CONNECTTEL
CUSTOMER CHURN PREDICTION

James Ehiabhi

ABOUT US

ConnectTel is a leading telecommunications company at the forefront of innovation and connectivity solutions.

With a strong presence in the global market, ConnectTel has established itself as a trusted provider of reliable voice, data, and Internet services. Offering a comprehensive range of telecommunications solutions, including mobile networks, broadband connections, and enterprise solutions, ConnectTel caters to both individual and corporate customers, they are committed to providing exceptional customer service and cutting-edge technology. ConnectTel ensures seamless communication experiences for millions of users worldwide. Through strategic partnerships and a customer-centric approach, ConnectTel continues to revolutionize the telecom industry, empowering individuals and businesses to stay connected and thrive in the digital age.

James Ehiabhi

PROBLEM OVERVIEW

ConnectTel Telecom Company faces the pressing need to address customer churn, which poses 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.

To overcome this challenge, **ConnectTel** aims to develop a robust customer churn prediction system for which you have been contacted to handle as a Data Scientist. By leveraging advanced analytics and machine learning techniques on available customer data, the company seeks to accurately forecast customer churn and implement targeted retention initiatives.

This proactive approach will enable **ConnectTel** to reduce customer attrition, enhance customer loyalty, and maintain a competitive edge in the highly dynamic and competitive telecommunications industry.

James Ehiabhi

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.

James Ehiabhi

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.

James Ehiabhi

PROJECT ANALYSIS / GUIDE

1. **Problem definition**: clearly articulate the problem that is to be solved with your data mining. How will the company benefit from your solution?

2. **Perform exploratory data analysis in Python:**
   a) Visualize relationships between the label and some key features
   b) Explore correlations
   c) Conduct univariate, bivariate, and multivariate analysis as much as is feasible

3. **Perform feature engineering:**
   a) Encoding categorical variables
   b) Create new features from existing features where necessary, depending on insights from your EDA

4. **Model selection, training, and validation:**
   a) Train and test at least 3 supervised learning model

5. **Model evaluation:**
   a) Analyze the results of your trained model
   b) What metrics are most important for the problem? Should the business be more concerned with better results on false negatives or true positives?

6. **Submission:**
   a) Publish your Jupyter Notebook to your GitHub profile and Google Classroom.
   b) In the readme file, include a description of the project and summarize the steps you took and the challenges you faced.
   c) share the link with your instructor.

James Ehiabhi

THANK YOU
AND
LET'S CONNECT!

James Ehiabhi
Data Scientist | ML & AI Expert

kingjamesuwe@gmail.com
GitHub : JamesEhiabhi
James Ehiabhi