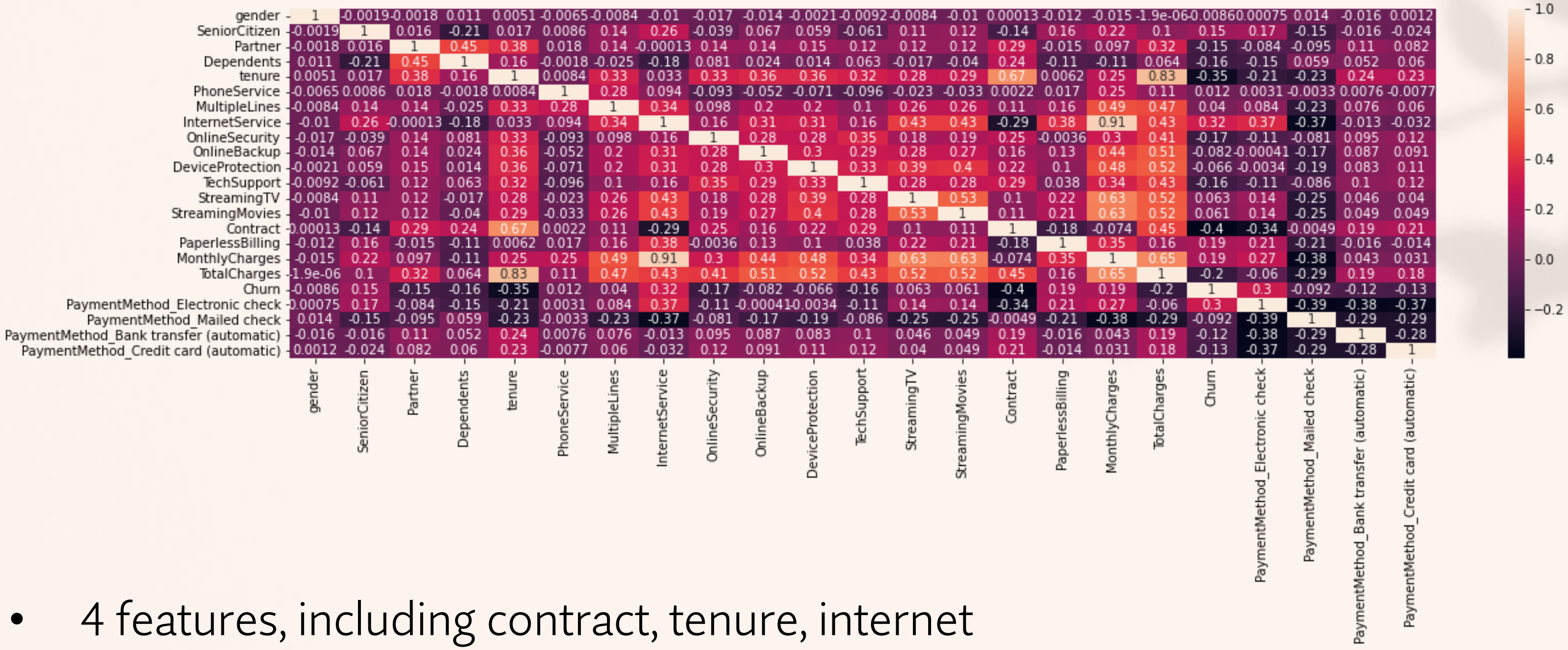
Exploratory data analysis (EDA)



- “No internet service” appears in 6 variables, all with the same frequency of 1,526.

5. APPENDIX

Exploratory data analysis (EDA)

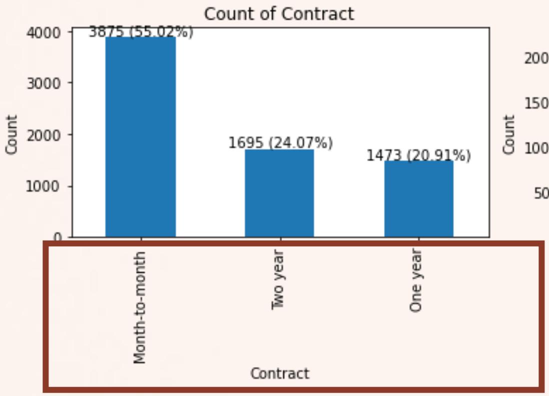


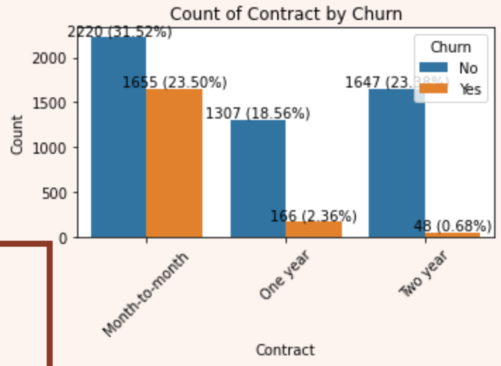

```
Churn                                      1.000000
Contract                                   0.396713
tenure                                     0.352229
InternetService                            0.316846
PaymentMethod_Electronic check             0.301919
TotalCharges                               0.199037
MonthlyCharges                             0.193356
PaperlessBilling                           0.191825
OnlineSecurity                             0.171226
TechSupport                                0.164674
Dependents                                 0.164221
SeniorCitizen                              0.150889
Partner                                    0.150448
PaymentMethod_Credit card (automatic)      0.134302
PaymentMethod_Bank transfer (automatic)    0.117937
PaymentMethod_Mailed check                 0.091683
OnlineBackup                               0.082255
DeviceProtection                           0.066160
StreamingTV                                0.063228
StreamingMovies                            0.061382
MultipleLines                              0.040102
PhoneService                               0.011942
gender                                     0.008612
Name: Churn, dtype: float64
```

- 4 features, including contract, tenure, internet service, and electronic payment method, exhibit a high correlation (greater than 0.3 in absolute terms) with the target variable.

5. APPENDIX

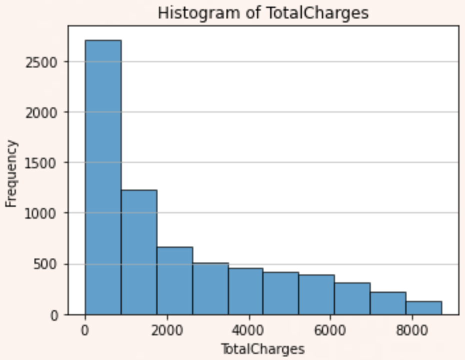
Exploratory data analysis (EDA)

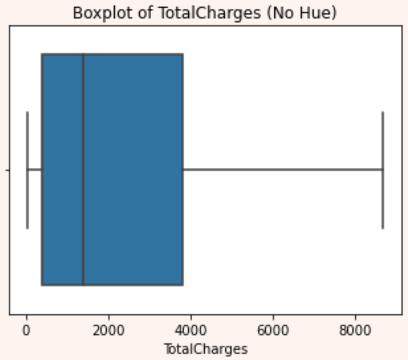


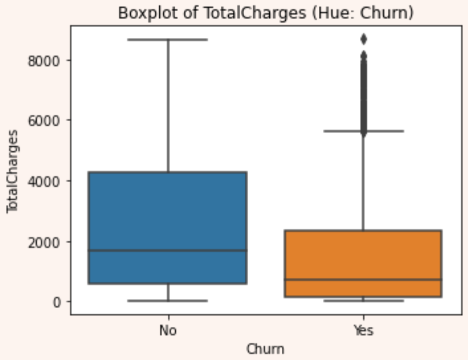


0, 1, 2

- We can represent the "Contract" variable as an ordinal variable with values 0, 1, and 2.







- The distribution of “TotalCharges” is right-skewed. Use the median to replace missing values instead of the mean.

5. APPENDIX

Model evaluation

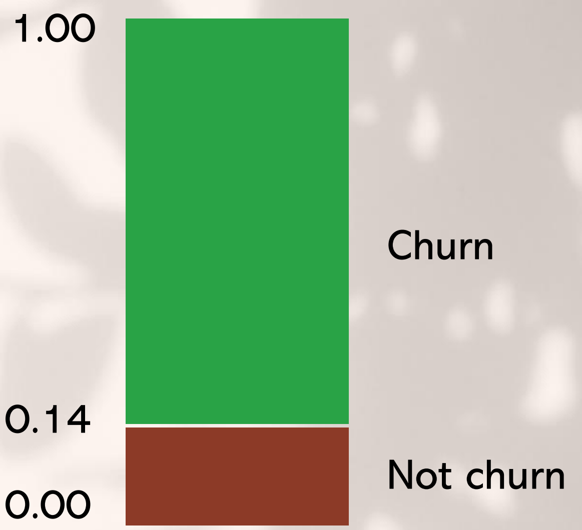
	Model	Observation	TP	TN	FP	FN	Precision	Recall	AUCROC	Best Threshold	Validation Gain	Test Gain
0	Dummy Model (All Churn)	1409	371	0	1038	0	0.263307	1.000000	0.500000	0.00	85900	81700
1	Dummy Model (All Not Churn)	1409	0	1038	0	371	NaN	0.000000	0.500000	0.00	-189000	-185500
2	Random Forest (SMOTE)	1409	230	873	165	141	0.582278	0.619946	0.844267	0.09	117500	109200
3	LightGBM (SMOTE)	1409	215	881	157	156	0.577957	0.579515	0.831627	0.01	119100	106000
4	XGBoost (SMOTE)	1409	258	880	158	113	0.620192	0.695418	0.861914	0.14	119600	110900
5	Random Forest (ADASYN)	1409	227	871	167	144	0.576142	0.611860	0.837344	0.14	117900	108200
6	LightGBM (ADASYN)	1409	216	884	154	155	0.583784	0.582210	0.829656	0.01	115300	107900
7	XGBoost (ADASYN)	1409	255	875	163	116	0.610048	0.687332	0.859215	0.12	119800	108700
8	Random Forest (RUS)	1409	319	755	283	52	0.529900	0.859838	0.863704	0.22	122100	109100
9	LightGBM (RUS)	1409	317	738	300	54	0.513776	0.854447	0.863652	0.30	118600	109100
10	XGBoost (RUS)	1409	319	735	303	52	0.512862	0.859838	0.864441	0.36	121100	110800

- Before performing threshold tuning, XGBoost(RUS) seems good because it has a highest AUCROC.
- However, prior to that, XGBoost (SMOTE) is the better model as it achieves maximum gains.
- Ultimately, for the final decision, we choose XGBoost (SMOTE) to deploy.

5. APPENDIX

Threshold selection

	Model	Observation	TP	TN	FP	FN	Precision	Recall	AUCROC	Best Threshold	Validation Gain	Test Gain
0	Dummy Model (All Churn)	1409	371	0	1038	0	0.263307	1.000000	0.500000	0.00	85900	81700
1	Dummy Model (All Not Churn)	1409	0	1038	0	371	NaN	0.000000	0.500000	0.00	-189000	-185500
2	Random Forest (SMOTE)	1409	230	873	165	141	0.582278	0.619946	0.844267	0.09	117500	109200
3	LightGBM (SMOTE)	1409	215	881	157	156	0.577957	0.579515	0.831627	0.01	119100	106000
4	XGBoost (SMOTE)	1409	258	880	158	113	0.620192	0.695418	0.861914	0.14	119600	110900
5	Random Forest (ADASYN)	1409	227	871	167	144	0.576142	0.611860	0.837344	0.14	117900	108200
6	LightGBM (ADASYN)	1409	216	884	154	155	0.583784	0.582210	0.829656	0.01	115300	107900
7	XGBoost (ADASYN)	1409	255	875	163	116	0.610048	0.687332	0.859215	0.12	119800	108700
8	Random Forest (RUS)	1409	319	755	283	52	0.529900	0.859838	0.863704	0.22	122100	109100
9	LightGBM (RUS)	1409	317	738	300	54	0.513776	0.854447	0.863652	0.30	118600	109100
10	XGBoost (RUS)	1409	319	735	303	52	0.512862	0.859838	0.864441	0.36	121100	110800



- Cost of promotion: 100
- Cost of loss: 500
- Savings per customer: 500
- Goal: Find the optimal threshold that results in the highest gain

5. APPENDIX

SHAPley

Features	Sign	Meaning
Contract	+	Customers on contract or top-up plans are more likely to churn, possibly due to the ease of changing numbers for top-up customers.
Internet service	-	Customers without internet in their package are at a higher churn risk, as competitors offer inclusive internet services, reflecting its growing importance.
Tenure	+	Long-time customers have a higher churn probability, potentially due to elderly individuals reducing phone usage to cut expenses.
Payment method (electronic check)	-	Non-users of electronic payment for bills are more likely to churn, possibly due to the inconvenience of payment, prompting a switch to other providers.
Monthly charges	-	Customers paying lower charges are more likely to churn, influenced by industry trends towards appealing, low-priced packages from other operators.