Customer Churn Analysis and Retention Strategy

Presented by: Jana Mohamed El-Sayed Mohamed

Date: 30 November 2025

Project Idea: Addressing Customer Attrition

The Problem

Businesses consistently face the challenge of customer attrition, commonly known as churn. This phenomenon directly impacts revenue streams, hinders growth, and erodes customer lifetime value. Proactively identifying high-risk customers before they defect is crucial for sustained business success.

Proposed Solution

Predictive Model

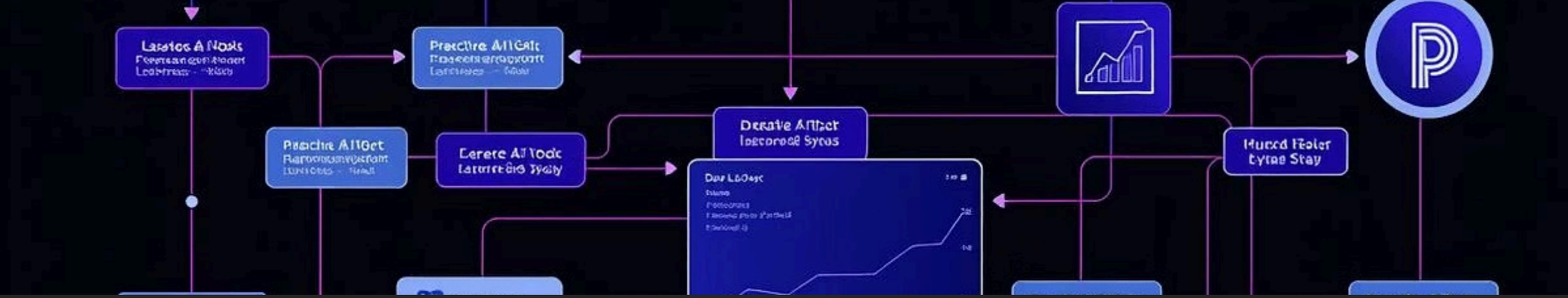
Development of a robust predictive model designed to analyse customer data.

Churn Forecasting

Accurately forecast churn risk for individual customers based on their behaviour patterns.

Data-Backed Strategy

Provide a data-driven retention strategy to mitigate identified churn risks effectively.



Project Wireframe: Conceptual Data Flow & User Interface

Our solution moves beyond mere observation, offering actionable, predictive insights through a deployable application for immediate risk assessment.

1

1. Input Data

Business users input customer data (tenure, usage, complaints).

2

2. Process

Deployed model processes the input data in real-time.

3

3. Output Prediction

Instant churn prediction (High/Low risk) and probability displayed.

4

4. Action

Users receive recommended retention strategies (e.g., offer discount).

End Users and Key Features

Marketing/Retention Team Managers

Need: To identify and prioritise vulnerable customer segments for targeted campaigns.

- Key Feature: Comprehensive Model Evaluation Report (Precision/Recall/F1) to ensure reliability for budget allocation decisions.

Customer Service Representatives (CSRs)

Need: Real-time churn risk assessment during customer interactions to guide conversations.

- Key Feature: Intuitive Deployed Model Interface (Streamlit/Flask app) for instant churn predictions.

Data Analysts/Scientists

Need: To understand the primary drivers of churn and continuously refine predictive capabilities.

- Key Feature: Detailed Feature Importance Plots for interpreting customer behaviour patterns and model transparency.

Data Structure and Flow

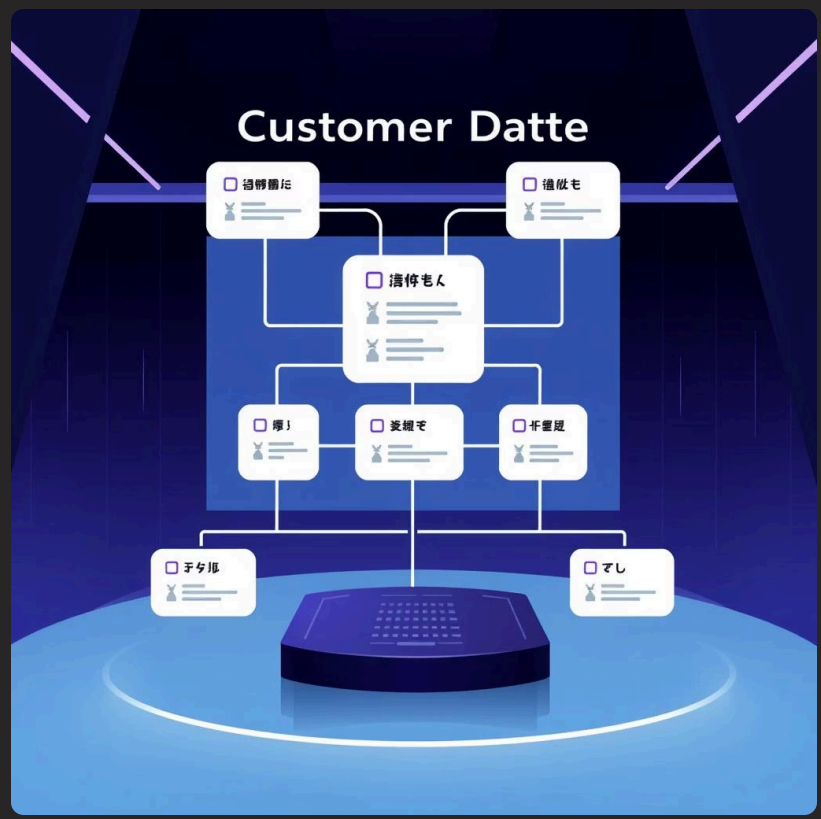
Database Architecture

Our initial modelling approach assumes a flat-file or CSV-based data architecture, providing flexibility and ease of integration for diverse datasets. This structure is designed for straightforward data handling.

Key Entities & Relationships

The core of our data model revolves around a single entity: **Customer**. This entity encapsulates critical features such as:

- **customer_id:** Unique identifier
- **tenure:** Length of customer relationship
- **usage:** Customer activity metrics
- **complaints:** Record of customer issues
- **churn_status:** The target variable (binary: churned/not churned)

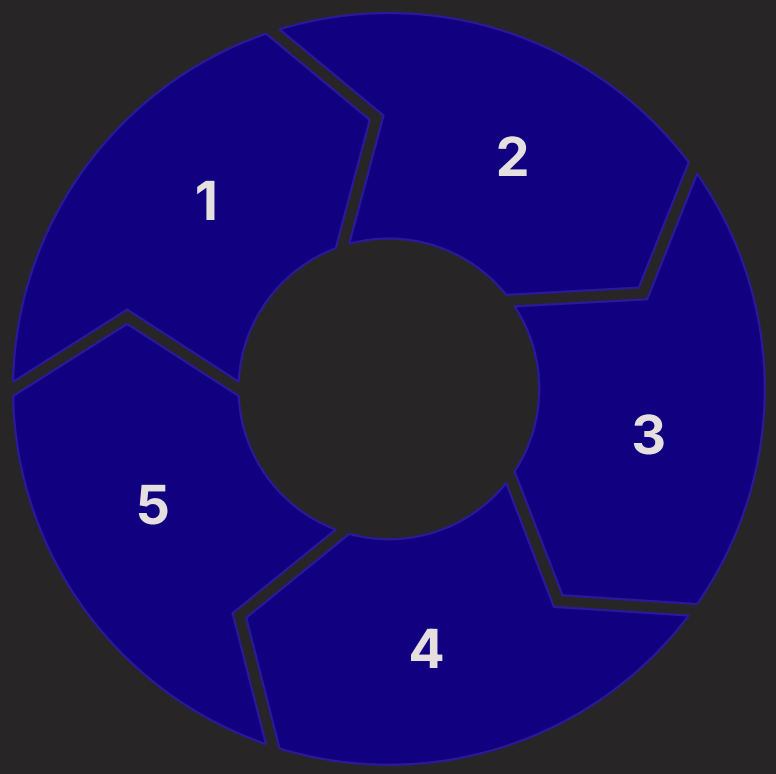


Data Collection

Utilising public or provided datasets.

Model Storage

Saving trained models as deployable artifacts.



Preprocessing

Encoding categorical features, scaling numeric features, handling missing data, and addressing class imbalance.

Data Split

Dividing data into training and testing sets.

Model Training

Feeding preprocessed data into chosen algorithms.

Programming Languages & Frameworks

Python

The primary language for all ML/Data Science tasks, leveraging its extensive ecosystem.

Scikit-learn

Essential for implementing foundational ML algorithms such as Logistic Regression and Random Forest.

XGBoost

Utilised for advanced gradient boosting techniques to enhance model performance and accuracy.

Streamlit/Flask

Chosen for building a simple, intuitive web application for model deployment and user interaction.

Google Colab

The collaborative development environment for initial exploration, prototyping, and model training.

Supporting Libraries

Including Pandas for data manipulation, and Matplotlib/Seaborn for comprehensive data visualisation.

Live Application & Testing Phases

Our project is currently in the crucial modelling and deployment phase, with active development on Google Colab. The ultimate goal is a fully functional, deployed model interface.

1 Model Development & Training

Initial model building and iterative training on prepared datasets.

2 Hyperparameter Tuning

Systematic optimisation of model parameters for peak performance, focusing on chosen classification algorithms like Random Forest.

3 Model Evaluation

Rigorous assessment using key metrics: Accuracy, Precision, Recall, F1-score, and ROC-AUC curve analysis.

4 Deployment Interface

Transitioning the trained model into a user-friendly application for real-time predictions.

5 User Acceptance Testing

Gathering feedback to ensure the model's predictions and interface meet business objectives and user needs for effective retention strategies.

Key Deliverables & Project Timeline

Reports & Documentation

- **EDA Notebook:** Visualisations and insights from exploratory data analysis.
- **Summary Report:** Key customer behaviour insights driving churn.
- **Model Evaluation Report:** Detailed performance metrics and model robustness.
- **Retention Strategy Report:** Data-backed actionable strategies.

Other Final Products

- Cleaned dataset ready for further modelling.
- Trained model with complete training code.
- Deployed model interface for user interaction.

Timeline for Deliverables (Milestones)

Milestone 1

Data Preprocessing & EDA Notebook complete.

Milestone 2

Trained model & Feature importance plots complete.

Milestone 3

Deployed model interface & Retention strategy report complete.

Project Team & Roles

Jana Mohamed El-Sayed Mohamed (21032500)

Steven Tamer Soliman (21081767)

Mohab Sherif Mohamed (21039536)

Arwa Mohamed (21098099)

Lamis Abdallah Essmat Abdelhamid (21044350)

Malak Mahmoud Shehata Mahmoud (21086742)

Key Responsibilities

Roles within the team were dynamically distributed, encompassing Data Collection, Preprocessing, Model Development, Deployment Engineering, and Report Writing, ensuring comprehensive coverage of all project phases.

Collaboration Methods

Our team leveraged a collaborative work environment, primarily Google Colab, to facilitate seamless integration of efforts and real-time project progression.

Thank You

We appreciate your time and attention to our project.

Contact Information

Janamohamed7178@gmail.com

We welcome any questions or feedback you may have.