**NAAN MUDHALAVAN PHASE - 5**

**IBM COGNOS DATA ANALYSIS**

**CUSTOMER CHURN PREDICTION**

**OBJECTIVE:**

The project involves using IBM Cognos to predict customer churn and identify factors influencing customer retention. The goal is to help businesses reduce customer attrition by understanding the patterns and reasons behind customers leaving. This project includes defining analysis objectives, collecting customer data, designing relevant visualizations in IBM Cognos, and building a predictive model

.  
**Analysis Objective:**

**1: Identify Potential Churners:**

Identifying customers who are likely to churn is the primary objective. By analysing historical data and customer behaviour patterns, businesses aim to accurately classify customers into potential churners and non-churners.

**2: Understand Churn Patterns and Trends**

The analysis should aim to identify patterns and trends in customer churn. This involves studying historical churn data over different time periods to recognize fluctuations and seasonal trends. Understanding when churn rates are typically high or low provides insights for timely interventions and resource allocation.

**3: Determine Key Churn Drivers**

Determining the factors leading to churn is crucial. By analysing customer data, feedback, and interactions, businesses can identify specific factors such as poor customer service, pricing issues, or product dissatisfaction. Pinpointing these drivers is essential for devising strategies that directly address the root causes of churn.

**4: Develop Predictive Models**

Developing predictive models using machine learning algorithms is a significant objective. These models utilize historical data and identified churn drivers to forecast future churn. Accurate predictive models enable businesses to proactively identify potential churners, allowing for timely and personalized retention efforts.

**5: Segment Customers for Targeted Strategies**

Segmentation of the customer base based on behaviour, demographics, or other relevant factors is essential. Analysing churn within these segments helps in tailoring retention strategies.

**Design thinking**

**1.Analysis Objectives:**

Define the specific objectives of predicting customer churn, such as identifying potential churners and understanding the key factors contributing to churn.

**2.Data Collection:**

Determine the sources and methods for collecting customer data, including customer demographics, usage behaviour, and historical interactions.

**3.Visualization Strategy:**

Plan how to visualize the insights using IBM Cognos, showcasing factors affecting churn and retention rates.

**4.Predictive Modelling:**

Decide on the machine learning algorithms and features to use for predicting customer churn.

**Development Phases**

**Phase 1:**

This phase deals with designing an approach to handle the project and appropriate ways to deal with the data.

**Phase 2:**

This phase involves the development and application of novel approaches, techniques, or technologies to better understand, predict, and address customer churn within a business or organization.

**Phase 3:**

In this Phase, the customer data has been collected from the provided source: Telco Customer Churn Dataset. The dataset contains various customer-related attributes such as customer ID, gender, tenure, monthly charges, contract type, payment method, and churn status.

This Phase also involves cleaning and Preprocessing of the dataset which is obtained from the source mentioned above.

The code for preprocessing the dataset is as below:

*import pandas as pd*

*data = pd.read\_csv(‘telco\_customer\_churn.csv’)*

*print(data.head())*

*print(data.info())*

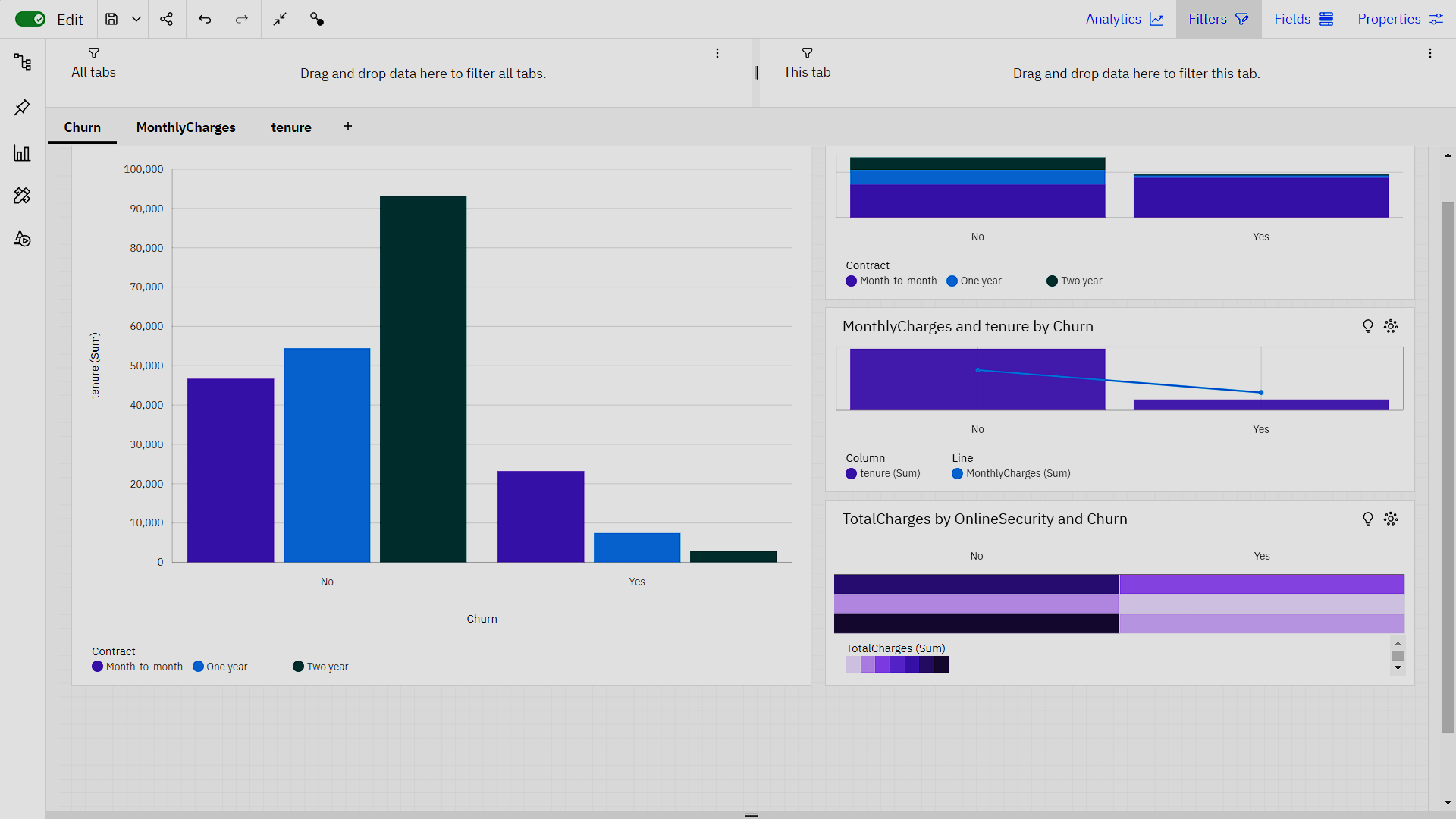
*print(“Duplicate Rows:”, data.duplicated().sum())*

*data\_encoded = pd.get\_dummies(data, columns=[‘gender’, ‘contract’, ‘payment\_method’])*

*data\_encoded.to\_csv(‘cleaned\_telco\_customer\_churn.csv’, index=False)*

**Phase 4:**

This Phase is all about visualizing the data using various suitable formats and the process of visualization is done using the IBM Cognos. The visualization is mainly on the areas such as churn patterns, retention rates, and key factors influencing churn.



**Visualization**

Along with these appropriate reports are created for factors such as customer segments, tenure, monthly charges, and contract details. 

**Report**

**PREDICTIVE MODEL**

The predictive model used for this analysis is **Random Forest Classifier**. Insights derived from data analysis and predictive models can significantly help businesses reduce customer churn in several ways:

1. Identifying Churn Patterns:

Insights: Analysing historical data can reveal patterns in customer behaviour leading to churn. For example, certain demographics, service usage patterns, or contract lengths might be associated with higher churn rates.

Prediction Model: Predictive models can identify these patterns automatically and highlight which factors contribute most significantly to churn.

2. Segmentation and Targeted Marketing:

Insights: Customer segmentation based on behaviour and demographics can identify high-churn-risk segments.

Prediction Model: Predictive models can categorize customers into segments, indicating which segments are more likely to churn.

Application: Businesses can then target these high-risk segments with personalized marketing campaigns, offers, or loyalty programs to increase retention.

3. Product and Service Improvements:

Insights: Analysing feedback and service usage data can pinpoint areas where customers are dissatisfied.

Prediction Model: Predictive analytics can predict which aspects of the product or service might lead to churn in the future.

Application: Insights can guide product/service improvements, leading to higher customer satisfaction and reduced churn.

4. Early Warning System:

Insights: Historical data analysis can identify early signs of customer dissatisfaction.

Prediction Model: Predictive models can provide real-time predictions of potential churn based on ongoing customer interactions.

Application: An early warning system allows businesses to proactively address customer issues, providing timely solutions and preventing churn.

5. Personalized Customer Engagement:

Insights: Customer behaviour data can reveal preferences, interests, and communication channel preferences.

Prediction Model: Predictive models can forecast which communication channels or engagement strategies are likely to retain specific customers.

Application: Businesses can engage with customers through their preferred channels, with personalized content, offers, and support, improving customer satisfaction and loyalty.

6. Customer Feedback Analysis:

Insights: Analysing customer feedback (reviews, surveys) can reveal common pain points.

Prediction Model: Sentiment analysis models can predict customer sentiment in real-time.

Application: By addressing negative feedback promptly and proactively, businesses can prevent customers from churning due to unresolved issues.

7. Optimizing Customer Support:

Insights: Analysing support ticket data can reveal common customer issues.

Prediction Model: Predictive models can forecast support ticket volumes and types based on historical patterns.

Application: Businesses can optimize support operations, ensuring quick and effective responses to customer issues, thereby improving customer satisfaction and reducing churn.

By leveraging insights from data analysis and predictive models, businesses can take proactive measures, optimize their operations, and engage with customers more effectively. This customer-centric approach leads to higher satisfaction, increased loyalty, and ultimately, reduced churn.

The code used for the predictive model is as follows:

*import pandas as pd*

*import numpy as np*

*from sklearn.model\_selection import train\_test\_split*

*from sklearn.ensemble import RandomForestClassifier*

*from sklearn.metrics import accuracy\_score, classification\_report*

*from sklearn.preprocessing import OneHotEncoder*

*data = pd.read\_csv('cleaned\_telco\_customer\_churn.csv')*

*data = data.drop(columns=['customerID']*

*X = data.drop(columns=['Churn'])*

*y = data['Churn'] # Target variable*

*non\_numeric\_cols = X.select\_dtypes(include=['object']).columns*

*X[non\_numeric\_cols] = X[non\_numeric\_cols].astype('category')*

*encoder = OneHotEncoder()*

*X\_encoded = encoder.fit\_transform(X).toarray()*

*X\_train, X\_test, y\_train, y\_test = train\_test\_split(X\_encoded, y, test\_size=0.2, random\_state=42)*

*model = RandomForestClassifier(random\_state=42)*

*model.fit(X\_train, y\_train)*

*predictions = model.predict(X\_test)*

*accuracy = accuracy\_score(y\_test, predictions)*

*report = classification\_report(y\_test, predictions)*

*print(f'Accuracy: {accuracy}')*

*print(f'Classification Report:\n{report}')*

**Output**:

*Accuracy: 0.7934705464868701*

*Classification Report:*

*precision recall f1-score support*

*No 0.83 0.91 0.87 1036*

*Yes 0.66 0.46 0.54 373*

*accuracy 0.79 1409*

*macro avg 0.74 0.69 0.70 1409*

*weighted avg 0.78 0.79 0.78 1409*

**Conclusion:**

The Customer Churn Prediction project represents a comprehensive exploration of data analytics and machine learning techniques to address a critical business challenge: reducing customer churn. Through meticulous data preprocessing, feature engineering, and the implementation of advanced machine learning algorithms, valuable insights have been derived, enabling a deep understanding of customer behaviour patterns.

In summary, this project serves as a testament to the power of data-driven decision-making. By leveraging sophisticated analytics and machine learning techniques, businesses can gain a competitive edge in retaining customers, fostering loyalty, and ensuring sustainable growth in today's dynamic market landscape.