Project Title: Churn Prediction in Telecom Business

Problem Definition:

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.

customerI	gender	SeniorCitiz	Partner	Dependen	tenure	PhoneServ	MultipleLi	InternetSe	OnlineSec	OnlineBac	DevicePro	TechSuppc	Streaming	StreamingI	Contract	PaperlessE	PaymentM	MonthlyCl	TotalCharg	Churn
7590-VHVI	Female	0	Yes	No	1	No	No phone	DSL	No	Yes	No	No	No	No	Month-to-	Yes	Electronic	29.85	29.85	No
5575-GNV	Male	0	No	No	34	Yes	No	DSL	Yes	No	Yes	No	No	No	One year	No	Mailed che	56.95	1889.5	No
3668-QPYI	Male	0	No	No	2	Yes	No	DSL	Yes	Yes	No	No	No	No	Month-to-	Yes	Mailed che	53.85	108.15	Yes
7795-CFO(	Male	0	No	No	45	No	No phone	DSL	Yes	No	Yes	Yes	No	No	One year	No	Bank trans	42.3	1840.75	No
9237-HQIT	Female	0	No	No	2	Yes	No	Fiber optic	No	No	No	No	No	No	Month-to-	Yes	Electronic	70.7	151.65	Yes
9305-CDS	Female	0	No	No	8	Yes	Yes	Fiber optic	No	No	Yes	No	Yes	Yes	Month-to-	Yes	Electronic	99.65	820.5	Yes
1452-KIOV	Male	0	No	Yes	22	Yes	Yes	Fiber optic	No	Yes	No	No	Yes	No	Month-to-	Yes	Credit carc	89.1	1949.4	No
6713-OKO	Female	0	No	No	10	No	No phone	DSL	Yes	No	No	No	No	No	Month-to-	No	Mailed che	29.75	301.9	No
7892-POO	Female	0	Yes	No	28	Yes	Yes	Fiber optic	No	No	Yes	Yes	Yes	Yes	Month-to-	Yes	Electronic	104.8	3046.05	Yes
6388-TABC	Male	0	No	Yes	62	Yes	No	DSL	Yes	Yes	No	No	No	No	One year	No	Bank trans	56.15	3487.95	No
9763-GRS	Male	0	Yes	Yes	13	Yes	No	DSL	Yes	No	No	No	No	No	Month-to-	Yes	Mailed che	49.95	587.45	No
7469-LKBC	Male	0	No	No	16	Yes	No	No	No interne	No interne	No interne	No interne	No interne	No interne	Two year	No	Credit carc	18.95	326.8	No
8091-TTV/	Male	0	Yes	No	58	Yes	Yes	Fiber optic	No	No	Yes	No	Yes	Yes	One year	No	Credit carc	100.35	5681.1	No
0280-XJGE	Male	0	No	No	49	Yes	Yes	Fiber optic	No	Yes	Yes	No	Yes	Yes	Month-to-	Yes	Bank trans	103.7	5036.3	Yes
5129-JLPIS	Male	0	No	No	25	Yes	No	Fiber optic	Yes	No	Yes	Yes	Yes	Yes	Month-to-	Yes	Electronic	105.5	2686.05	No
3655-SNQ	Female	0	Yes	Yes	69	Yes	Yes	Fiber optic	Yes	Yes	Yes	Yes	Yes	Yes	Two year	No	Credit carc	113.25	7895.15	No
8191-XWS	Female	0	No	No	52	Yes	No	No	No interne	No interne	No interne	No interne	No interne	No interne	One year	No	Mailed che	20.65	1022.95	No
9959-WOF	Male	0	No	Yes	71	Yes	Yes	Fiber optic	Yes	No	Yes	No	Yes	Yes	Two year	No	Bank trans	106.7	7382.25	No
4190-MFLI	Female	0	Yes	Yes	10	Yes	No	DSL	No	No	Yes	Yes	No	No	Month-to-	No	Credit carc	55.2	528.35	Yes
4183-MYF	Female	0	No	No	21	Yes	No	Fiber optic	No	Yes	Yes	No	No	Yes	Month-to-	Yes	Electronic	90.05	1862.9	No
8779-QRD	Male	1	No	No	1	No	No phone	DSL	No	No	Yes	No	No	Yes	Month-to-	Yes	Electronic	39.65	39.65	Yes
1680-VDC	Male	0	Yes	No	12	Yes	No	No	No interne	No interne	No interne	No interne	No interne	No interne	One year	No	Bank trans	19.8	202.25	No
1066-JKSG	Male	0	No	No	1	Yes	No	No	No interne	No interne	No interne	No interne	No interne	No interne	Month-to-	No	Mailed che	20.15	20.15	Yes
3638-WEA	Female	0	Yes	No	58	Yes	Yes	DSL	No	Yes	No	Yes	No	No	Two year	Yes	Credit carc	59.9	3505.1	No
6322-HRPI	Male	0	Yes	Yes	49	Yes	No	DSL	Yes	Yes	No	Yes	No	No	Month-to-	No	Credit carc	59.6	2970.3	No

Design Thinking:

1. Project Objectives:

Objective 1: Churn Prediction

- Develop a machine learning model to predict which customers are likely to churn.
- Achieve a high level of prediction accuracy and reliability to proactively address churn.

Objective 2: Identify Key Churn Drivers

- Determine the key factors and variables that contribute most to customer churn.
- Prioritize these factors based on their impact on churn to inform targeted strategies.

Objective 3: Recommend Retention Strategies

- Propose actionable strategies to reduce churn based on model insights.
- Create targeted retention plans for at-risk customer segments.

2. Analysis Approach:

Step 1: Data Collection and Integration

- Gather comprehensive customer data, including demographics, usage patterns, billing history, customer interactions, and historical churn records.
- Ensure data quality and consistency through data cleaning and validation processes.

Step 2: Data Preprocessing

- Handle missing data by imputing values or removing incomplete records.
- Address outliers and inconsistencies.
- Engineer relevant features from the data, such as customer tenure, call duration, plan details, customer complaints, and past interactions.
- Normalize or scale features as needed to make them suitable for modeling.

Step 3: Model Development

- Select suitable machine learning algorithms for churn prediction (e.g., logistic regression, decision trees, random forests).
- Split the dataset into training and testing sets to evaluate model performance.
- Train and validate the model, optimizing for high accuracy and precision in identifying potential churners.
- Experiment with hyperparameter tuning to improve model performance.

Step 4: Feature Importance Analysis

- Determine the importance of each feature in predicting churn using techniques like feature importance scores or SHAP (SHapley Additive exPlanations) values.
- Identify key drivers of churn, such as contract length, usage patterns, and customer satisfaction.

Step 5: Retention Strategy Recommendations

- Utilize model insights to recommend targeted retention strategies.
- Segment customers based on their likelihood to churn, allowing for tailored retention campaigns.
- Continuously monitor and update the model to adapt to changing customer behaviors and preferences.

3. Visualization Selection:

Line Charts and Trends:

- Visualize historical churn rates over time to identify trends, seasonality, and long-term patterns.
- Show how churn rates change across different time periods (e.g., monthly, quarterly).

Bar Charts and Feature Importance Plots:

- Present the importance of different features in predicting churn.
- Highlight key factors influencing churn, making it clear to stakeholders.

Heatmaps and Correlation Matrices:

- Illustrate correlations between various customer attributes (e.g., contract length, customer complaints) and churn.
- Visualize relationships between features to identify potential interactions.

Interactive Dashboards:

- Create interactive dashboards that enable stakeholders to monitor real-time churn predictions and explore data insights interactively.
- Include visualizations, summary statistics, and scenario planning tools.

Conclusion Questions:

1. How accurate is the churn prediction model in identifying customers at risk of churning?
2. What are the top factors influencing customer churn in the telecom business, and how significant are their impacts?
3. Are there specific customer segments that are more prone to churn than others?
4. How effective are the recommended retention strategies in reducing churn rates, and how do they perform in different customer segments?
5. How has churn behavior evolved over time, and are there any noticeable trends or seasonal patterns?
6. Can the model provide early warnings for potential churn, allowing for proactive

retention efforts?

7. Which customer attributes have the most significant impact on churn, and how can they be leveraged for targeted retention campaigns?
8. What is the Return on Investment (ROI) of implementing the churn prediction model and associated retention strategies, and how does it compare to the cost of customer acquisition?
9. This detailed approach provides a comprehensive framework for planning and executing a Churn Prediction project in the telecom industry, ensuring that objectives are clearly defined, data is processed effectively, models are developed and validated, and insights are communicated through appropriate visualization techniques.