greatlearning
Power Ahead

Capstone Presentation on Customer Churn

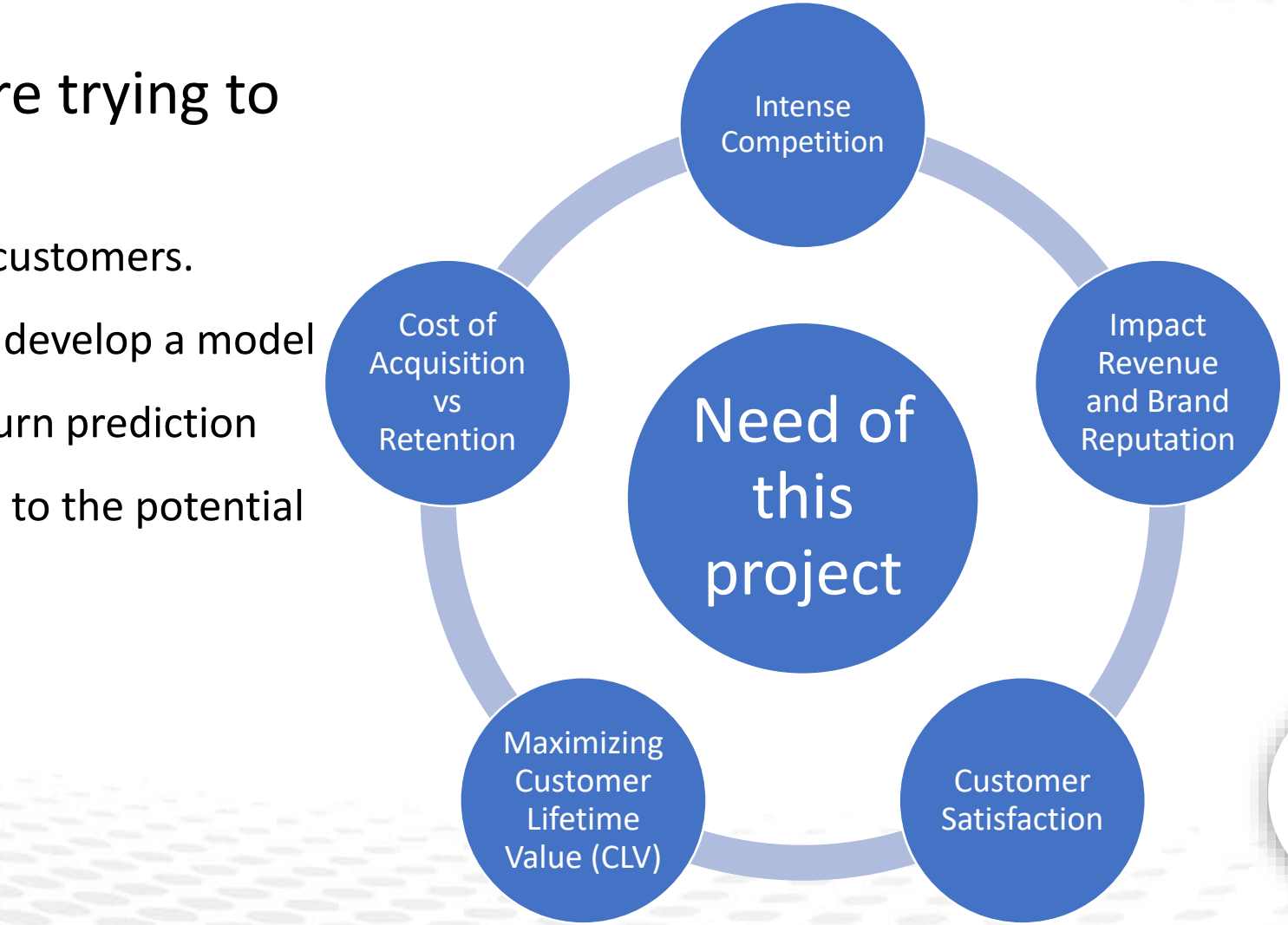
Prepared By :

- **Govind Singh Rawat**
- **PGPDSBA.O.MAY23.A**

Business Problem Understanding

Business problem we are trying to solve:

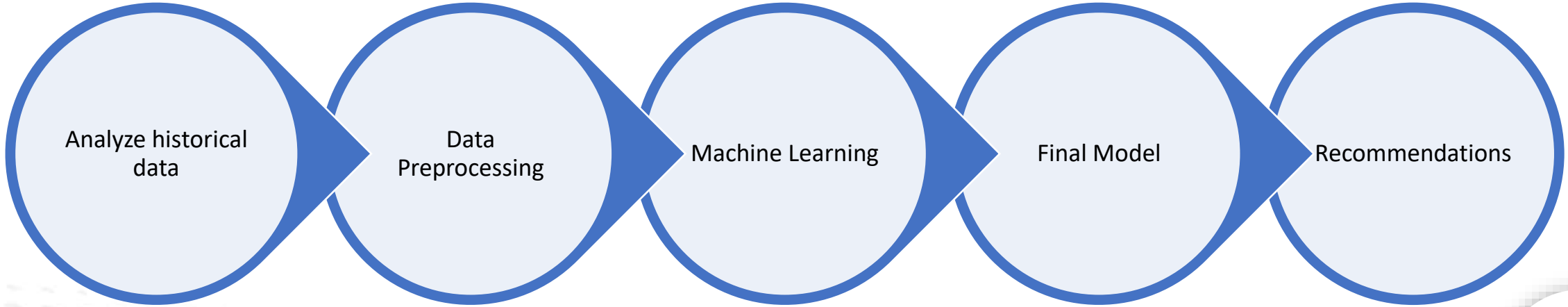
- The challenge is to retain the customers.
- Hence, the company wants to develop a model through which they can do churn prediction
- And provide segmented offers to the potential churners.



Business Problem Understanding

Scope

- Supervised learning classification problem.



The Dataset

- The dataset comprises historical data of customers.
- The time of data collection is not specified.
- There are 11,260 observations in the dataset.
- It consists of 19 columns, including one unique identifier (AccountID).
- Impurities have been detected in the dataset.
- No duplicate observations.

Data Cleaning

Removed Special Characters

- Replaced with null values

Handling Irregular data

- like "Regular +" and "Regular Plus"

Missing Value imputation

- KNN imputer for numeric column
- Mode for categorical column

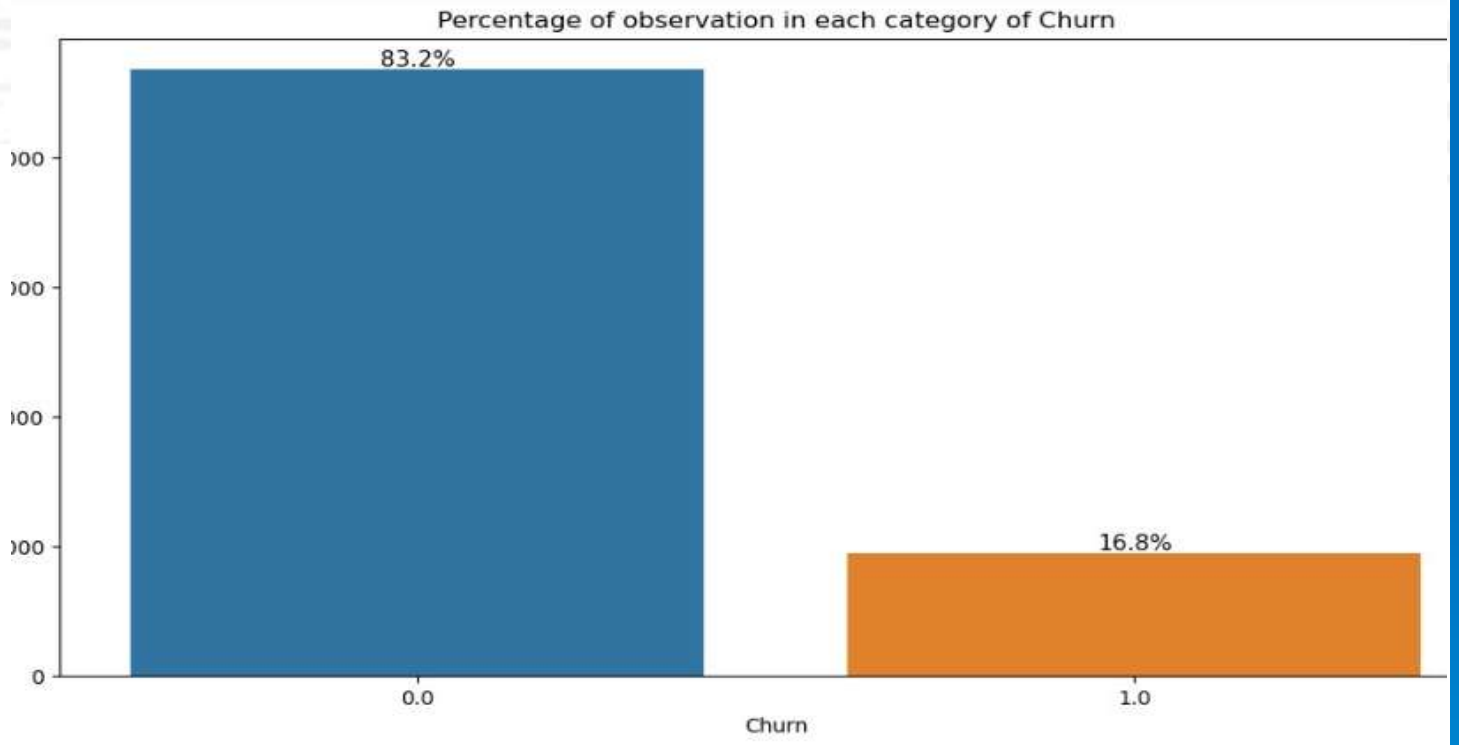
Outlier Treatment

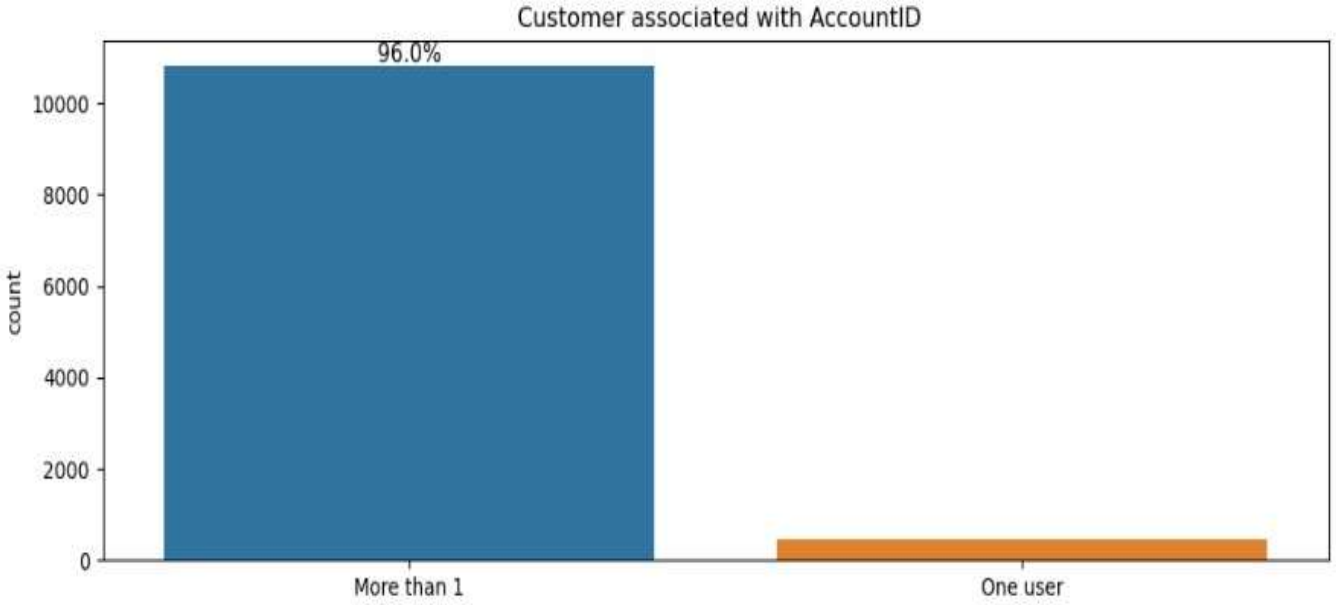
- Clipped into IQR

Exploratory Data Analysis

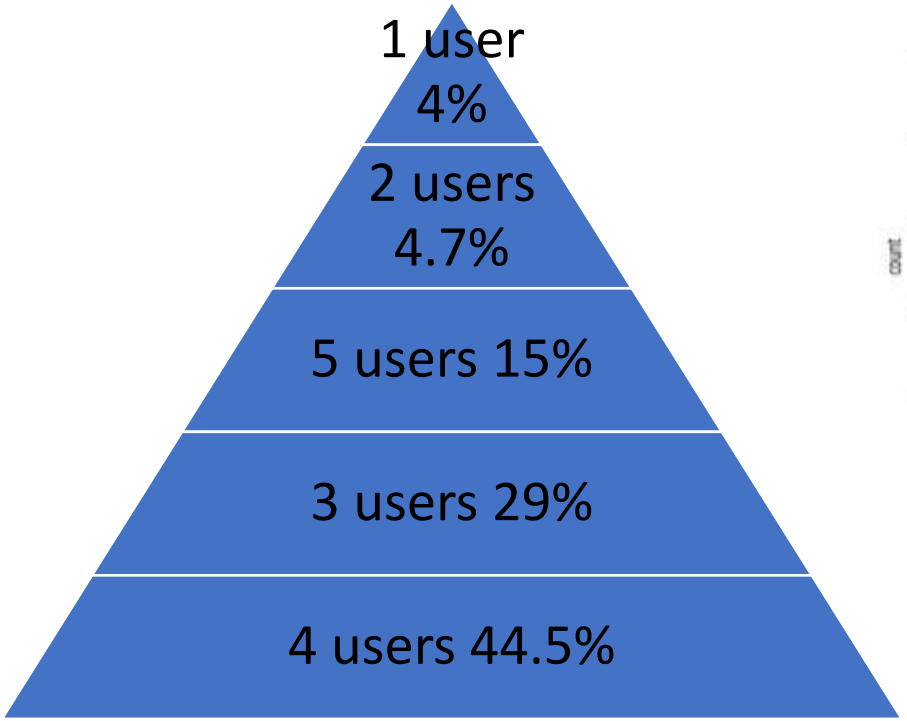
- Only 17% of total customers have churned, indicating an imbalanced dataset.

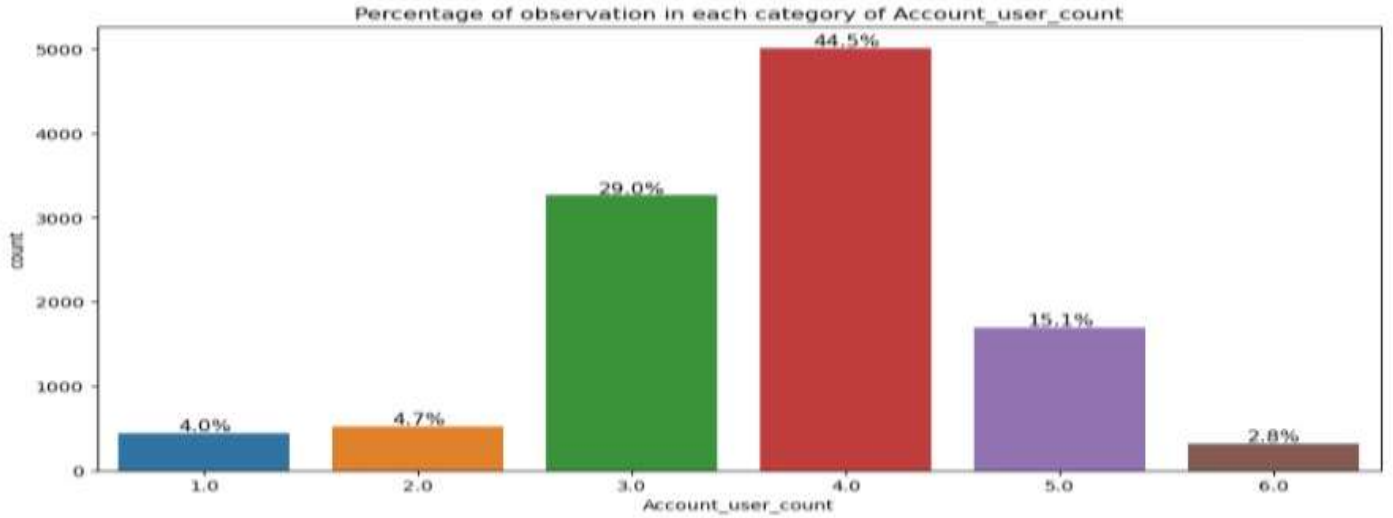
- If one AccountID has churned, there are 96% chances that we lost more than 1 user.

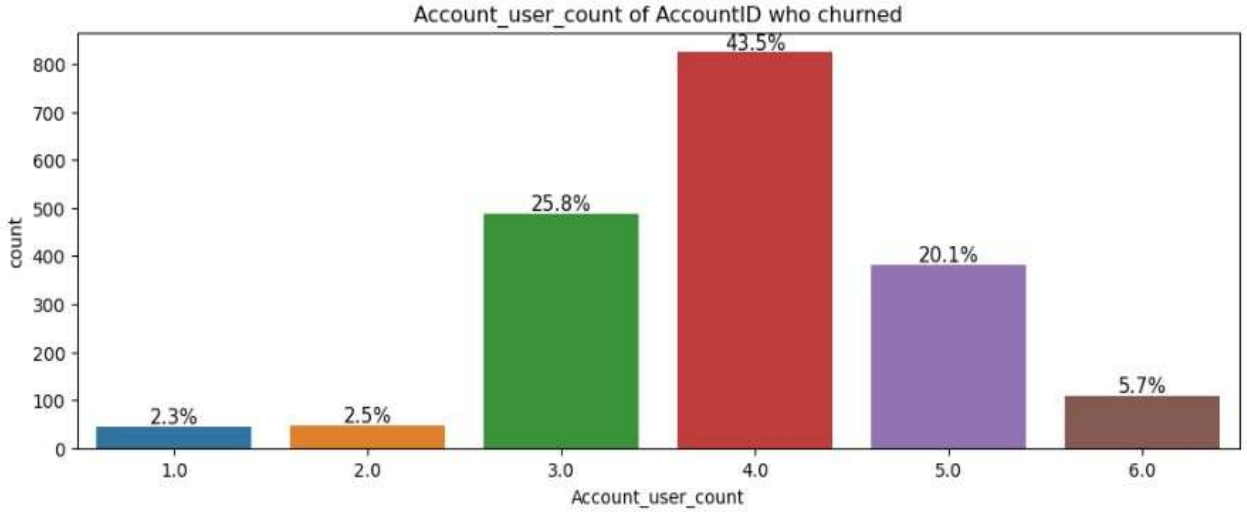




Exploratory Data Analysis



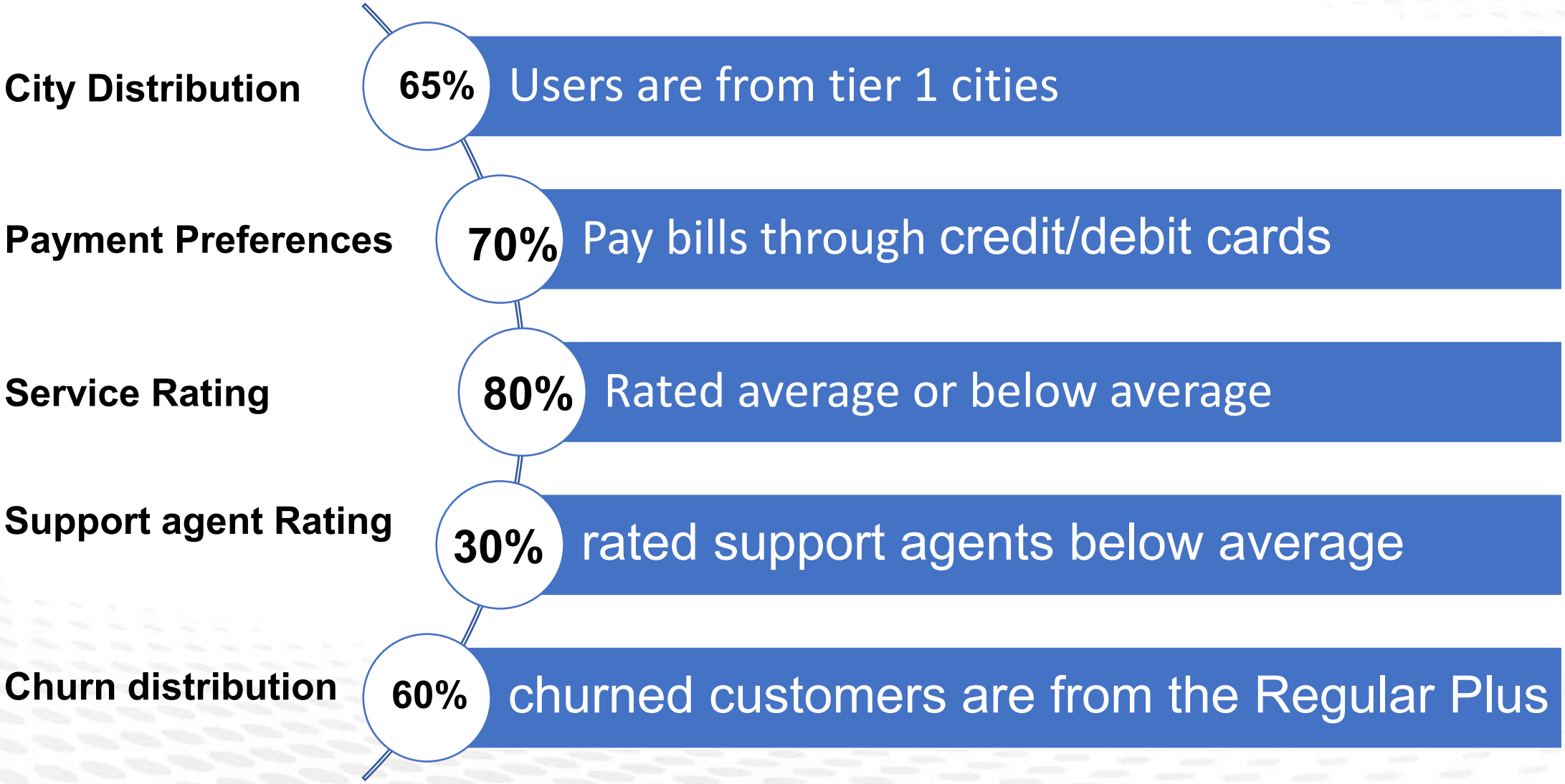




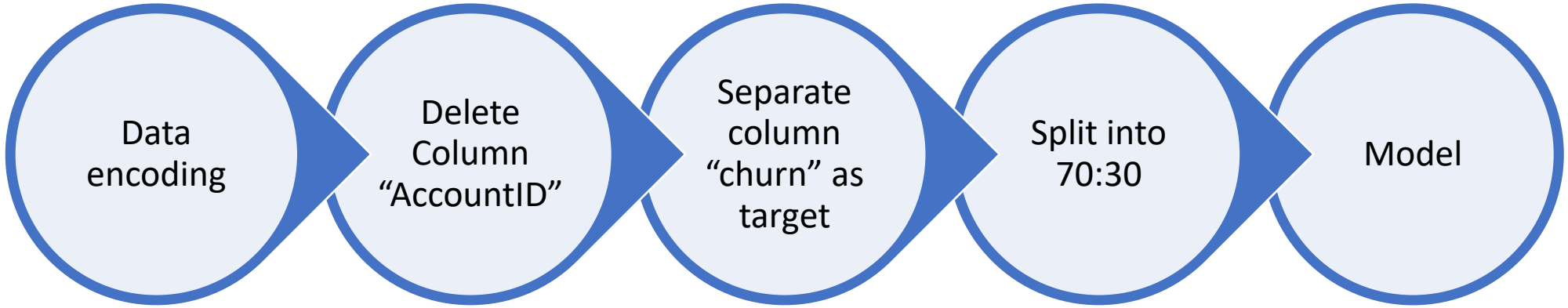
Churned Accounts

- 4 users associated with 43% of the churned accounts.

Exploratory Data Analysis



Modelling – Preprocessing



Modelling Approach

6 Different Models

- Logistic Regression
- Linear Discriminant Analysis
- K Neighbours Classifier
- Decision Tree Classifier
- Random Forest Classifier
- Bagging Classifier

```
LogisticRegression
LogisticRegression(max_iter=10000, n_jobs=2, solver='newton-cg', verbose=True)
```

```
LinearDiscriminantAnalysis
LinearDiscriminantAnalysis()
```

```
KNeighborsClassifier(weights='distance')
```

```
DecisionTreeClassifier
DecisionTreeClassifier(random_state=1)
```

```
RandomForestClassifier
RandomForestClassifier(random_state=1)
```

```
BaggingClassifier
BaggingClassifier(random_state=1)
```

Modelling Results

	Training data(70%)					Test data(30%)				
Model	**Accuracy**	**AUC**	**Precision**	**Recall**	F1	**Accuracy**	**AUC**	**Precision**	**Recall**	**F1**
Logistic Regression	76.18	85.3	40	80	53	88.51	85.6	76	41	58
Logistic Regression(Balanced data)	78.63	86.1	77	82	79	77.8	84.9	76	81	79
LDA	87.99	86.5	77	41	53	87.92	84.6	77	40	53
KNN	100	100	100	100	100	90.08	91.9	76	60	67
Decision Tree	100	100	100	100	100	94.76	91	84	85	85
Random Forest	100	100	100	100	100	97.36	99.3	98	86	92
Bagging Classifier(default parameters)	99.74	100	100	98	99	95.82	97.9	94	80	87
Random Forest(tuned with GridSearch CV)	100	100	100	100	100	100	100	100	100	100

Modelling Results – Final Model

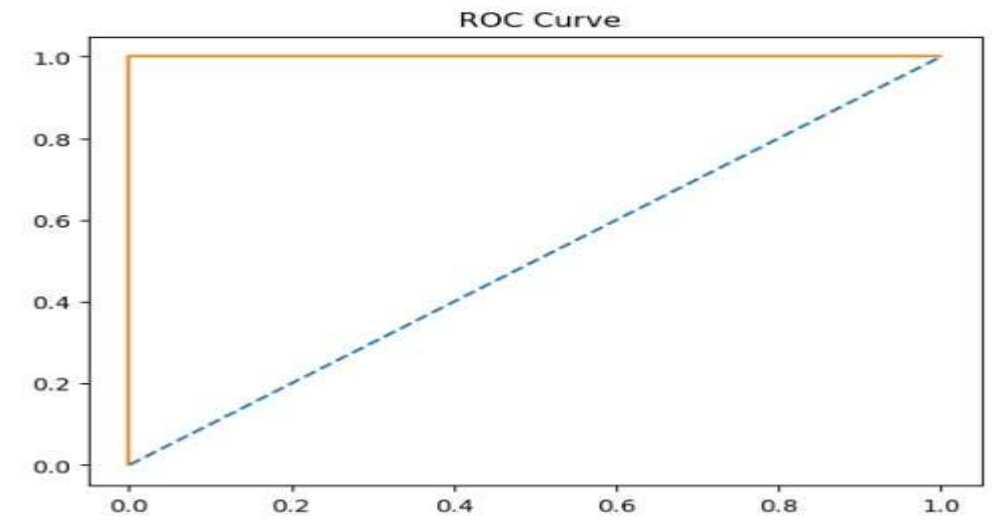
Results on test data

```
GridSearchCV(cv=3, estimator=RandomForestClassifier(random_state=1), n_jobs=-1,
             param_grid={'max_depth': [None, 3, 5, 10],
                         'n_estimators': [10, 20, 25, 50, 100]},
             scoring='f1')
```

estimator: RandomForestClassifier

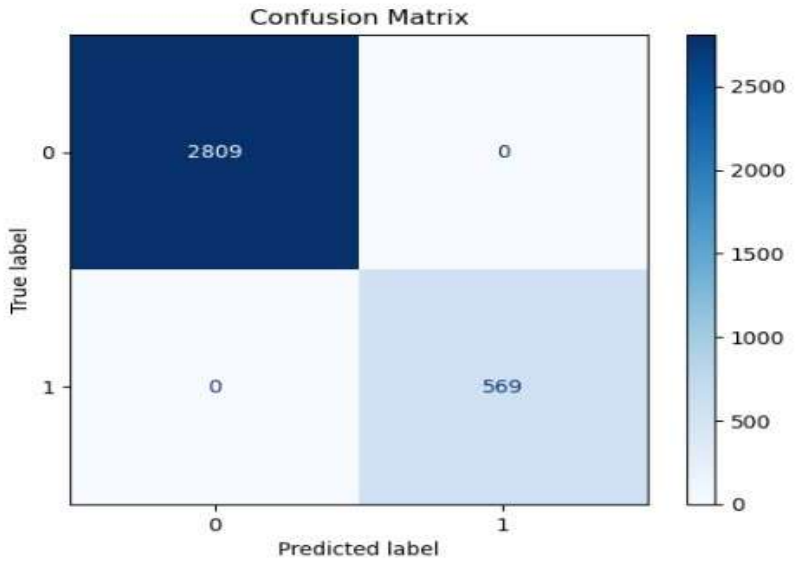
```
RandomForestClassifier(random_state=1)
```

```
AUC: 1.000
ROC curve
```



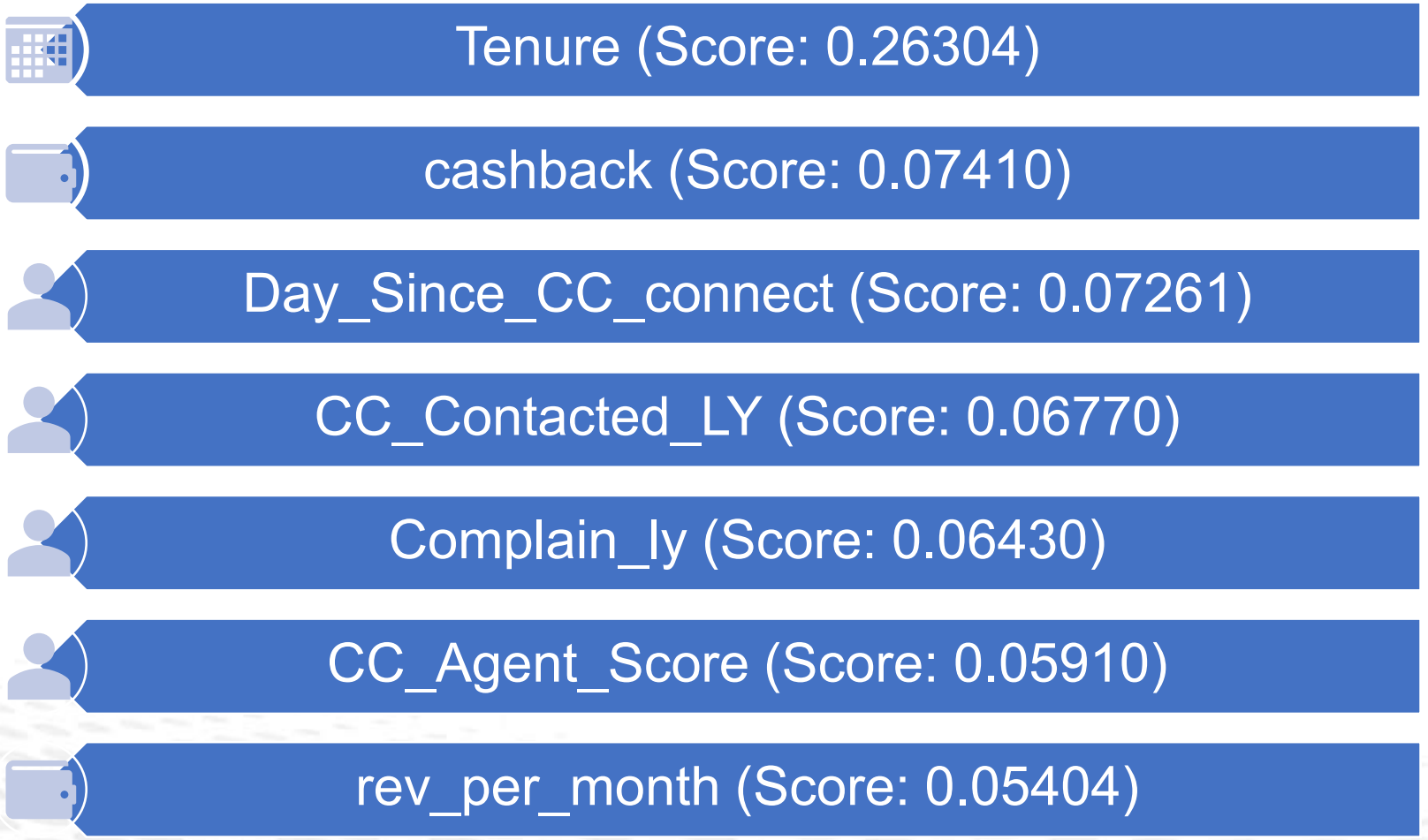

```
Best Model
RandomForestClassifier(random_state=1)
Best parameters
{'max_depth': None, 'n_estimators': 100}
```

Confusion Matrix



Insights from Analysis

Feature Importance



Insights from Analysis

City Distribution	Churn Patterns	Ratings	Segment	Important features
• High concentration in urban areas	• Urban areas experience significant churn	• 30% Rated CC below average • Service ratings are below avg 80% of the time	• Regular Plus segment has high Tendency to churn	• tenure • revenue per month • Cashback • Variables related to customer satisfaction

Recommendations

Enhance Service Quality and Customer Support:

- Monitor and evaluate
- Improve the performance and quality

Segment-Specific Retention Strategies:

- Different needs and preferences

Urban Market:

- Increasing customer engagement in urban areas
- Target customers with personalized offers and incentives

Customer Engagement and Loyalty Programs:

- Prioritize initiatives aimed at enhancing customer engagement and loyalty.
- Design loyalty programs and special offers

Thank You