CUSTOMER CHURN IN CREDIT CARD SERVICE

Zhixing Liu, Jiapeng Wang, Zifu Wang, Ge Li, Arthur Hu

Contents

INTRODUCTION

METHODS

OUTCOMES

Topic

- More and more customers leaving their credit card services from a bank
- Figuring out better services to keep these customers

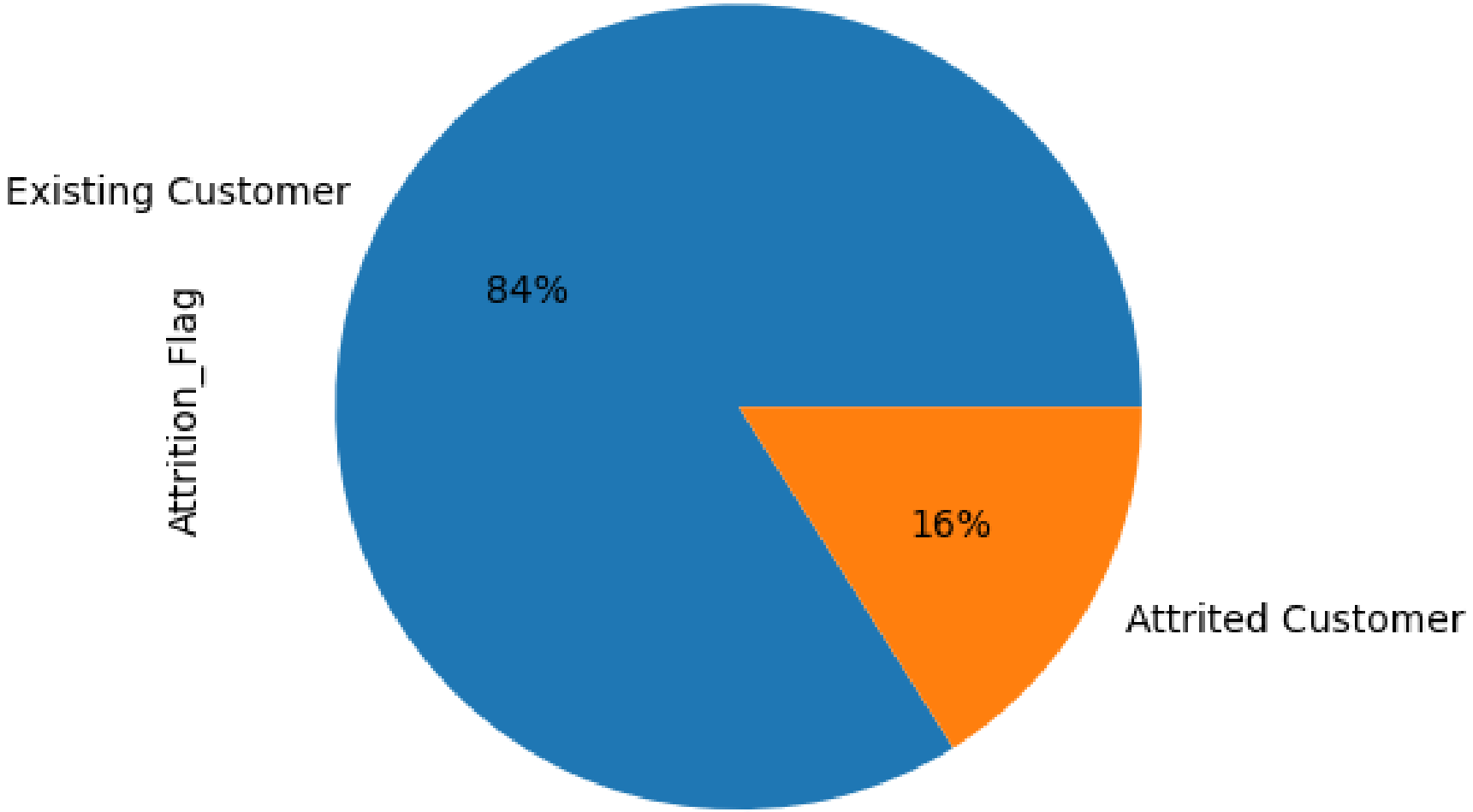
Image source: https://www.cnbc.com/select/how-to-cancel-credit-card

Data

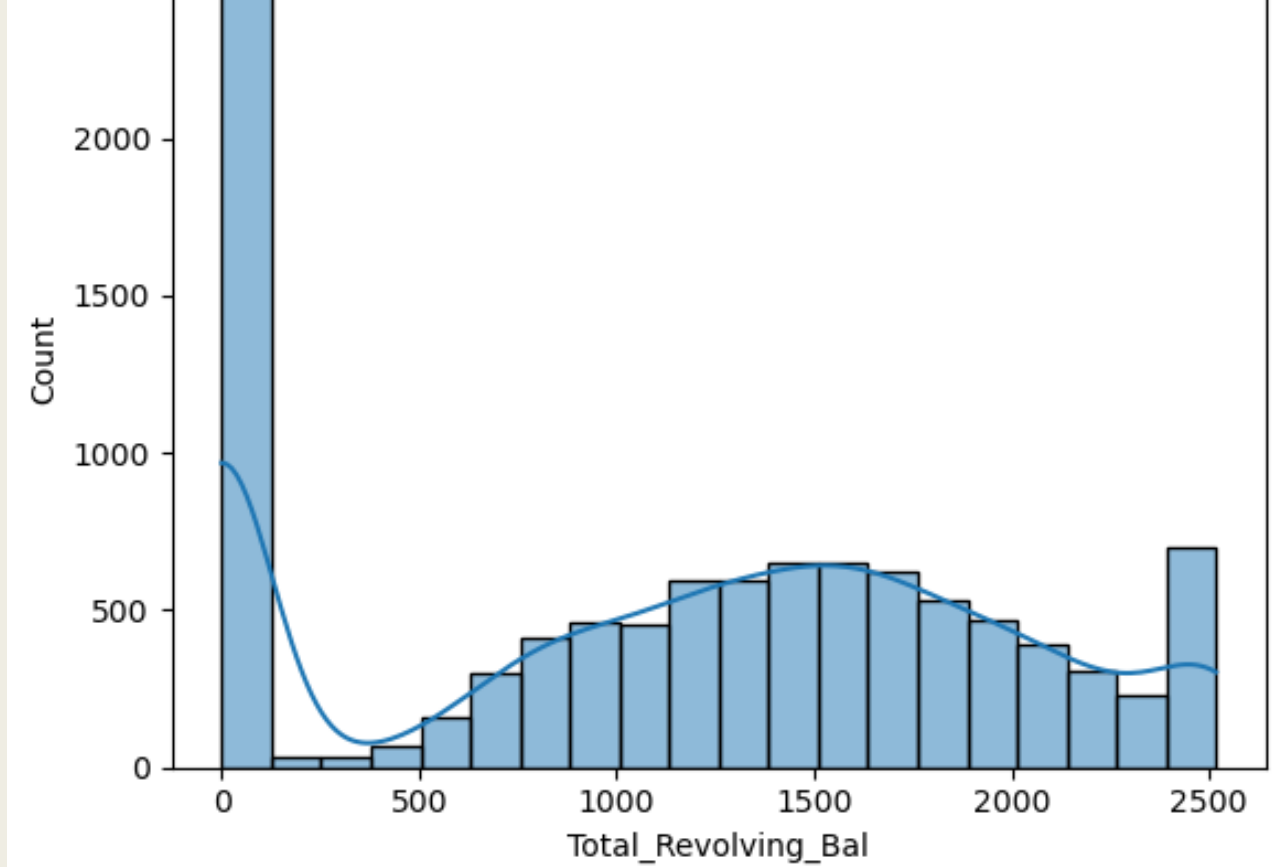
- Source: https://www.kaggle.com/datasets/sakshigoyal7/credit-card-customers

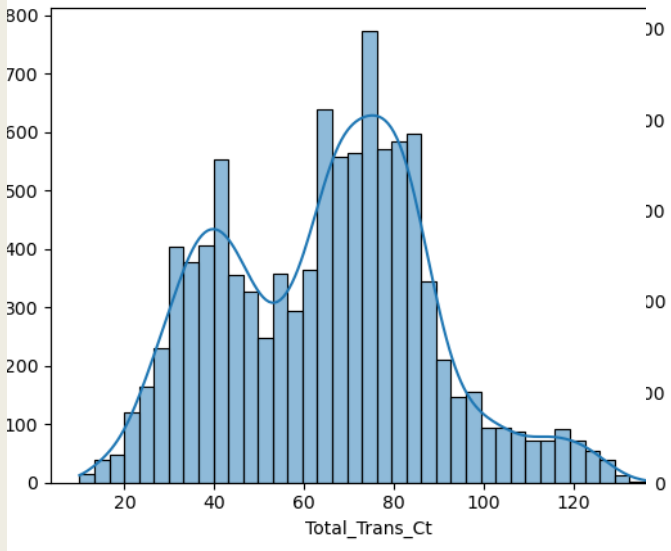
- There are 10,000 customers, with 23 distinct features

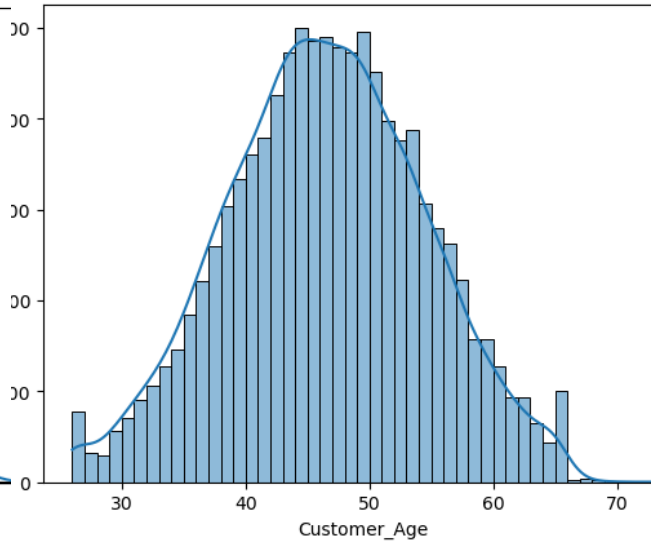
- These features offer a comprehensive view of customer attributes, including **age**, **salary**, **marital status**, **credit card limit**, and **category etc.**

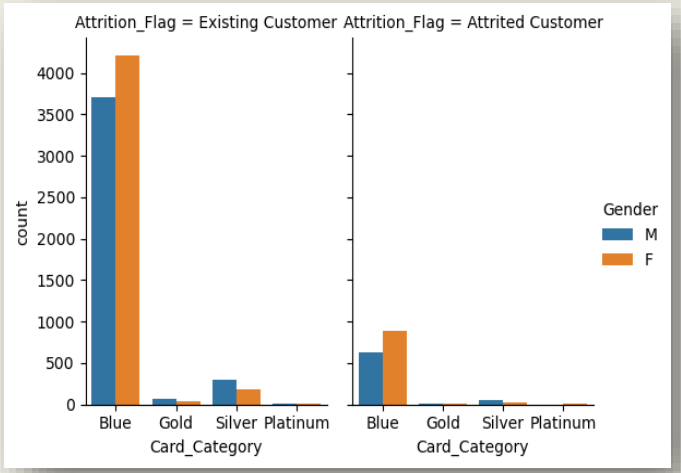

Existing Customer
Attrition_Flag
84%
16%
Attrited Customer

- Total Revolving Balance
- Total Transaction Count(Last 12 months)
- Customer Age

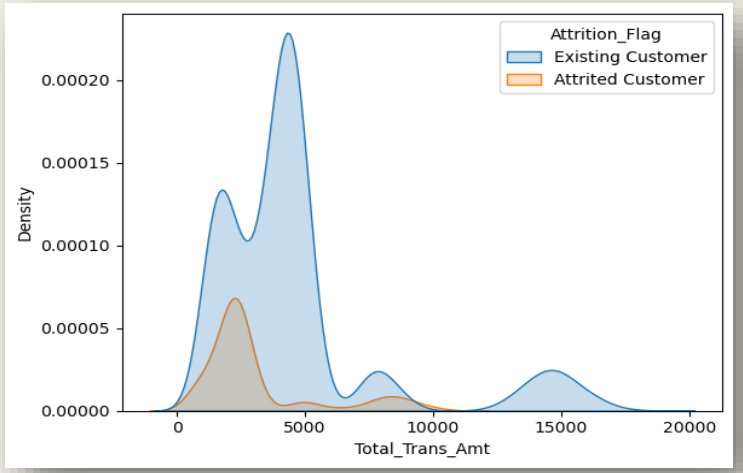




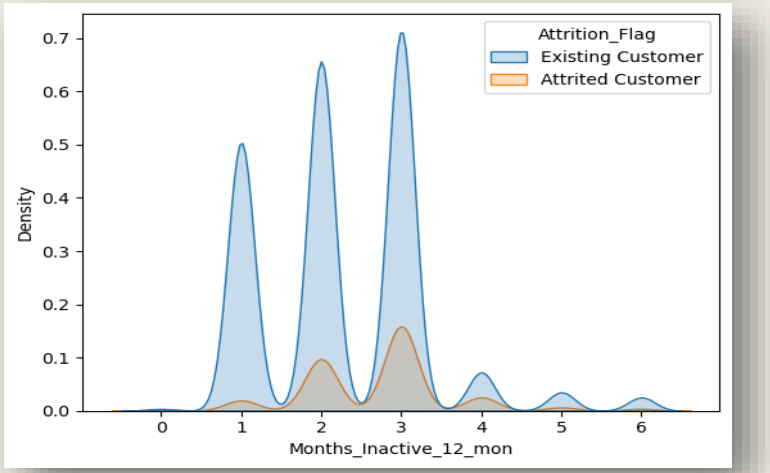




Existing Customer vs. Attrited Customer on Card Category



Existing Customer vs. Attrited Customer on Total Transaction Amount



Existing Customer vs. Attrited Customer on Months Inactive

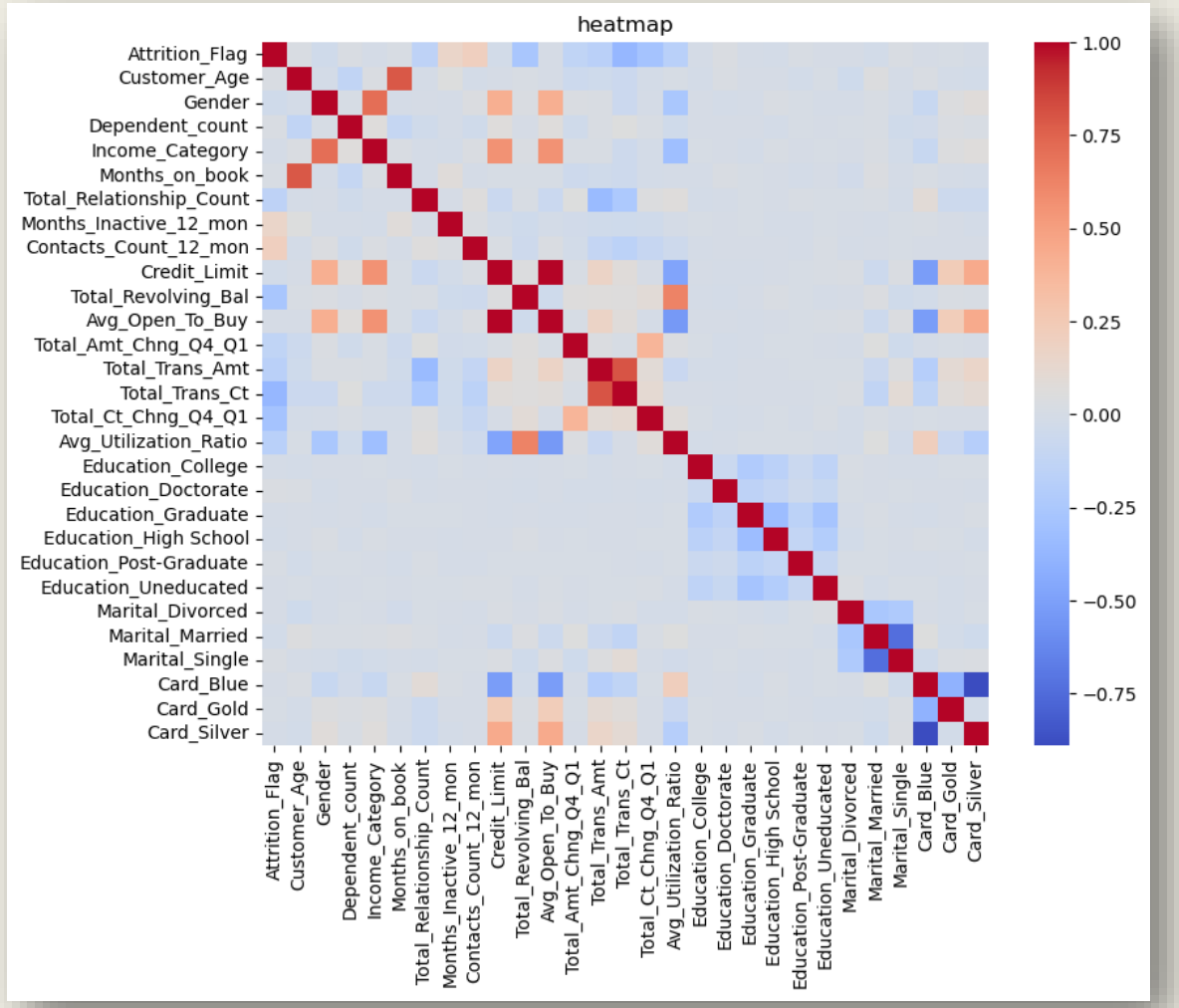
Features

Before processing

	Attrition_Flag	Customer_Age	Gender	Income_Category	Card_Category
0	Existing Customer	45	M	$60K-80K$	Blue
1	Existing Customer	49	F	Less than $40K	Blue
2	Existing Customer	51	M	$80K-120K$	Blue
3	Existing Customer	40	F	Less than $40K	Blue
4	Existing Customer	40	M	$60K-80K$	Blue

After processing

	Attrition_Flag	Customer_Age	Gender	Income_Category	Card_Silver	Card_Gold	Card_Blue
0	0	-0.165406	1	7.0	0	0	1
1	0	0.333570	0	2.0	0	0	1
2	0	0.583058	1	10.0	0	0	1
3	0	-0.789126	0	2.0	0	0	1
4	0	-0.789126	1	7.0	0	0	1



Heat map of processed features

Imbalanced Data Handling

	Accuracy	Recall	Precision
Original Data	0.903	0.602	0.73
Resampled Data	0.844	0.841	0.50

Project Goals

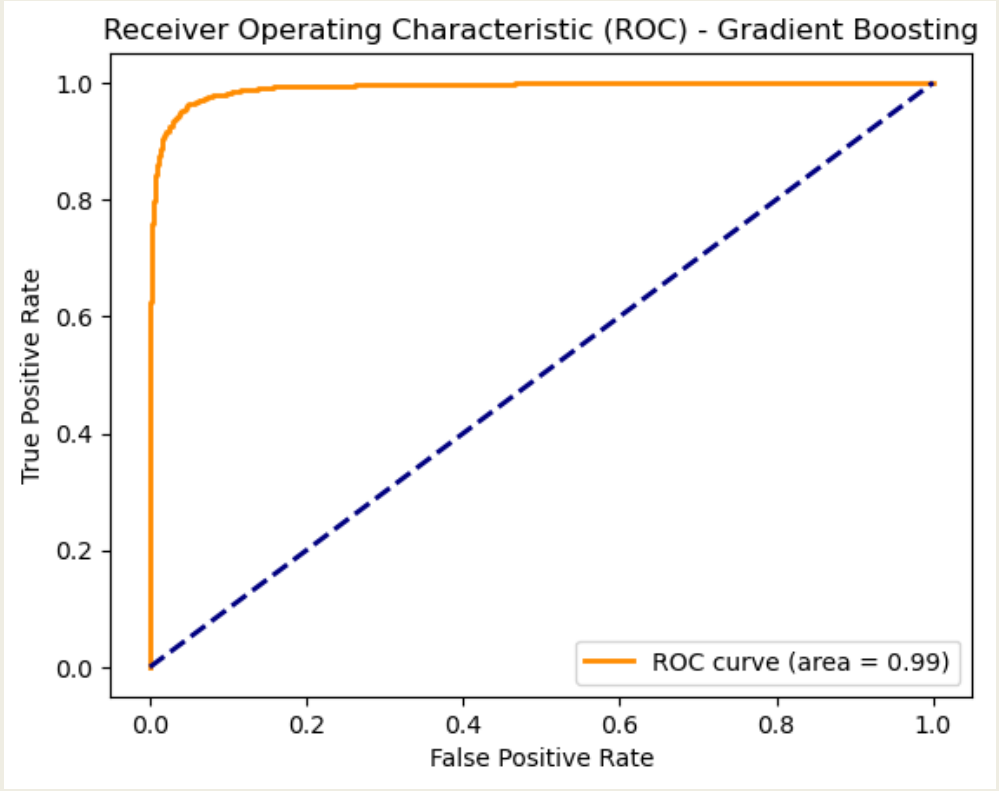
- Develop a model to predict customers at risk of leaving
- Identify the features with the most significant impact on customers churn
- Evaluate the advantages and disadvantages of various feature processing and machine learning methods applied to this dataset

Methods Applied

- kNN
- SVM
- Logistic Regression
- Decision Tree
- Random Forest
- Gradient Boosting
- Stacking

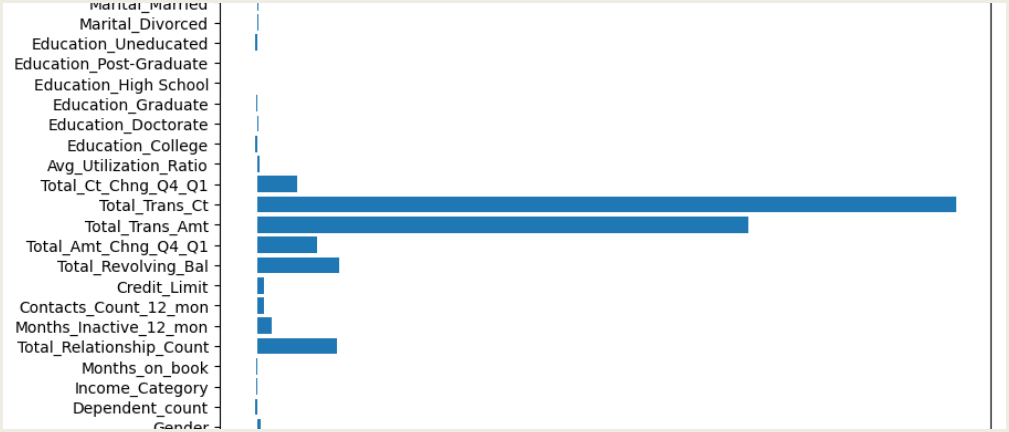
Model Results Comparing

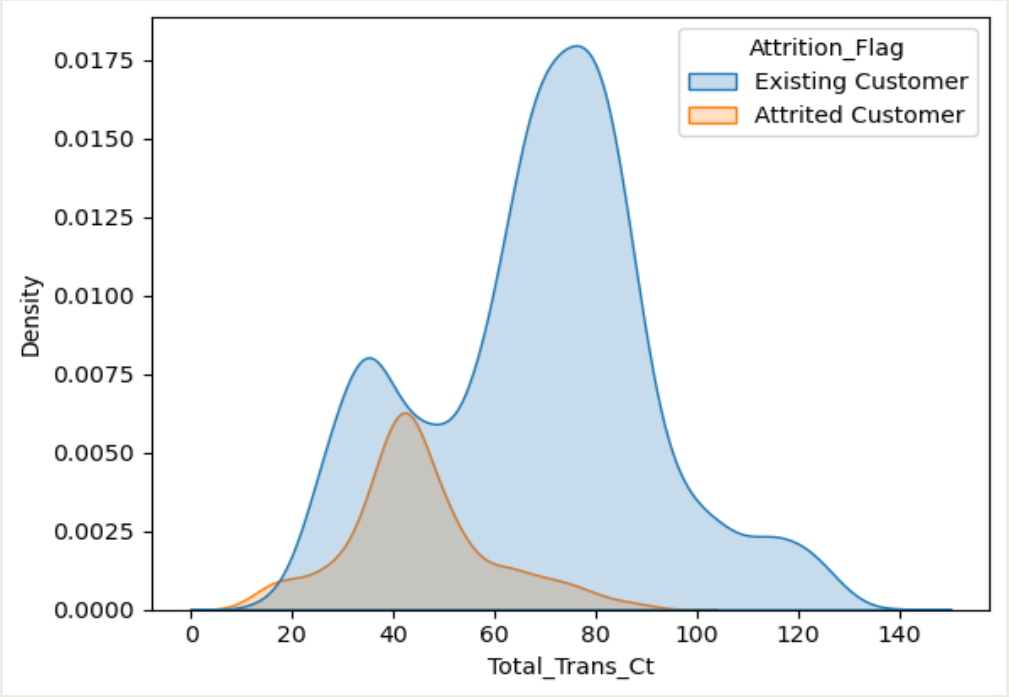
KNN	0.905
SVM	0.933
Logistic Regression	0.903
Decision Tree	0.927
Random Forest	0.959
Gradient Boosting	0.969
Stacking	0.971



Evaluations

- Most significant feature: Total Transaction





THANK YOU