

DATA ANALYTICS WITH COGNOS – GROUP 3

CUSTOMER CHURN PREDICTION

PHASE 5

PROJECT DOCUMENTATION AND SUBMISSION

PROJECT OBJECTIVE:

Customer churn prediction involves developing models to forecast which customers are likely to leave a service or product. The project analyzes historical data, identifies relevant features, and employs machine learning algorithms to create a predictive model. The goal is to enable proactive retention strategies and enhance customer satisfaction. This project involves using IBM Cognos to identify factors influencing customer retention and relevant visualization Analysis.

DESIGN THINKING:

- Data Collection and preparation
- Exploratory Data Analysis (EDA)
- Test and Train data split
- Feature Scaling
- Model building
- Prediction
- Evaluating the model
- Data Visualization using Cognos

DEVELOPMENT PHASES:

PART 1: PREPROCESSING AND VISUALIZATION

We loaded the dataset and checked for null values and also derived insights from the data, then we split the dataset into features and target variable.

We visualized the dataset using Cognos and in cognos we used the waterfall plot, bar plot and line and column chart and analysed the relationship between distinct features and the target variable.

PART 2: MODEL BUILDING AND VISUALIZATION

We built the model of Random Forest Classifier by fitting the training data and after predicting using this model we evaluate the accuracy, confusion matrix and the overall classification report of the model.

We also analyse by visualizing the dataset using bullet chart and bar plot.

ANALYSIS OBJECTIVE:

The specific objective of analysing customer churn is to acknowledge the organization of the churners to enhance their business strategies and make decisions based on it.

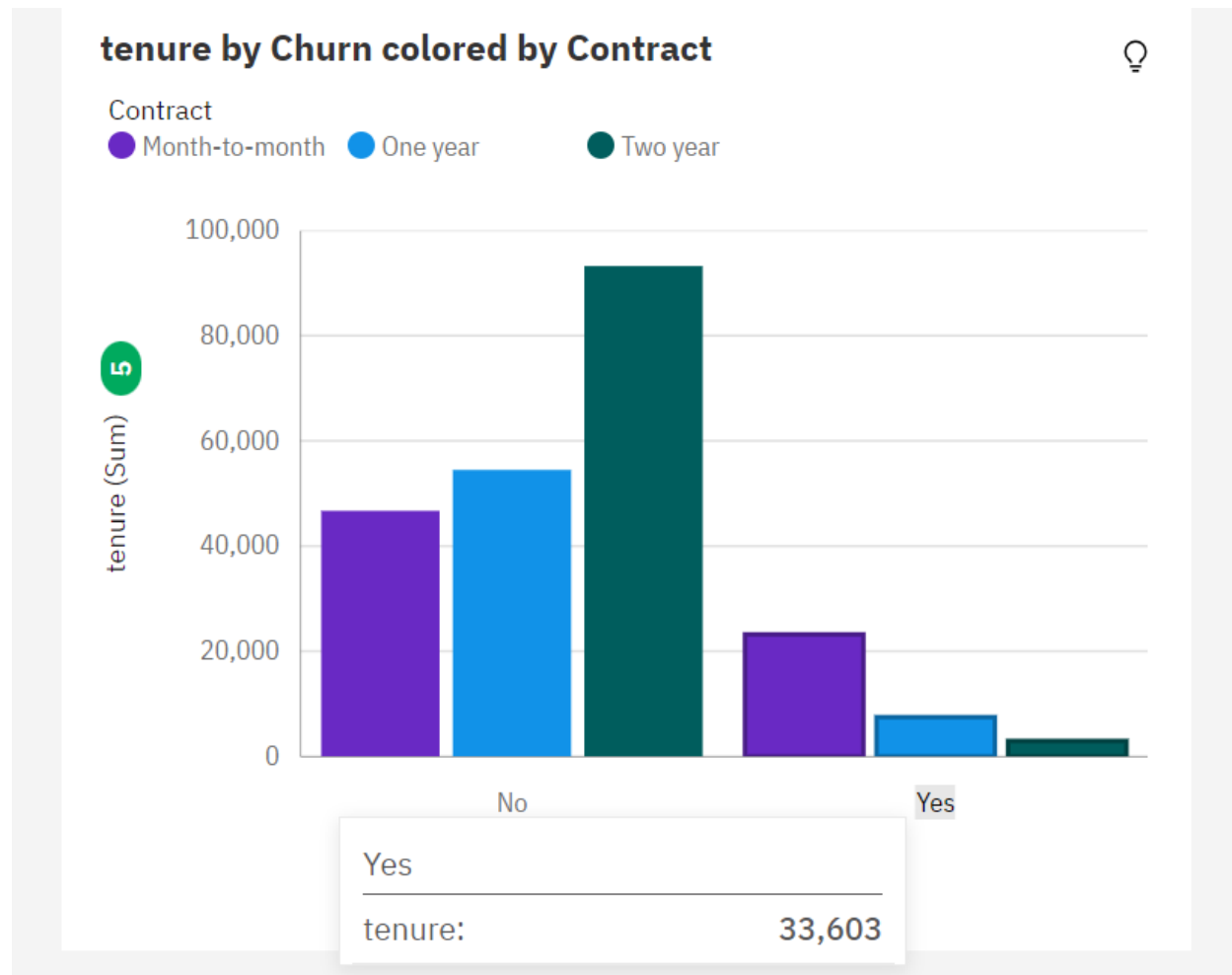
DATA COLLECTION:

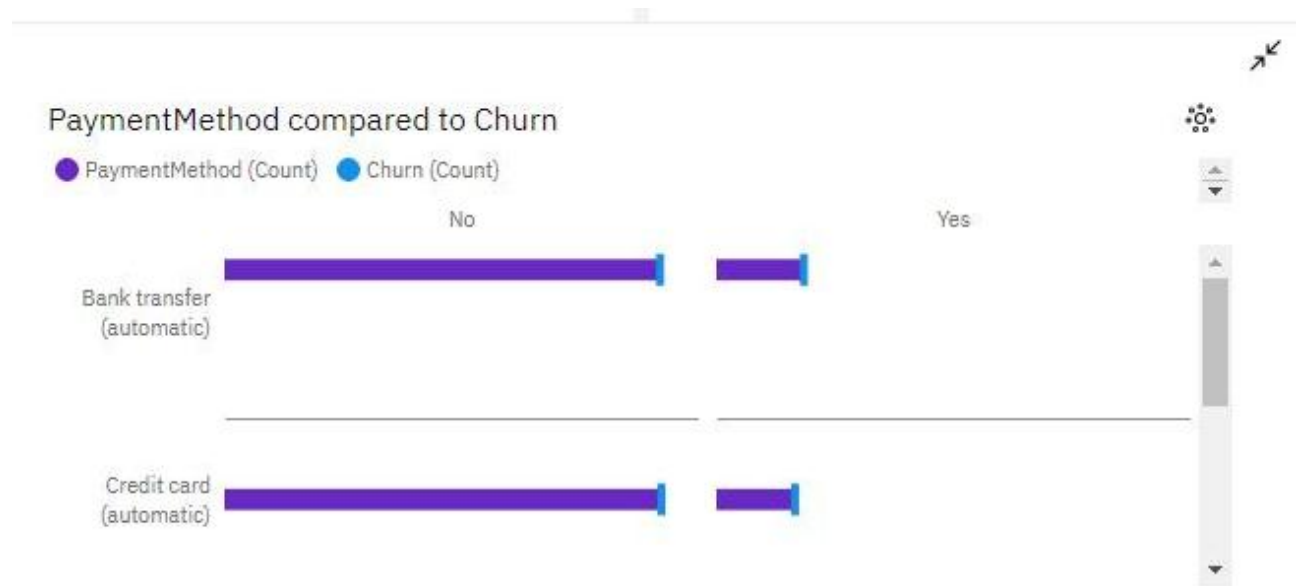
The data is collected based on the customer streaming time, The total charges, Device Protection, Online security, Online backup, etc., to determine whether the customer churns or not.

DATA VISUALIZATION USING COGNOS:

To convey the churn rates of the customer with clarity, We have employed various visualization techniques from IBM such as waterfall chart, bar plot,etc ensuring that the chosen methods are well-suited to the specific aspects of the data being portrayed.

Examples:





PYTHON CODE INTEGRATION:

The python code includes stages like

- Importing Libraries
- Loading the dataset
- Data Preprocessing
- Encode Categorical Variables
- Train_test_split
- Feature Scaling
- Building and training the model
- Making Prediction
- Evaluating the model

Examples:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
data= pd.read_csv("/content/churn_dataset.csv")
data.head()
```

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	...	DeviceProtection
0	7590-VHVEG	Female	0	Yes	No	1	No	No phone service	DSL	No	...	No
1	5575-GNVDE	Male	0	No	No	34	Yes	No	DSL	Yes	...	Yes
2	3668-QPYBK	Male	0	No	No	2	Yes	No	DSL	Yes	...	No
3	7795-CFCSW	Male	0	No	No	45	No	No phone service	DSL	Yes	...	Yes

```
[22] #importing libraries to design the model and the data to fit the model
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder, StandardScaler
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score, classification_report, confusion_matrix

[23] # to encode categorical variables
label_encoders = {}
for column in X.select_dtypes(include=['object']).columns:
    label_encoders[column] = LabelEncoder()
    X[column] = label_encoders[column].fit_transform(X[column])

[24] # Splitting data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Feature Scaling
scaler = StandardScaler()
X_train = scaler.fit_transform(X_train)
X_test = scaler.transform(X_test)

[26] #Building the random forest model and fitting train data
model = RandomForestClassifier(random_state=46).fit(X_train,y_train)

[27] # Making Prediction
y_pred = model.predict(X_test)

#attaining the accuracy of the model
print("Accuracy score",accuracy_score(y_test,y_pred))

Accuracy score 0.794180269694819

[20] #attaining the confusion matrix
print("Confusion Matrix:/n",confusion_matrix(y_test,y_pred))

Confusion Matrix:/n [[944  92]
 [198 175]]
```

We have trained a model to learn the dataset and predict the churn of the customers.

TAKEAWAY FROM INSIGHTS AND HOW IT HELPS THE WEBSITE OWNERS TO IMPROVE USER EXPERIENCE:

Predicting customer churn and gaining insights from the analysis can help website owners improve user experience in the following ways:

- 1. Identify At-Risk Users:** By analyzing customer churn patterns, website owners can identify users who are at risk of leaving. This allows them to proactively target these users with tailored retention strategies to keep them engaged.
- 2. Improve Content and Features:** Analysis can reveal which content or features are less engaging or may be causing users to leave. Website owners can then focus on enhancing or removing these elements to improve the overall user experience.
- 3. Personalization:** Insights from churn analysis can help in creating personalized user experiences. By understanding user preferences and behaviors, website owners can provide content and recommendations that are more relevant to individual users, increasing user satisfaction.
- 4. A/B Testing:** Website owners can use the insights to conduct A/B testing to assess the impact of different changes on user retention. This data-driven approach helps in making informed decisions about design and content changes.

5. Customer Feedback: Churn analysis can also prompt website owners to seek feedback from departing users. This can provide valuable insights into specific issues and areas for improvement.

6. Customer Support: Identify common issues or complaints that lead to churn and improve customer support resources to address these concerns promptly.

7. Optimize Performance: Analyzing the causes of churn can also lead to improvements in website performance, loading times, and overall user experience.

By leveraging insights from customer churn analysis, website owners can make data-driven decisions to enhance their websites, retain users, and ultimately provide a better user experience.

CONCLUSION:

This project on whole predicts the Churning of Customers based on the features such as tenure, total charges, etc. and also analyses the data by visualizing using Cognos. It helps the owners to make data driven decision and improve the customer experience.