10Alytics
data & strategy

Machine-Learning-for-Churn-Detection

CONNECTTEL

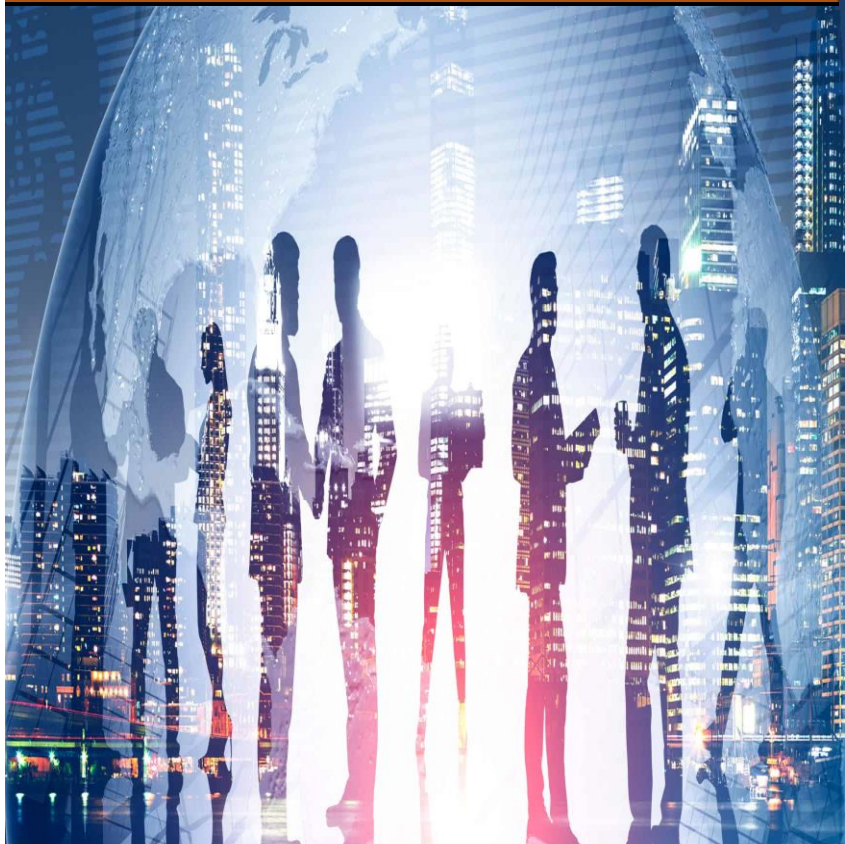
BACKGROUND

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 faces an imminent challenge in addressing the issue of customer churn, which poses a significant threat to the sustainability and growth of its business. The urgency lies in the fact that the current customer retention strategies employed by the company are not as precise and effective as needed, resulting in the departure of valuable customers to rival competitors. This situation not only jeopardizes the company's current market standing but also hinders its potential for expansion and long-term success. It is imperative for ConnectTel to swiftly reassess and bolster its retention strategies to ensure customer satisfaction, loyalty, and sustained competitiveness in the dynamic telecommunications landscape.

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.

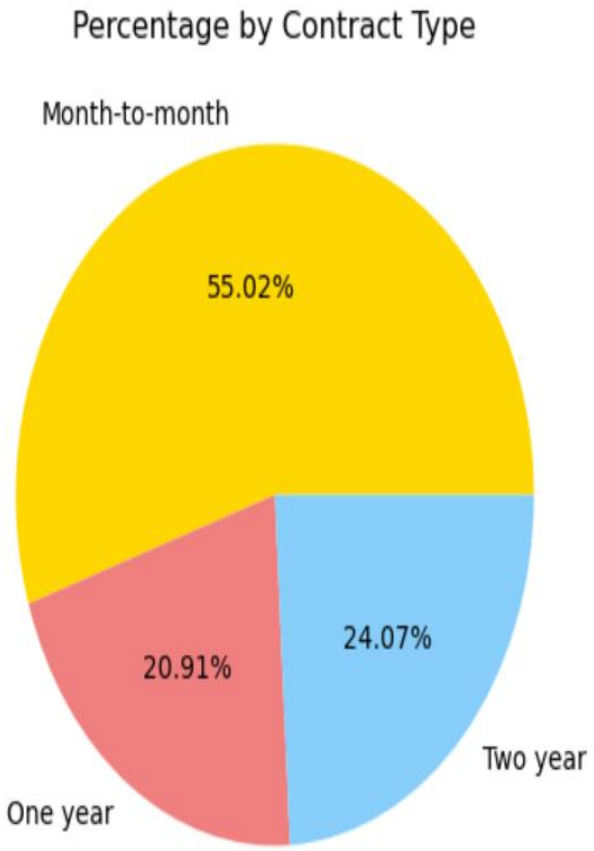
18. PaymentMethod: Indicates the method of payment used by the customer, such as electronic checks, mailed checks, bank transfers, or credit cards. This attribute helps in evaluating the impact of payment methods on churn.

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

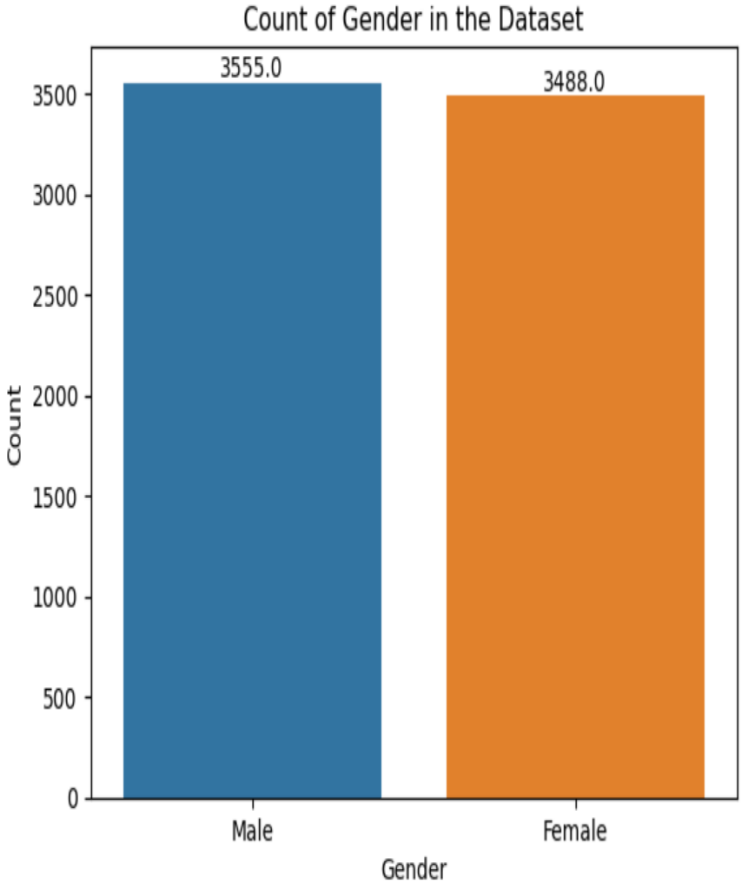


OBSERVATIONS:

ConnectTel Telecom Company is currently grappling with a pressing challenge that demands immediate attention – the issue of customer churn, which poses a formidable threat to the sustainability and expansion of its business. The urgency stems from the realization that the existing customer retention strategies employed by the company lack the required precision and effectiveness. This inadequacy has, unfortunately, translated into the departure of valuable customers who are opting for the services of competitors. The repercussions of this situation extend beyond mere customer loss; they directly impact the company's current market standing and future growth prospects. To mitigate this threat, ConnectTel must promptly reassess and fortify its customer retention strategies. By doing so, the company can not only stem the outflow of valuable clientele but also enhance customer satisfaction, loyalty, and overall competitiveness in the ever-evolving telecommunications industry. The strategic recalibration of retention initiatives is imperative to secure ConnectTel's position in the market and pave the way for sustained growth and success.

.

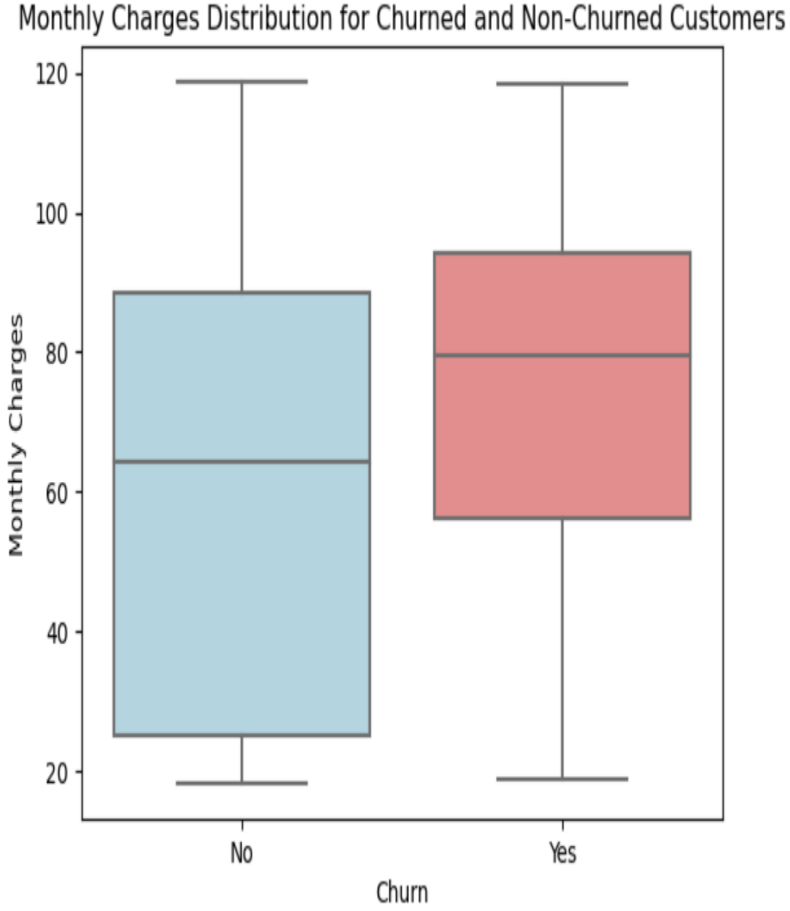
VISUALIZATION OF KEY FEATURES



GENDER OBSERVATIONS:

The company boasts a customer base comprising 3,555 males, outpacing the 3,488 females who also patronize its services. This gender distribution highlights a diverse clientele and provides valuable insights for marketing strategies and service offerings tailored to the preferences and needs of both male and female customers. The relatively balanced representation of both genders underscores the company's commitment to inclusivity and catering to a broad demographic spectrum. Analyzing and understanding these patronage patterns can contribute to the development of targeted and effective initiatives, ensuring enhanced customer satisfaction and sustained engagement across diverse segments of the population.

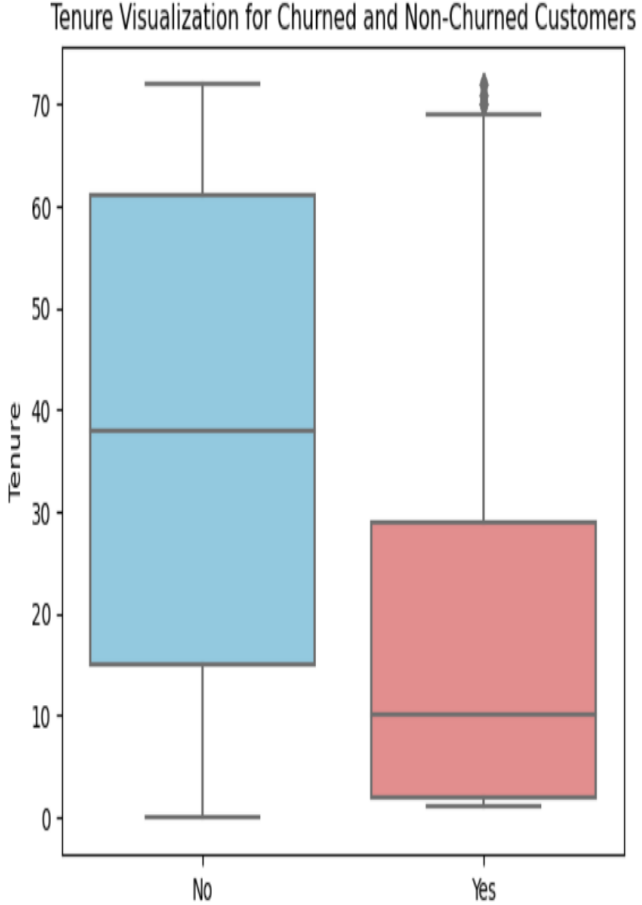
VISUALIZATION OF KEY FEATURES



OBSERVATION:

The median value for customers who have churned stands notably higher, indicating that this particular group tends to incur elevated monthly charges compared to their counterparts who have not churned. This observation provides valuable insights into the spending patterns of customers who ultimately choose to discontinue their services. The higher median suggests that a significant portion of customers opting for churn have, on average, been utilizing services associated with increased monthly charges. Understanding these distinctions in expenditure can guide the formulation of targeted retention strategies, aiming to address the concerns and preferences of customers with higher monthly charges and thereby mitigate the likelihood of churn.

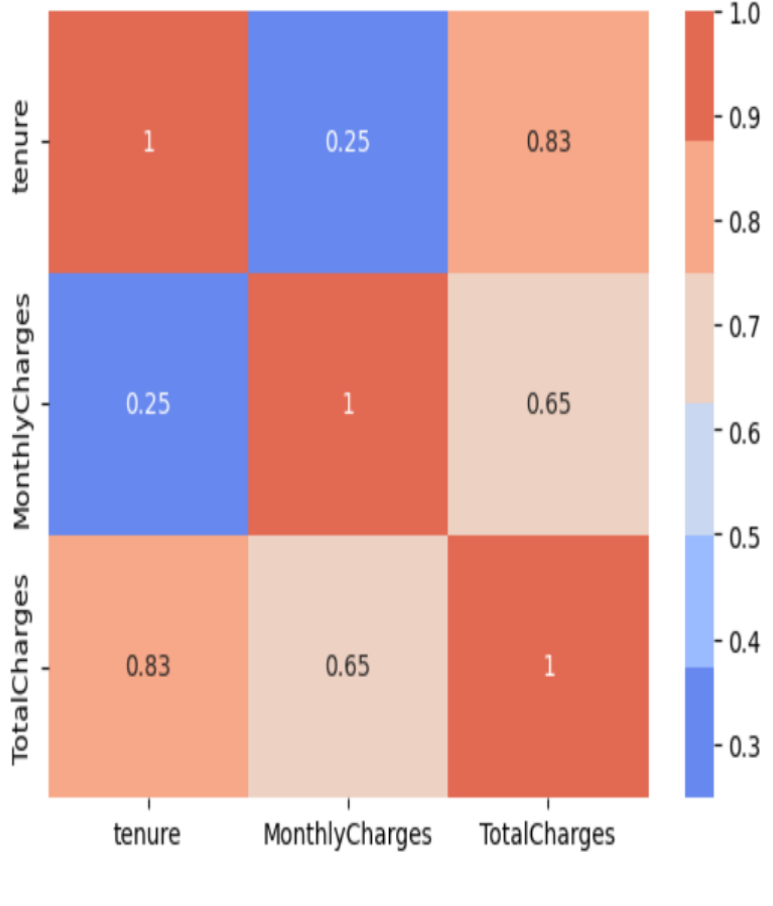
VISUALIZATION OF KEY FEATURES



OBSERVATION:

Upon careful observation of the dataset, it becomes apparent that there are outliers, signifying instances where certain customers had a notably short tenure before deciding to discontinue their services. This observation prompts the need for a more in-depth investigation to discern the underlying factors contributing to this phenomenon. These outliers suggest a unique subset of customers who, for various reasons, exhibited a rapid churn behavior. To gain a comprehensive understanding, further analysis is essential to uncover the specific attributes, patterns, or external factors that may have accelerated their decision to churn. By delving into these aspects, the company can acquire valuable insights that will aid in tailoring targeted retention strategies to address the distinct needs

VISUALIZATION OF KEY FEATURES



OBSERVATION:

The heatmap provides a visual representation that underscores a substantial and discernible correlation between Tenure and Total Charges. Beyond this, Monthly Charges also display a notably high positive correlation with Total Charges. In essence, the heatmap vividly communicates the presence of a positive correlation across all these variables. This interconnectedness among Tenure, Monthly Charges, and Total Charges suggests that as the tenure of customers increases, there is a corresponding rise in both monthly and total charges. Understanding these correlations is pivotal, as it informs the company about the intricate relationships between different aspects of customer engagement and billing. Such insights can prove invaluable in refining pricing strategies, designing tailored subscription plans, and optimizing customer satisfaction by aligning services with the evolving needs of the clientele

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

Among the four evaluation models listed earlier, the Logistic Regression Model stands out with superior metric results. This distinction is crucial, given the primary objective of predicting customer churn. The rationale behind favoring the Logistic Regression Model lies in its ability to prioritize key metrics such as recall, F1 score, and AUC-ROC.

In the context of predicting customer churn, it is paramount to place emphasis on models that exhibit higher recall, F1 score, and AUC-ROC values. These metrics play a pivotal role in accurately identifying customers who have churned, ensuring a more precise and reliable prediction. Specifically, a higher recall implies the model's effectiveness in capturing a greater proportion of actual churn cases, minimizing the occurrence of false negatives.

The F1 score, which balances precision and recall, provides a holistic evaluation of the model's predictive performance. Additionally, the AUC-ROC metric measures the model's ability to distinguish between positive and negative cases, further contributing to the overall accuracy of churn prediction.

In summary, prioritizing the Logistic Regression Model based on these robust metrics is instrumental in developing a reliable and effective churn prediction system, crucial for proactive customer retention strategies and ensuring the sustained success of the business.

- **Logistic Regression**

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

OBSERVATIONS

For the business lies in placing a heightened priority on rectify The imperative ing False Negative (FN) results. These instances inaccurately suggest that customers have not churned when, in reality, they have. This misprediction poses a significant challenge by diminishing the efficacy of customer retention strategies.

Addressing False Negative results is crucial for two primary reasons. Firstly, it ensures that customers who have genuinely churned are accurately identified, allowing the business to tailor retention initiatives to those specific individuals. Failure to do so may result in the neglect of valuable customers who require targeted efforts to prevent churn.

Secondly, mitigating False Negatives is essential for maintaining the overall precision and reliability of the churn prediction system. A higher rate of False Negatives implies that the model may be overlooking actual instances of customer churn, leading to missed opportunities for intervention and retention.

By prioritizing the resolution of False Negative results, the business can enhance the accuracy of its churn prediction system, subsequently bolstering the effectiveness of customer retention strategies and fortifying its position in the market.

THANK YOU