10Alytics
data & strategy

Machine-Learning-for-Churn-Detection

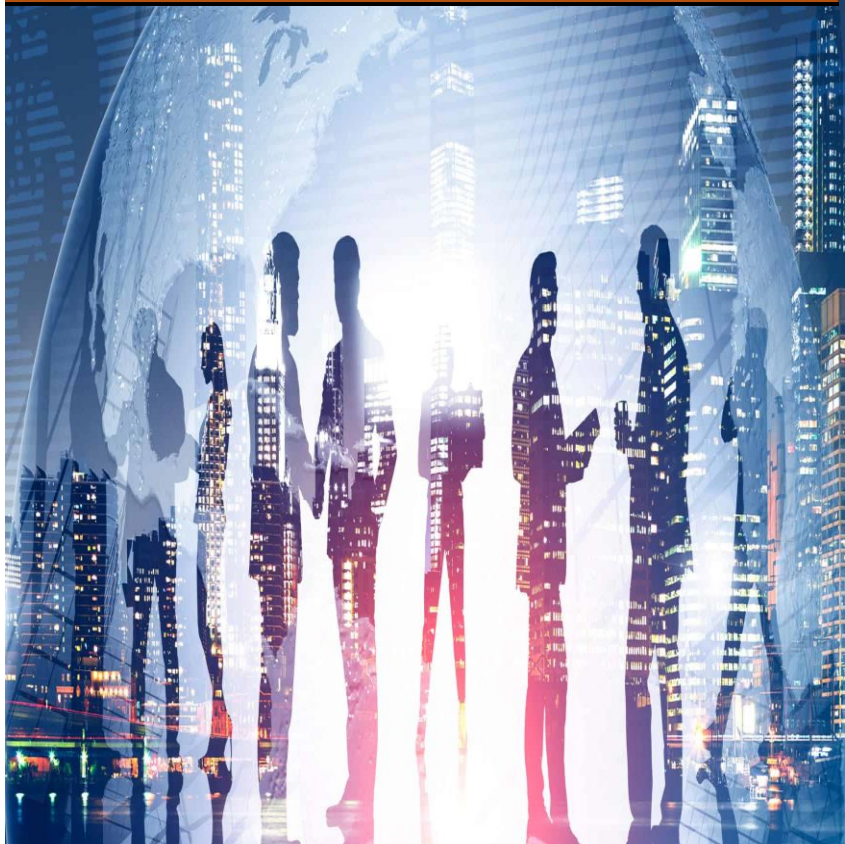
CONNECTTEL

BACKGROUND

CONNECTTEL
CUSTOMER CHURN PREDICTION

ConnectTel stands as a pioneering telecommunications enterprise, leading the way in innovative connectivity solutions and maintaining a robust global market presence. Recognized as a dependable provider of voice, data, and Internet services, ConnectTel offers a diverse array of telecommunications solutions. This includes mobile networks, broadband connections, and enterprise solutions, catering to both individual and corporate clientele.

Distinguished by an unwavering commitment to exceptional customer service and cutting-edge technology, ConnectTel ensures seamless communication experiences for millions of users across the globe. Through strategic partnerships and a customer-centric approach, ConnectTel continues to drive transformative change in the telecom industry, empowering individuals and businesses alike to stay seamlessly connected.
and thrive in the digital.



GOALS AND PROBLEM DEFINITION

ConnectTel Telecom Company is confronted with the urgent task of tackling customer churn, a substantial threat to its business sustainability and expansion. The existing customer retention strategies of the company lack accuracy and efficacy, leading to the defection of valuable customers to competitors.

DATA DICTIONARY

1. CustomerID:	A unique identifier assigned to each telecom customer, enabling tracking and identification of individual customers.
2. Gender:	The gender of the customer, which can be categorized as male, or female. This information helps in analyzing gender-based trends in customer churn.
3. SeniorCitizen:	A binary indicator that identifies whether the customer is a senior citizen or not. This attribute helps in understanding if there are any specific churn patterns among senior customers.
4. Partner:	Indicates whether the customer has a partner or not. This attribute helps in evaluating the impact of having a partner on churn behavior.
5. Dependents:	Indicates whether the customer has dependents or not. This attribute helps in assessing the influence of having dependents on customer churn.
6. Tenure:	The duration for which the customer has been subscribed to the telecom service. It represents the loyalty or longevity of the customer's relationship with the company and is a significant predictor of churn.
7. PhoneService:	Indicates whether the customer has a phone service or not. This attribute helps in understanding the impact of phone service on churn.

DATA DICTIONARY

8. MultipleLines: Indicates whether the customer has multiple lines or not. This attribute helps in analyzing the effect of having multiple lines on customer churn.

9. InternetService: Indicates the type of internet service subscribed by the customer, such as DSL, fiber optic, or no internet service. It helps in evaluating the relationship between internet service and churn.

10. OnlineSecurity: Indicates whether the customer has online security services or not. This attribute helps in analyzing the impact of online security on customer churn.

11. OnlineBackup: Indicates whether the customer has online backup services or not. This attribute helps in evaluating the impact of online backup on churn behavior.

12. DeviceProtection: Indicates whether the customer has device protection services or not. This attribute helps in understanding the influence of device protection on churn.

13. TechSupport: Indicates whether the customer has technical support services or not. This attribute helps in assessing the impact of tech support on churn behavior.

14. StreamingTV: Indicates whether the customer has streaming TV services or not. This attribute helps in evaluating the impact of streaming TV on customer churn.

DATA DICTIONARY

15. StreamingMovies: Indicates whether the customer has streaming movie services or not. This attribute helps in understanding the influence of streaming movies on churn behavior.

16. Contract: Indicates the type of contract the customer has, such as a month-to-month, one-year, or two-year contract. It is a crucial factor in predicting churn as different contract lengths may have varying impacts on customer loyalty.

17. PaperlessBilling: Indicates whether the customer has opted for paperless billing or not. This attribute helps in analyzing the effect of paperless billing on customer churn.

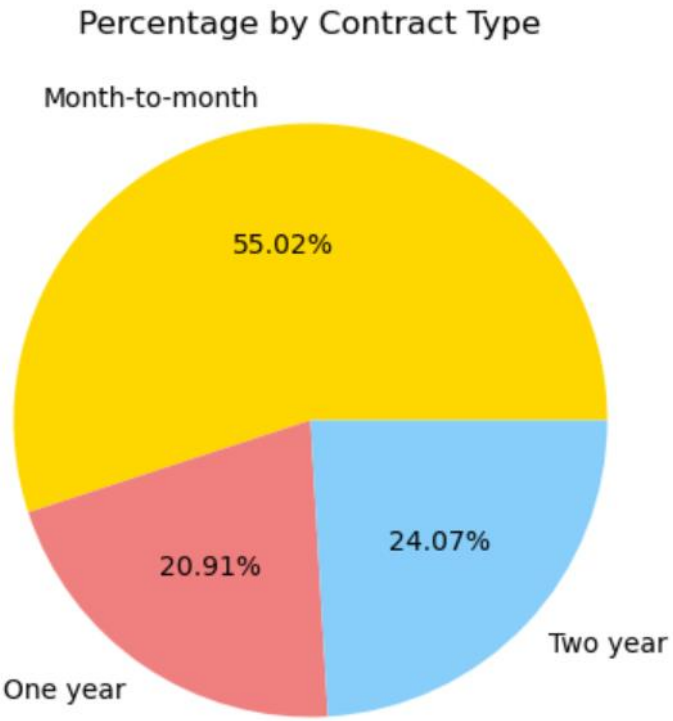
18. PaymentMethod: Indicates the method of payment used by the customer, such as electronic checks, mailed checks, bank transfers, or credit cards. This attribute helps in evaluating the impact of payment methods on churn.

19. MonthlyCharges: The amount charged to the customer on a monthly basis. It helps in understanding the relationship between monthly charges and churn behavior.

20. TotalCharges: The total amount charged to the customer over the entire tenure. It represents the cumulative revenue generated from the customer and may have an impact on churn.

21. Churn: The target variable indicates whether the customer has churned (canceled the service) or not. It is the main variable to predict in telecom customer churn analysis.

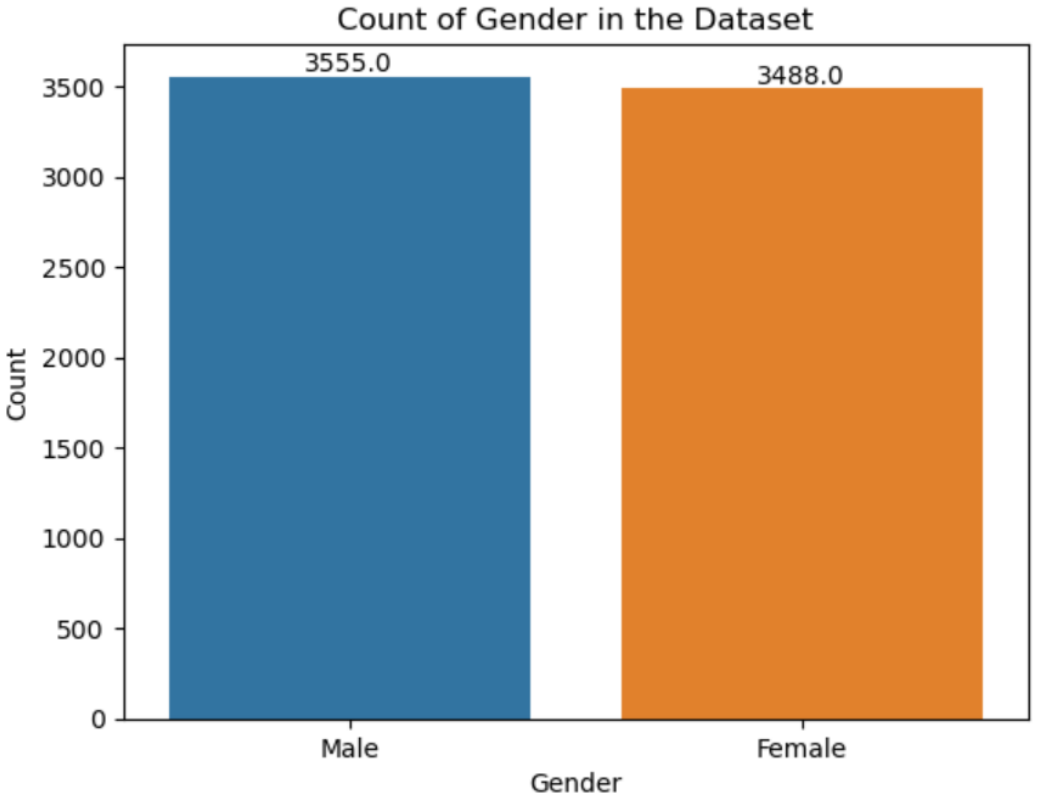
Visualization of key features



Observations:

ConnectTel Telecom Company faces an immediate challenge in addressing customer churn, a significant threat to its business sustainability and growth. The company's current customer retention strategies lack precision and effectiveness, resulting in the loss of valuable customers to competitors.

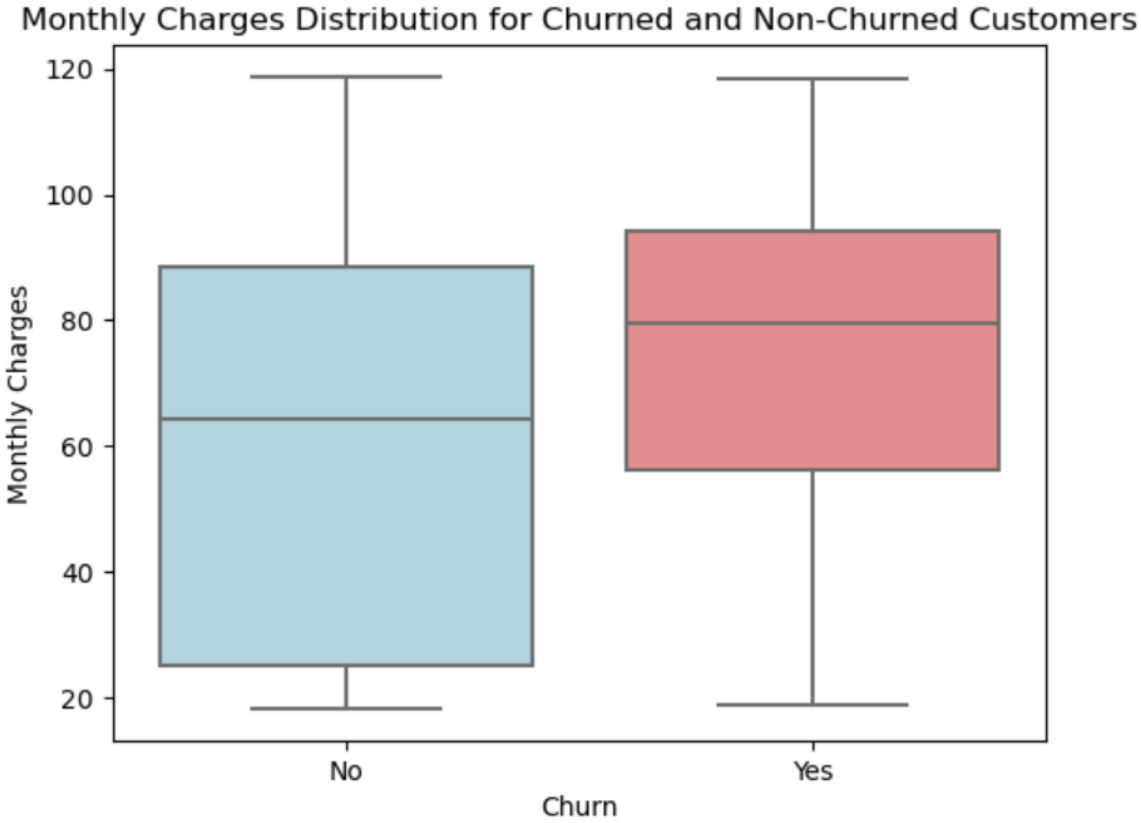
VISUALIZATION OF KEY FEATURES



GENDER OBSERVATIONS:

The company is patronized by 3,555 males compared to 3,488 females

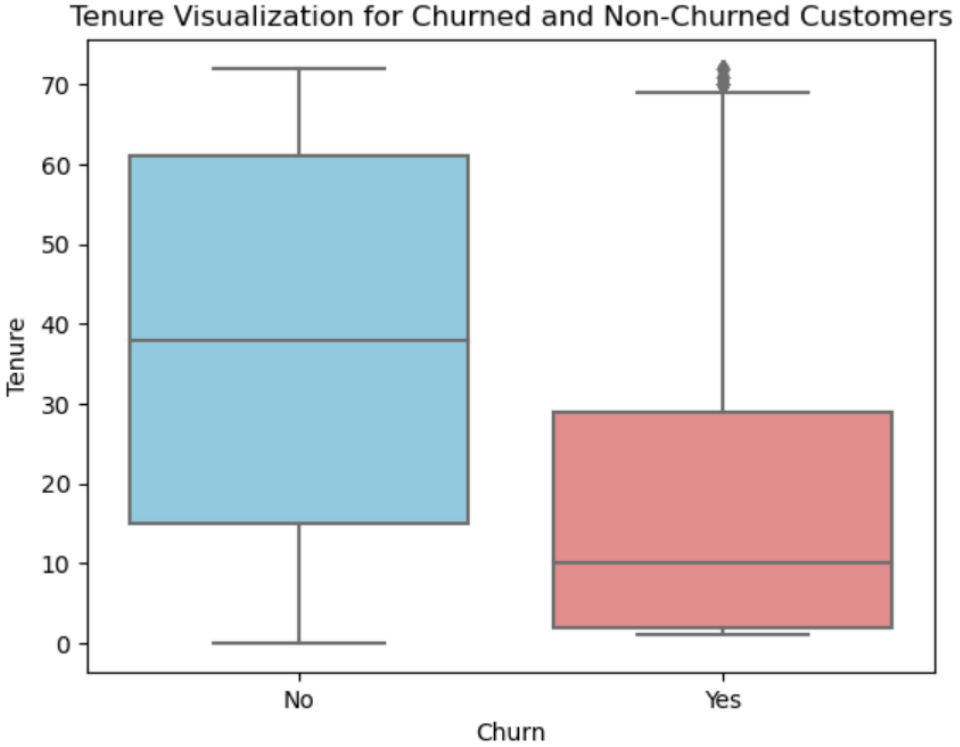
VISUALIZATION OF KEY FEATURES



OBSERVATION:

The median value for customers who have churned is elevated, suggesting they experience higher monthly charges compared to those who have not churned.

VISUALIZATION OF KEY FEATURES



OBSERVATION:

Observing the presence of outliers, it is evident that some customers experienced a relatively brief tenure before churning. Therefore, further investigation is necessary to understand the factors contributing to this phenomenon.

Visualization of key features

OBSERVATION:

The heatmap illustrates a notably strong correlation between Tenure and Total Charges. Additionally, Monthly Charges exhibit a high positive correlation with Total Charges. Overall, there is a positive correlation among all these variables.

EVALUATION MODELS

I will employ four distinct evaluation models to develop a churn detection system.

- **Logistic Regression Model**
- **Random Forest Model**
- **SGD Classifier Model**
- **Decision Tree Model**

EVALUATION MODELS

From the four Evaluation Models above Logistic Regression Model has better Metric results . This is because the goal is to predict customer churn. In that case, it is best to prioritize the model with higher recall, F1 score, and AUC-ROC. These metrics are particularly important for identifying customers who have churned accurately. These metrics help to identify customers who have churned accurately and minimize false negatives.:

- **Logistic Regression**

Accuracy: 0.8088026502602934
Precision: 0.6270491803278688
Recall: 0.5795454545454546
F1-score: 0.6023622047244095
AUC-ROC: 0.7323594780613708

OBSERVATIONS

The business should prioritize addressing False Negative (FN) results, as they incorrectly indicate that customers have not churned when they actually have. This misprediction diminishes the effectiveness of customer retention strategies.

THANK YOU