# CUSTOMER CHURN PREDUCTION

# PHASE -5

# DOCUMENT SUBMISSION

NAME: BALISETTY CHENCHU NITHIN

[au723921104003]

## PROJECT: CUSTOMER CHURN PREDICTION

## PHASE -5: PROJECT DOCUMENTATION & SUBMISSION

**Topic:**

* Outline the project's objective, design thinking process, and development phases.
* Describe the analysis objectives, data collection process, data visualization using IBM Cognos, and predictive modelling.
* Explain how the insights and prediction model can help businesses reduce customer churn.

**Project’s 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.

Design Thinking:

Creating a project on customer churn prediction is a valuable endeavor for businesses looking to retain their customers. Below, I outline a step-by-step guide on how to approach such a project:

Step 1: Define the Objectives and Scope

Clearly define the objectives of your customer churn prediction project. What do you want to achieve? Determine the scope of your project, such as the specific customer data you will use, the time frame you will consider, and the metrics you will use to measure churn.

Step 2: Data Collection and Preprocessing

Data Collection: Gather historical customer data, including information such as demographics, transaction history, customer service interactions, and more. Ensure data privacy and compliance with relevant regulations (e.g., GDPR).

Data Cleaning: Clean the data to remove duplicates, missing values, and outliers. Ensure data quality for accurate predictions.

Feature Engineering: Create relevant features from the raw data that can help improve model performance. For example, calculate customer tenure, create usage frequency metrics, or segment customers into categories.

Data Visualization: Explore the data through visualization to gain insights into customer behavior and identify patterns that may lead to churn.

Step 3: Data Splitting

Split your dataset into training, validation, and test sets. The training set is used to train your churn prediction model, the validation set is used to tune hyperparameters, and the test set is used to evaluate the final model's performance.

Step 4: Model Selection

Choose appropriate machine learning or deep learning algorithms for customer churn prediction. Common algorithms include logistic regression, decision trees, random forests, gradient boosting, neural networks, and more.

Step 5: Model Training

Train your chosen model(s) on the training dataset using appropriate performance metrics such as accuracy, precision, recall, F1-score, and AUC-ROC.

Step 6: Model Evaluation

Evaluate the model(s) using the validation dataset. Fine-tune hyperparameters to improve model performance. Perform cross-validation if needed.

Step 7: Model Testing

Assess the final model's performance on the test dataset to ensure it generalizes well to new, unseen data.

Step 8: Interpretability

Explainable AI techniques can help you understand why the model makes specific predictions. This is crucial for gaining insights into customer behavior.

Step 9: Deployment

Once satisfied with the model's performance, deploy it in a production environment. Ensure real-time or batch predictions are integrated into your business systems.

Step 10: Monitoring and Maintenance

Implement a monitoring system to keep track of customer churn predictions over time. Update the model as new data becomes available and business conditions change.

Step 11: Reporting and Visualization

Create reports and dashboards to communicate insights from the churn prediction model to stakeholders within your organization.

Step 12: Actionable Strategies

Use the insights from the model to develop actionable strategies for retaining customers. Implement marketing campaigns, loyalty programs, or customer support improvements based on the predictions.

Step 13: Continuous Improvement

Continuously analyze the impact of your strategies and refine your churn prediction model based on the results. Iterate on the project to achieve better customer retention rates.

Remember that customer churn prediction is an ongoing process, and success depends on the ability to adapt and refine your strategies based on the evolving needs of your customers and changes in the market.

**Development Phases:**

Customer churn prediction typically involves several development phases to build an effective model. These phases help you identify and prevent customer churn in your business. Here's an overview of the typical phases involved:

1. Problem Definition:

- Define the problem you want to address, such as reducing customer churn in your business.

- Determine the key objectives and goals of your churn prediction model.

2. Data Collection and Preprocessing:

- Gather relevant data from various sources, including customer information, transaction history, and behaviour data.

- Clean and preprocess the data, handling missing values and outliers.

- Feature engineering: Create new features or transform existing ones to make them more suitable for modelling.

3. Exploratory Data Analysis (EDA):

- Conduct EDA to gain insights into your data and identify patterns that may be related to churn.

- Visualize data and calculate summary statistics to understand the characteristics of your customers and their behaviours.

4. Data Splitting:

- Split your data into training, validation, and test sets to train and evaluate your model.

5. Model Selection:

- Choose appropriate machine learning or statistical models for your churn prediction task. Common choices include logistic regression, decision trees, random forests, gradient boosting, and neural networks.

6. Model Training:

- Train your selected model(s) using the training dataset.

- Optimize hyperparameters and tune the model to improve its performance.

7. Model Evaluation:

- Evaluate your model on the validation dataset using appropriate metrics like accuracy, precision, recall, F1-score, and ROC-AUC.

- Consider the business context and the cost of false positives and false negatives when choosing evaluation metrics.

8. Model Selection and Fine-Tuning:

- If necessary, iterate on model selection and hyperparameter tuning based on validation performance.

9. Model Interpretation:

- Interpret the model to understand which features are most influential in predicting churn. This insight can guide business decisions and interventions.

10. Model Deployment:

- Deploy the final model into your production environment to make real-time predictions or automate churn identification.

11. Monitoring and Maintenance:

- Continuously monitor the model's performance in a production environment.

- Retrain the model periodically to account for changing customer behaviours.

12. Feedback Loop:

- Implement a feedback loop that collects customer feedback and actual churn data.

- Use this data to improve the model and business processes.

13. Business Actions:

- Based on the model predictions, take proactive measures to retain at-risk customers. This could include personalized offers, outreach, or other retention strategies.

14. Reporting and Visualization:

- Develop dashboards and reports to track the performance of your churn prediction model and the effectiveness of your retention strategies.

Each of these phases is crucial in the development of a customer churn prediction system. It's an ongoing process that requires continuous improvement and adaptation to changing customer behaviours and business needs.

**Data Collection Process:**

The data collection process for customer churn prediction involves gathering and aggregating relevant data from various sources to build a dataset that can be used to train and evaluate your churn prediction model. Here are the steps involved in the data collection process:

1. Define Data Requirements:

- Determine what data is essential for your customer churn prediction model. This could include customer information, transaction history, customer interactions, product usage, and any other relevant details.

2. Identify Data Sources:

- Identify where the required data can be found. Common data sources may include:

- Customer databases

- Billing and payment systems

- CRM (Customer Relationship Management) software

- Customer support and service records

- Website and app usage logs

- Surveys and feedback forms

- Social media data

3. Data Extraction:

- Extract data from the identified sources. Depending on the source, this may involve database queries, API calls, log file analysis, or manual data entry.

4. Data Integration:

- Combine data from different sources into a unified dataset. Ensure that data is properly aligned and that common identifiers (e.g., customer IDs) are consistent across sources.

5. Data Cleaning:

- Clean the collected data to address issues such as missing values, outliers, and data inconsistencies. Common data cleaning tasks include imputing missing values and removing duplicates.

6. Data Transformation:

- Prepare the data for analysis by performing transformations such as one-hot encoding, scaling, and feature engineering. This step may involve creating new features that can provide valuable insights into customer behaviour.

7. Data Quality Assessment:

- Assess the quality of the data to ensure it is suitable for modelling. This includes checking for data integrity, consistency, and accuracy.

8. Data Storage:

- Store the cleaned and transformed data in a suitable format or database for easy access and retrieval during model development and deployment.

9. Data Privacy and Compliance:

- Ensure that data collection and storage adhere to data privacy regulations and compliance standards, such as GDPR or HIPAA, as relevant to your business and data.

10. Data Sampling or Balancing:

- Depending on the class distribution (churned vs. non-churned customers), you may need to perform sampling or balancing techniques to ensure the dataset is not heavily biased.

11. Data Documentation:

- Document the data sources, data processing steps, and any data transformations. This documentation is essential for maintaining data lineage and understanding the data's context.

12. Data Version Control:

- Implement version control for your dataset to keep track of changes over time. This helps ensure data consistency and reproducibility of results.

13. Automate Data Collection (if possible):

- Consider setting up automated processes for regular data collection and updates, as customer behaviour and attributes can change over time.

14. Data Security:

- Implement data security measures to protect sensitive customer information, especially if you're dealing with personal or financial data.

The quality and relevance of the data you collect play a crucial role in the performance of your customer churn prediction model. Regularly monitor and update your data to account for changes in customer behaviour and to improve the accuracy of your predictions.

IBM Cognos is a business intelligence and analytics platform that provides tools for data visualization, reporting, and predictive modelling. Here's how you can use IBM Cognos for data visualization and predictive modelling:

**Data Visualization with IBM Cognos:**

1. Data Preparation:

- Start by preparing your data for visualization. Clean, transform, and format your data to ensure it's suitable for analysis.

2. Connect Data Sources:

- In IBM Cognos, you can connect to various data sources, including databases, spreadsheets, or cloud services. Import your cleaned data into Cognos.

3. Data Modelling:

- Create a data model in IBM Cognos Framework Manager, defining relationships between different data tables and optimizing the structure for reporting and visualization.

4. Create Reports and Dashboards:

- Use IBM Cognos Report Studio or IBM Cognos Dashboard to design and build reports and dashboards. You can drag and drop data elements onto the canvas to create visuals like charts, graphs, tables, and maps.

5. Data Visualization:

- Customize the appearance and style of your visuals to convey insights effectively. IBM Cognos offers a range of formatting and design options.

6. Interactivity:

- Add interactivity to your reports and dashboards, such as filters, prompts, and drill-through actions, so users can explore data and gain deeper insights.

7. Schedule and Share:

- Schedule reports to be generated and shared automatically with stakeholders through email, FTP, or other channels.

**Predictive Modelling with IBM Cognos:**

IBM Cognos also supports predictive modelling using IBM SPSS Modeler. Here's how you can perform predictive modelling with Cognos:

1. Data Preparation:

- Begin with data preparation, as predictive modeling requires clean and well-structured data.

2. Data Import:

- Import your dataset into IBM SPSS Modeler, which is often integrated with IBM Cognos.

3. Feature Selection and Engineering:

- Select relevant features and engineer new ones to improve model performance.

4. Model Building:

- Use SPSS Modeler to create and train predictive models. You can choose from various algorithms like decision trees, logistic regression, neural networks, and more.

5. Evaluation:

- Evaluate the performance of your predictive models by using metrics like accuracy, precision, recall, F1-score, ROC-AUC, and others.

6. Deployment:

- Once you have a model with satisfactory performance, deploy it in IBM Cognos for use in your reports and dashboards.

7. Integration:

- Integrate predictive insights into your visualizations and reports. For instance, you can display predictive scores or probabilities in a dashboard.

8. Continuous Monitoring

- Regularly monitor the performance of your predictive models, and retrain them as necessary to adapt to changing data patterns.

IBM Cognos and SPSS Modeler offer a comprehensive platform for both data visualization and predictive modeling, allowing you to create actionable insights and reports that help drive business decisions. The integration of these tools can provide a seamless workflow for data analysis and reporting.

**Insights and Prediction Model:**

Insights and prediction models can be invaluable tools for helping businesses reduce customer churn. Here's how they can work together to achieve this:

1. Identifying At-Risk Customers:

- Prediction models can analyze historical customer data and behaviour to identify patterns associated with churn. By examining factors like usage frequency, purchase history, customer support interactions, and customer demographics, the model can flag customers who are at risk of churning in the near future.

2. Segmentation:

- Once at-risk customers are identified, businesses can use insights to segment them into different groups based on the reasons for potential churn. For example, some customers may be price-sensitive, while others may have had a poor experience with customer support. This segmentation helps tailor retention strategies to each group.

3. Personalized Retention Strategies:

- Armed with insights, businesses can design personalized retention strategies for each customer segment. For instance:

- Offering discounts or promotions to price-sensitive customers.

- Providing exceptional customer service to those who had support issues.

- Recommending new products or services based on individual preferences.

- Encouraging engagement through targeted emails or notifications.

4. Timely Interventions:

- The prediction model can also predict the timeframe within which a customer is likely to churn. This information allows businesses to intervene proactively. For example, if the model predicts a customer will churn in the next 30 days, a timely discount offer or a customer satisfaction survey can be sent to retain the customer.

5. A/B Testing:

- Insights and prediction models enable businesses to run A/B tests to measure the effectiveness of different retention strategies. By comparing the outcomes of different interventions, companies can refine their approaches over time.

6. Feedback Loop:

- After implementing retention strategies, collect feedback from customers. Insights can be gathered from survey responses, customer feedback, and customer behavior. This feedback loop helps refine the strategies further.

7. Monitoring and Evaluation:

- Continuously monitor the performance of the prediction model and the effectiveness of retention strategies. Adjust the model and strategies as necessary to adapt to changing customer behavior.

8. Resource Allocation:

- Insights can help businesses allocate resources more effectively. For example, they can focus their retention efforts on high-value customers who have a higher impact on revenue.

9. Customer Lifetime Value (CLV) Optimization:

- Insights and prediction models can be used to not only reduce churn but also increase CLV. By identifying profitable customer segments, businesses can focus on acquiring and retaining those customers who contribute the most to long-term revenue.

10. Competitive Advantage:

- Businesses that effectively reduce churn gain a competitive advantage. A lower churn rate can lead to higher customer retention, increased revenue, and enhanced brand reputation.

11. Cost Reduction:

- Reducing customer churn can also lead to cost savings. Acquiring new customers is typically more expensive than retaining existing ones, so a focus on churn reduction can lead to improved cost-efficiency.

In summary, insights derived from predictive models help businesses understand the drivers of churn and the characteristics of at-risk customers. This understanding, combined with personalized retention strategies and timely interventions, enables businesses to reduce churn and enhance customer loyalty, leading to increased revenue and improved long-term profitability.

**SUBMITTED BY,**

BALISETTY CHENCHU NITHIN