Customer Churn Prediction – Phase 1

TEAM MEMBERS :

AKSHAYA K	**2021115011**
AKSHAYA SM	**2021115012**
PREETHAM V	**2021115077**
PRAVEEN M	**2021115075**
AKASH M	**2021115324**

PROBLEM STATEMENT:

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.

DESIGN THINKING:

1.) Analysis Objectives:

Using IBM Cognos, the specific objectives for predicting customer churn are:

Churn Prediction Model: Develop a predictive model within IBM Cognos to identify customers at risk of churning.

Churn Risk Score: Assign a churn risk score to each customer, facilitating targeted retention efforts.

Segmentation: Segment customers based on their churn risk levels for customized retention strategies.

Key Churn Drivers: Identify and visualize the primary factors contributing to customer churn using IBM Cognos analytics.

Retention Recommendations: Generate actionable insights and recommendations for retention strategies within the IBM Cognos environment.

Model Validation: Evaluate and validate the performance of the churn prediction model using IBM Cognos tools and visualizations.

2.) Data Collection:

We will be using the dataset provided by kaggle.com to carry on this project

https://www.kaggle.com/datasets/blastchar/telco-customer-churn

The above dataset contains necessary data like day, date etc. It also contains number of unique visits, first visits and returning visits which will be very helpful for us to enhance the user experience by identifying what they need the most.

3.) Visualization Strategy:

Visualization is a powerful tool in understanding customer churn data and conveying insights effectively. Here are some visualization strategies to consider when working with customer churn data:

Churn Rate Trends:

Line charts or time series plots can show how churn rates have evolved over time.

Compare churn rates among different customer segments using stacked area charts or grouped bar charts.

Customer Segmentation:

Create pie charts, bar charts, or treemaps to visually represent customer segments based on churn risk levels or demographics.

Use heatmaps to display correlation between different customer attributes and churn.

Customer Journey Mapping:

Flowcharts or Sankey diagrams can illustrate the customer journey, highlighting touchpoints where churn is more likely to occur.

4.) Predictive modeling

Predictive modeling is a process used in data science and machine learning to develop and train algorithms that can make predictions or classifications based on data. In the context of customer churn prediction, predictive modeling involves building a model that can forecast whether a customer is likely to churn (leave) or stay with a product or service in the future. Here are the key steps involved in predictive modeling for customer churn prediction:

Data Preparation:

Collect and clean the customer data, ensuring it's of high quality and ready for analysis.

Split the data into training and testing sets to evaluate model performance.

Feature Selection and Engineering:

Identify relevant features (variables) that could influence customer churn, such as demographics, usage patterns, or customer interactions.

Create new features or transform existing ones to enhance the predictive power of the model.

Model Selection:

Choose an appropriate machine learning algorithm for the task. Common choices include logistic regression, decision trees, random forests, support vector machines, or gradient boosting methods.

Experiment with multiple models to find the one that performs best for your dataset.

Model Training:

Use the training data to train the selected model. The model learns from historical data patterns to make predictions about future churn.

Model Evaluation:

Assess the model's performance using evaluation metrics such as accuracy, precision, recall, F1-score, ROC AUC, and others.

Validate the model on the testing dataset to ensure it generalizes well to new, unseen data.

Hyperparameter Tuning:

Fine-tune the model's hyperparameters to optimize its performance. Techniques like cross-validation can be employed to find the best hyperparameters.

Deployment:

Implement the trained model into your operational systems, allowing it to make real-time predictions on new customer data.

Monitoring and Maintenance:

Continuously monitor the model's performance in the production environment and retrain it as necessary to adapt to changing customer behaviors.

Interpretation and Actionable Insights:

Understand the model's predictions and the factors that contribute to churn, enabling the development of targeted retention strategies.