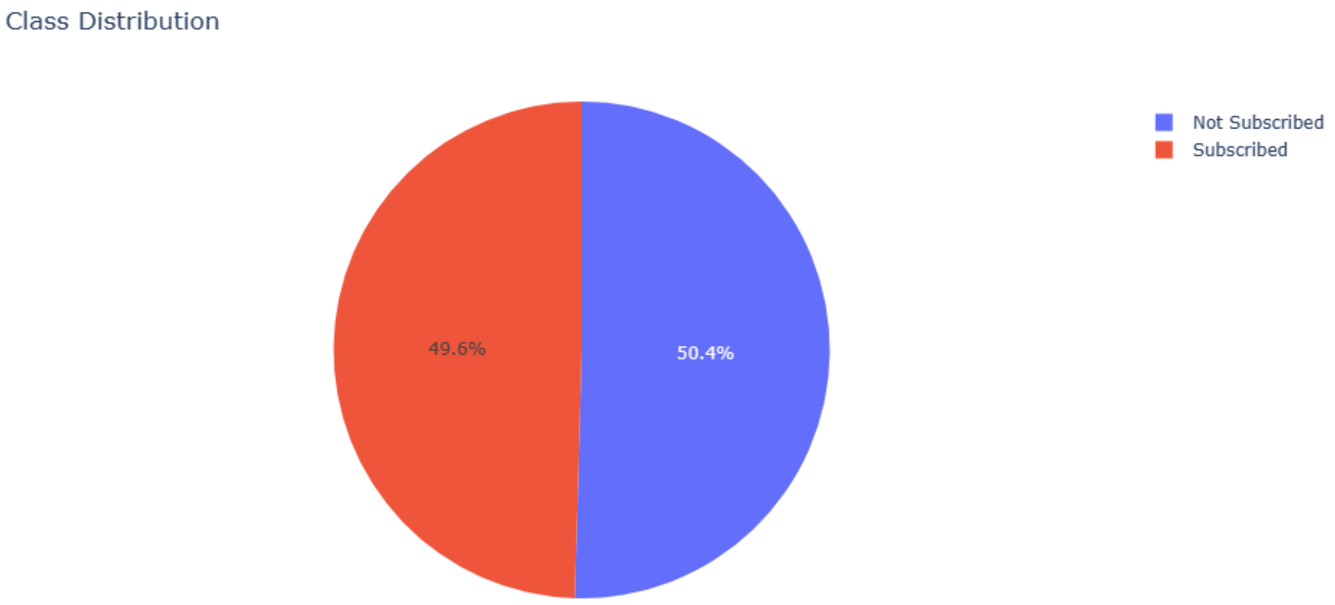
Wireless Industry Subscription Case Study

The problem of machine learning in wireless subscription involves predicting which customers who are likely to cancel their subscription, and taking proactive measures to retain them. This problem is critical for companies, as it directly impacts their revenue and customer base.

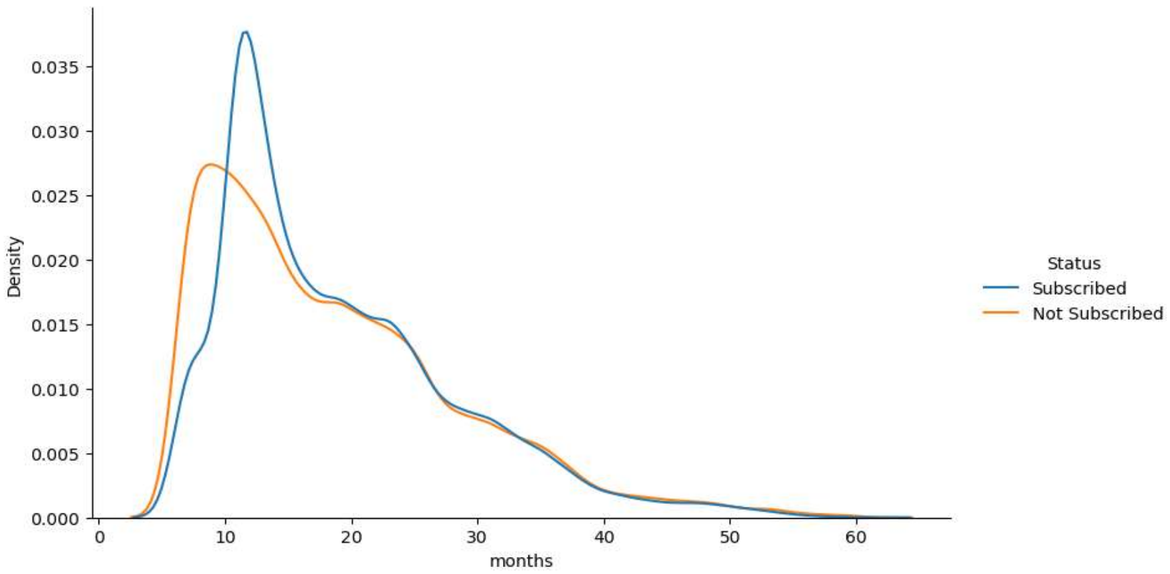
To solve this problem, telecom companies collect data on customer behavior, such as call duration, internet usage, customer demographics, and other variables. This data is used to train machine learning models to identify patterns that are indicative of customers who are likely to cancel their subscription. The models can then be used to predict which customers are at risk of cancellation of subscription, allowing companies to develop targeted retention strategies such as offering discounts or upgrading services to retain these customers.

There are several machine learning algorithms that can be used for this task, including logistic regression, decision trees, random forests, and neural networks. The choice of algorithm depends on the size and complexity of the data set, as well as the desired level of accuracy and interpretability of the results.

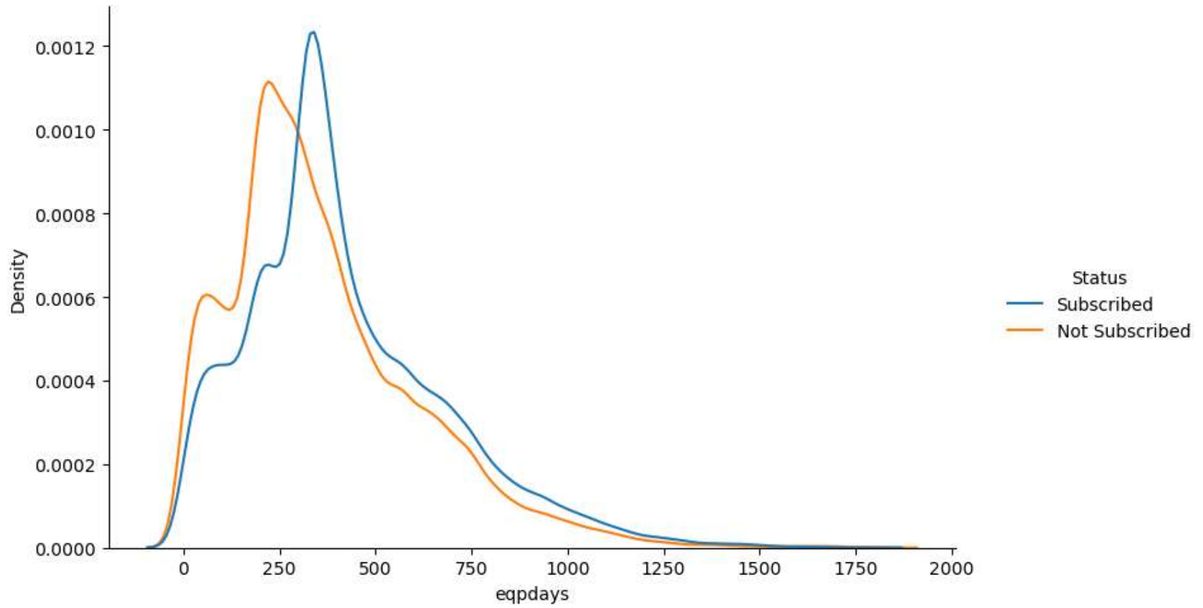


Status	No. of Customers
Not Subscribed	50438
Subscribed	49562
Subscription rate	49.56%

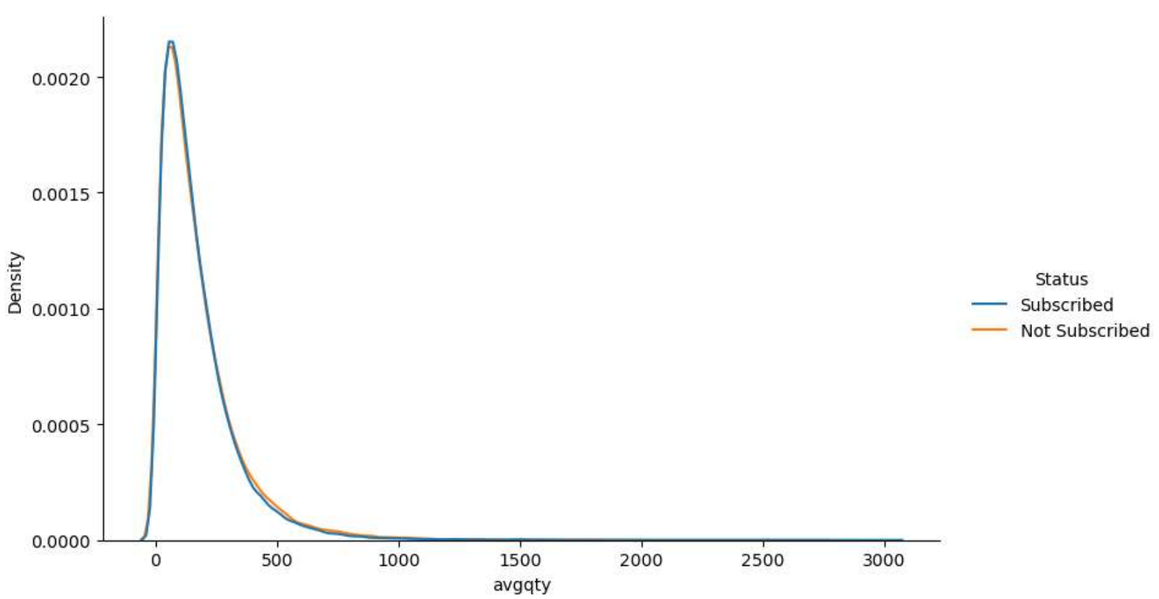
Features Responsible : Auto-ML



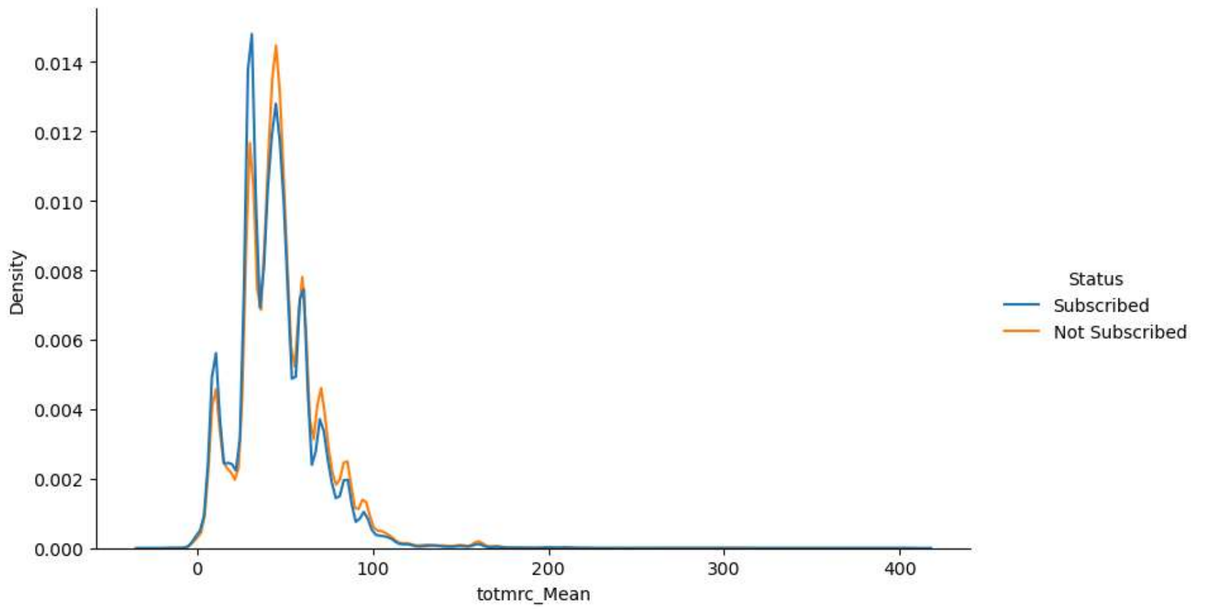
Months: Customers who have subscribed to the service for a longer period of time are more likely to continue their subscription in the future, while customers who have history of frequent un-subscriptions may be less likely to continue their subscription.



Equipment Age (in days): The age of equipment can be an important factor in determining subscription plans and pricing strategies in a telecom service, and understanding customer preferences and behavior can help inform product development and marketing decisions.



Average monthly number of calls over the life of the customer



The mean total monthly recurring charge can be an important factor in determining subscription plans and pricing strategies in a telecom service, and understanding customer preferences and behavior can help inform product development and marketing decisions.

Auto-ML Methodology Results

Algorithms	Test Accuracy (25 percentile)	Test Accuracy (50 percentile)	Test Accuracy (75 percentile)	Test Accuracy (90 percentile)
Decision Tree	55.07	55.28	54.6	55.21
Random Forest	65.6	66.49	66.37	66.73
XGBoost	66.17	66.66	66.74	67.25
MLP	50.1	53.88	55.17	53.53
RNN	50	50	51.51	57.04
Total Features	21	42	63	76
Avg. Accuracy	57.388	58.462	58.878	59.952

- **Based on our observation from the standard ML algorithms, 90 percentile has the best average accuracy**
- **XGBoost was the best performing algorithm at 75th percentile with 66.74% accuracy.**

Conclusion

In Wireless Industry, Auto-ML can help in analyzing customer data, such as past purchases, browsing history, and demographics, to provide personalized recommendations and tailor the customer's experience. This can help keep customers engaged and more likely to continue their subscription. The dataset contains 100,000 records with 14 Categorical feature and 63 Numerical features.

For classification, models were created with algorithms using Auto-ML techniques like Decision tree, Random forest, XGBoost, Multilayer Perceptron and Recurrent Neural Network. With these models, performance measurement values were obtained for feature sets of 21, 42, 63 and 76. The Auto-ML algorithms were able to predict whether the customer will subscribe or not with an average accuracy between 56% – 60% and helped to identify factors that determine the Subscription Rate.

The major factors include Months, Equipment Age, Average monthly number of calls and The mean total monthly recurring charge. In conclusion, AI can be a powerful tool for subscription businesses looking to reduce churn and improve retention.

Sensitivity Analysis

Actual Data Values

EQPDAYS	OVRMOU_MEAN	AVGMOU	AVGQTY	OVRREV_MEAN	STATUS
700.0	419.5	810.25	312.81	41.95	0
51 %	-26 %	-44 %	-32 %	25 %	

Status -0 : Not Subscribed

Adjusted Data Values

EQPDAYS	OVRMOU_MEAN	AVGMOU	AVGQTY	OVRREV_MEAN	STATUS
984.07 ↑ (284.07)	419.5 ↑ (0.0)	722.31 ↓ (-87.94)	276.82 ↓ (-35.99)	41.95 ↑ (0.0)	1

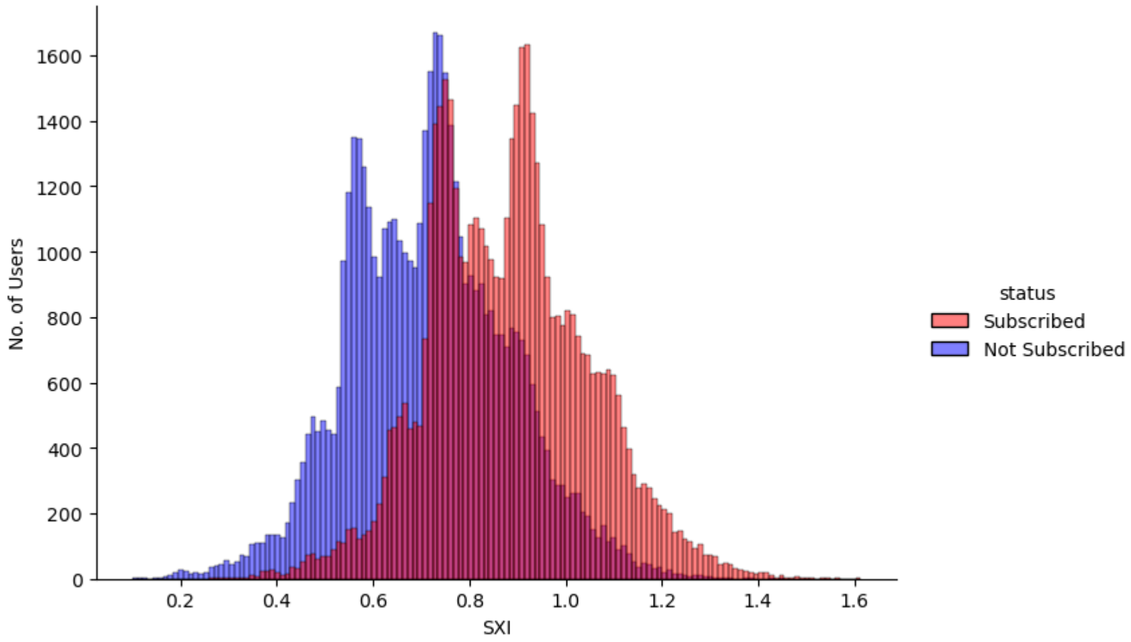
Status -1 : Subscribed

- The top 5 features :
  - Eqpdays : Age of the equipment in days.
  - Ovrmou_mean: Mean overage minutes (time spent for the extra charges after completion of a plan) of use.
  - Avgmou: Average monthly minutes of use over the life of the customer.
  - Avgqty: Average monthly number of calls over the life of the customer.
  - Ovrrev_mean: Revenue from the extra usages.

  For this user we can see:
  - If the age of the equipment with 51% increase from the actual .
  - If the average monthly minutes of use with 44 % decrease from the actual .
  - If the average monthly number of calls with 32% decrease from the actual.

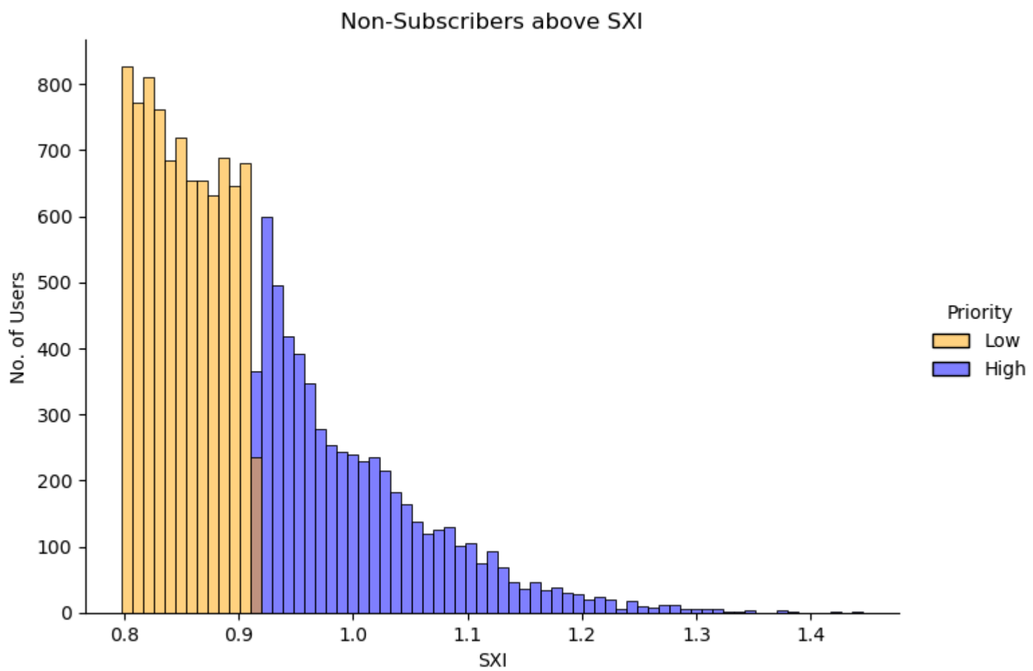
  Which will lead this user to be converted. (will make a subscription)

SXI Method



SXI	0.8
Top SXI	1.61
Minimum SXI	0.1
No. of Subscribers above SXI	32979
No. of Subscribers below SXI	16583
No. of non-Subscribers above SXI	14792
No. of non-Subscribers below SXI	35646
SXI Model Accuracy	98%

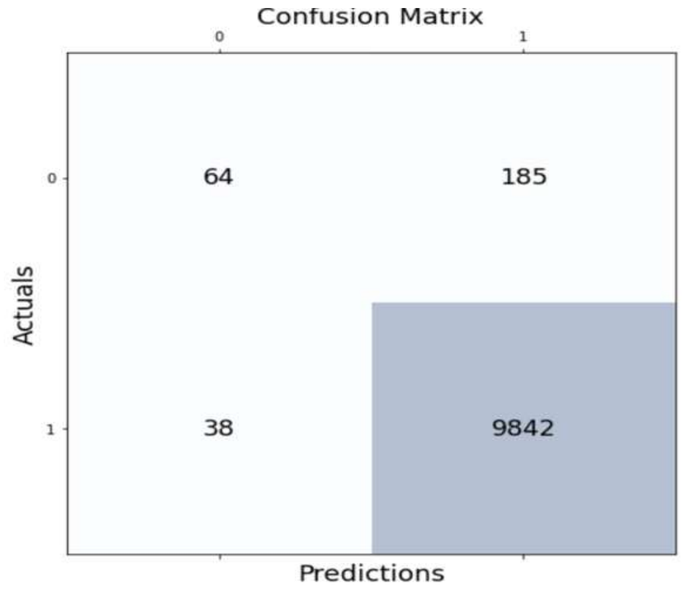
- Brown – Top Customers
- Maroon– Mix of Loosing Customers, Prospective Users and Non-Subscribers.
- Purple - Non-Subscribers



Non –Subscribers above SXI

- Users who have above SXI values and who are not subscribed are the prospective customers.
- In order to increase the subscription rate these users are one to focus on, because in future they have more chance to subscribe.
- The average SXI on “Non-Subscribers who are above SXI” is 0.91. Here we further categorize ,so user’s SXI above 0.91 are to be mostly retargeted in future.
- Analyzing the trend and behavior patterns of these users will improve customer acquisition, retention, and revenue growth over time.

SXI Method Accuracy is 98%



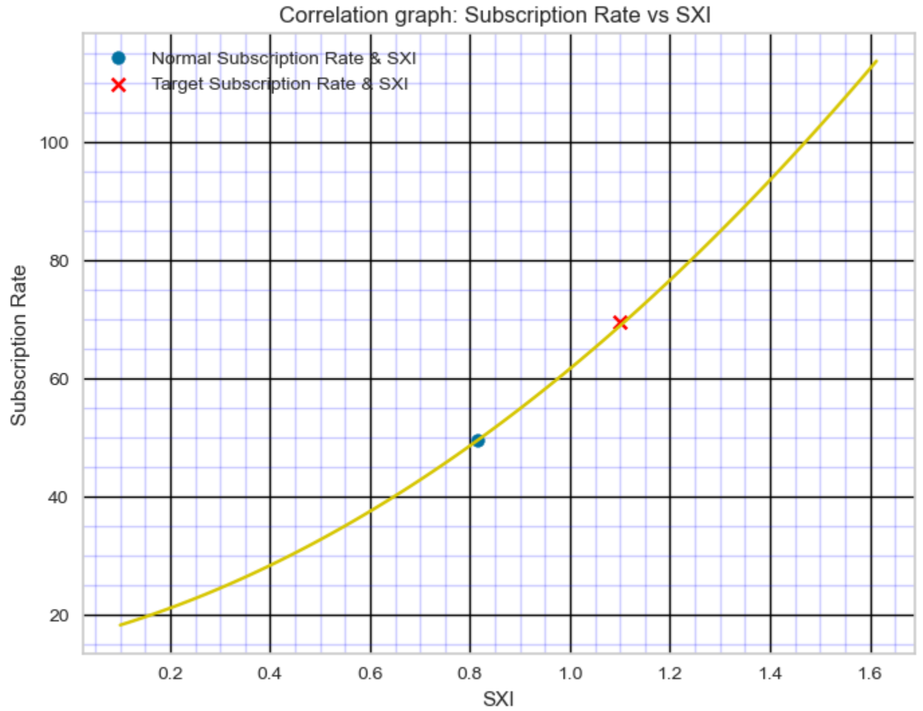
- Actual and Predicted was Subscribed (TP) : 9842
- Actual and Predicted was Not Subscribed (TN) : 64
- Actual Subscribed and Predicted Not Subscribed (FN): 38
- Actual Not Subscribed Predicted Subscribed (FP) : 185

Train Records	Test Records	Actual Train count for subscribed	Actual train count for not subscribed	Actual test count for subscribed	Actual test count for not subscribed	Predicted test count subscribed	Predicted test count not subscribed	Precision rate	Recall rate	Model Accuracy
40514	10129	39561	953	9880	249	10027	102	0.98	0.99	**0.98**

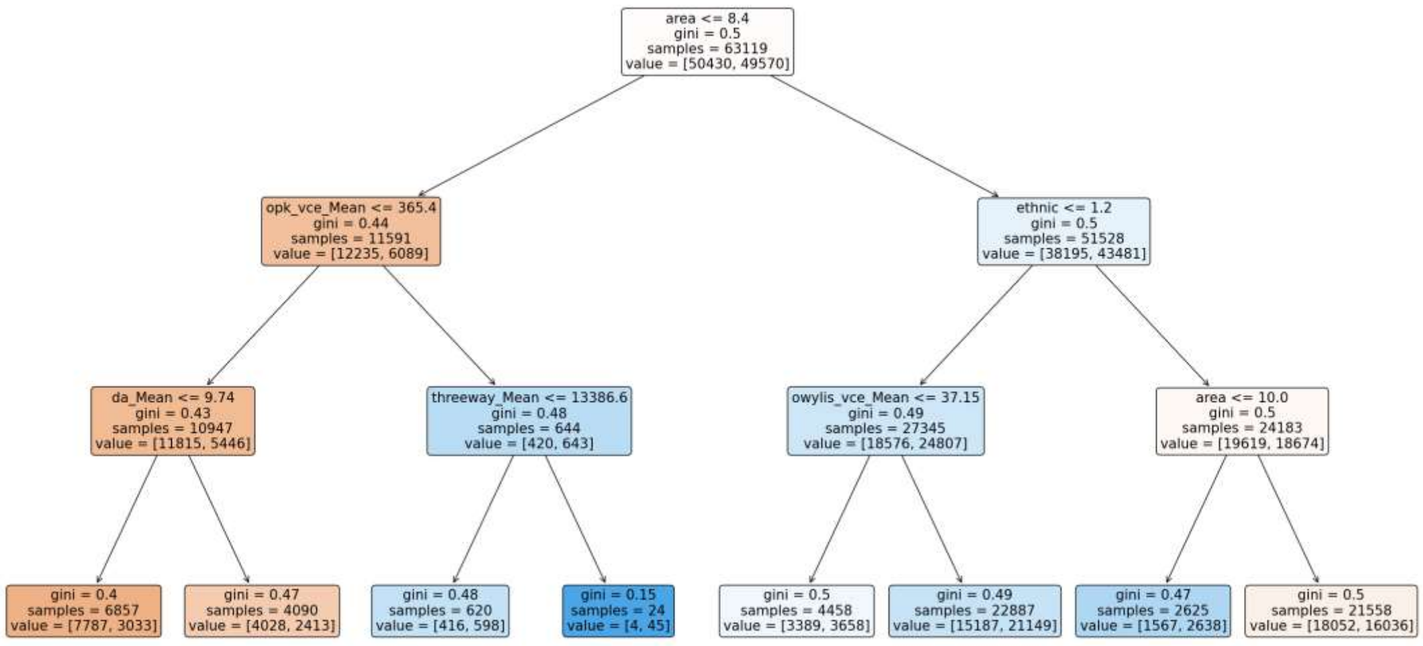
- **Precision rate**: Precision is defined as the ratio of actual Subscribed customers (True Positive) to a total number of predicted subscribed customers. ***TP / (TP + FP)***
- **Recall rate**: The recall is calculated as the ratio between the Actual numbers of subscribed customers to the total number of wrongly predicted subscribed customers as not subscribed customers plus actual number of subscribed customers. ***TP / (TP + FN)***
- **Model Accuracy**: It is the fraction of predictions where the model got right. ***(TP + TN) / (TP + FP + TN + FN)***

SXI Method - Conclusion

SXI	0.8
Target SXI	1.1
Subscription rate	49.56%
Target Subscription rate	69.56%



The correlation between SXI and Subscription rate is ***0.99***



Tree - Interpretation

- **Area =** Chicago, Great Lakes, Dallas**,**Tennessee, Philadelphia.
- **Mean number of off-peak voice calls <** 365
- **Mean number of three way calls >** 13386
- 24 samples leading to higher subscription rate.