



Customer Churn Prediction for a Subscription- Based Business

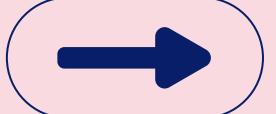
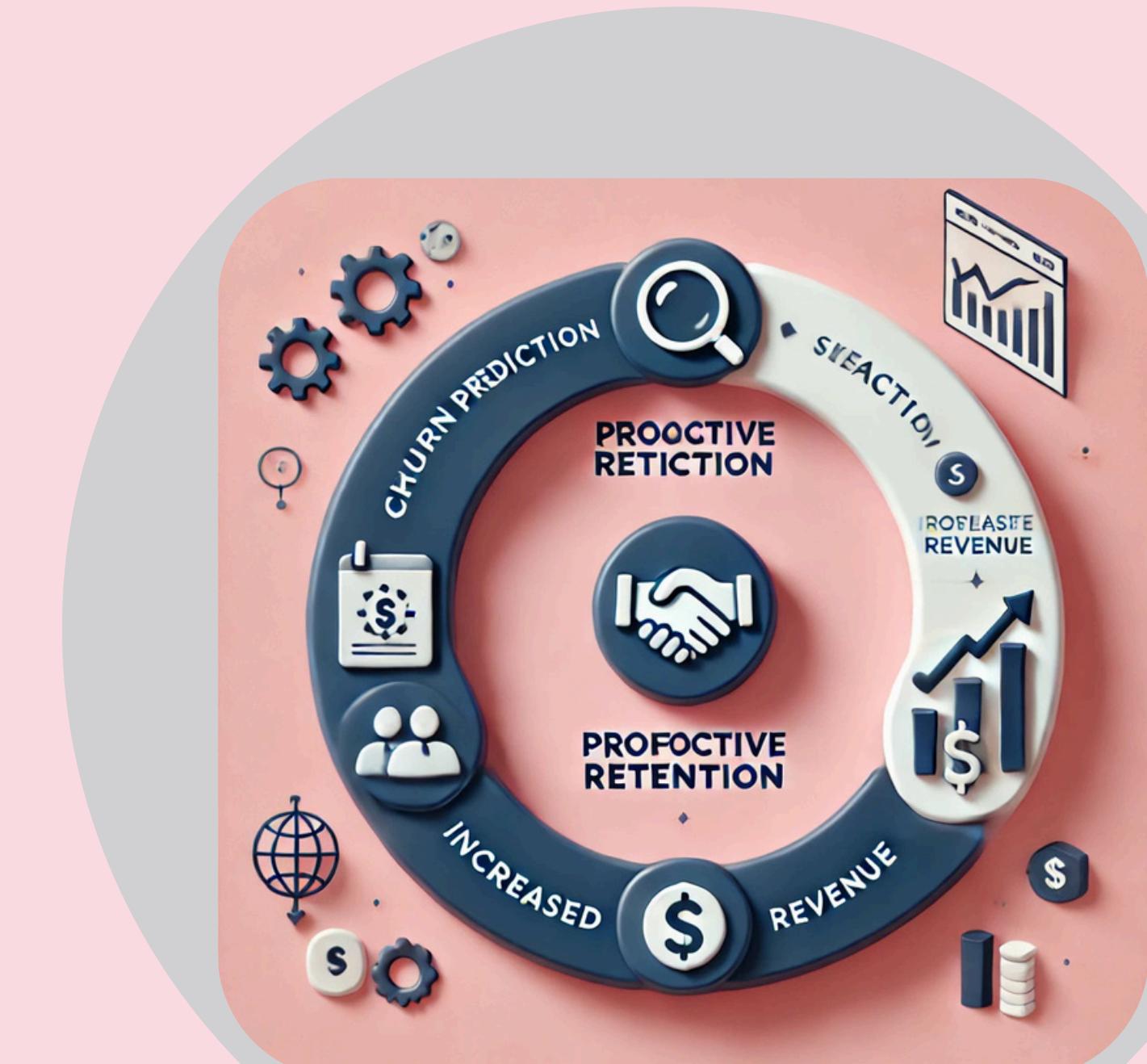
Using Machine Learning to Retain
Customers and Drive Growth

Presented by : Deepti Gupta



Introduction

- **Objective:** The goal of this project is to predict customer churn for a subscription-based streaming service. By identifying at-risk customers, the business can implement targeted retention strategies.
- **Why It Matters:** Reducing churn improves customer lifetime value (CLV), enhances revenue, and lowers acquisition costs.

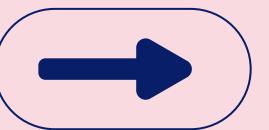
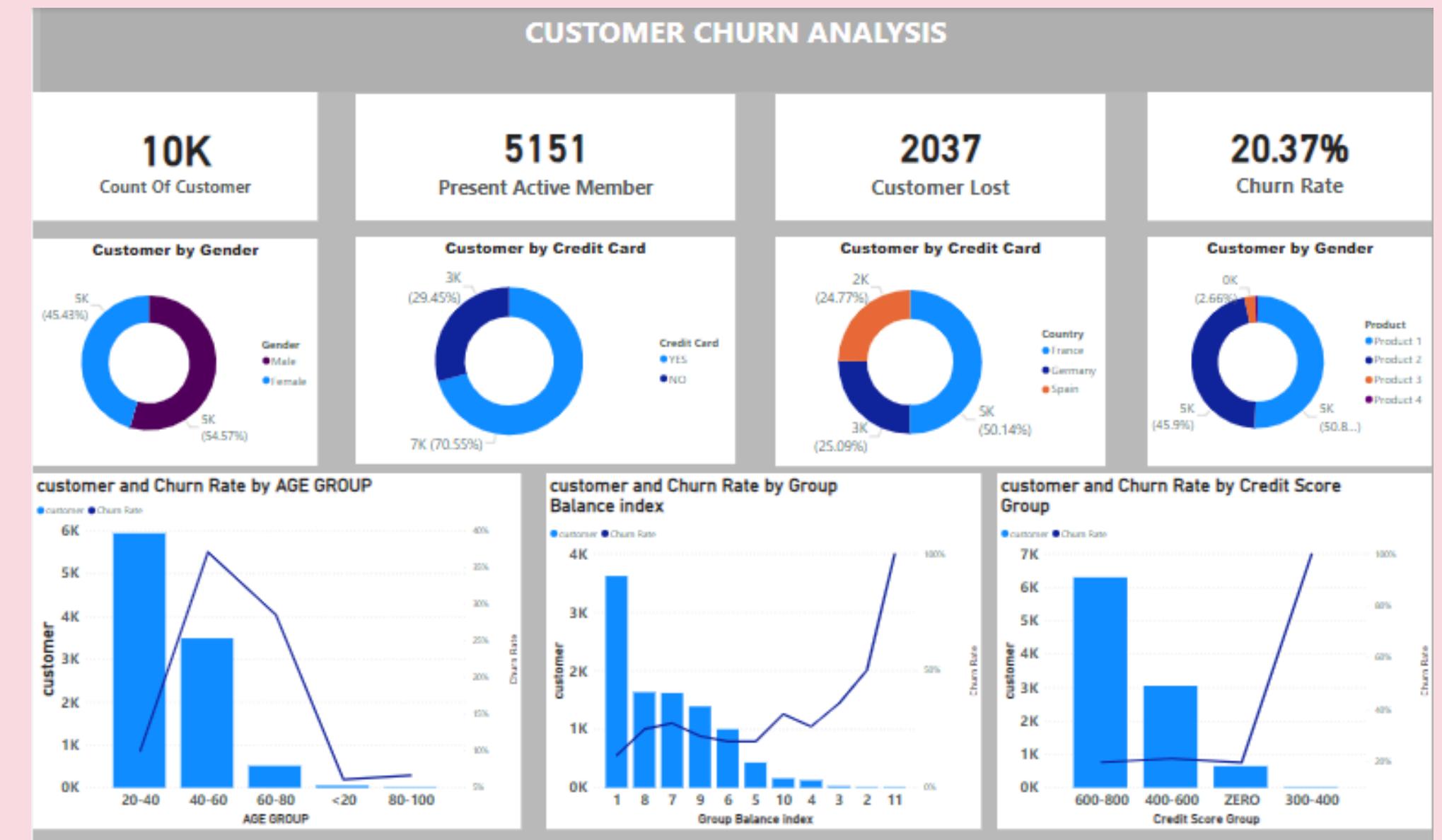


Dataset Overview

Description: The dataset includes 9 features: demographics, subscription details, payment method, and churn status (target).

Key Features:

- Customer ID (Unique identifier)
- Age, Gender (Demographics)
- Subscription Plan (Basic, Standard, Premium)
- Monthly Usage Hours (User engagement)
- Tenure (Months with service)
- Monthly Fee (Cost)
- Payment Method (Credit Card, PayPal, etc.)
- Churn (1 = churned, 0 = active)



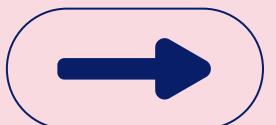
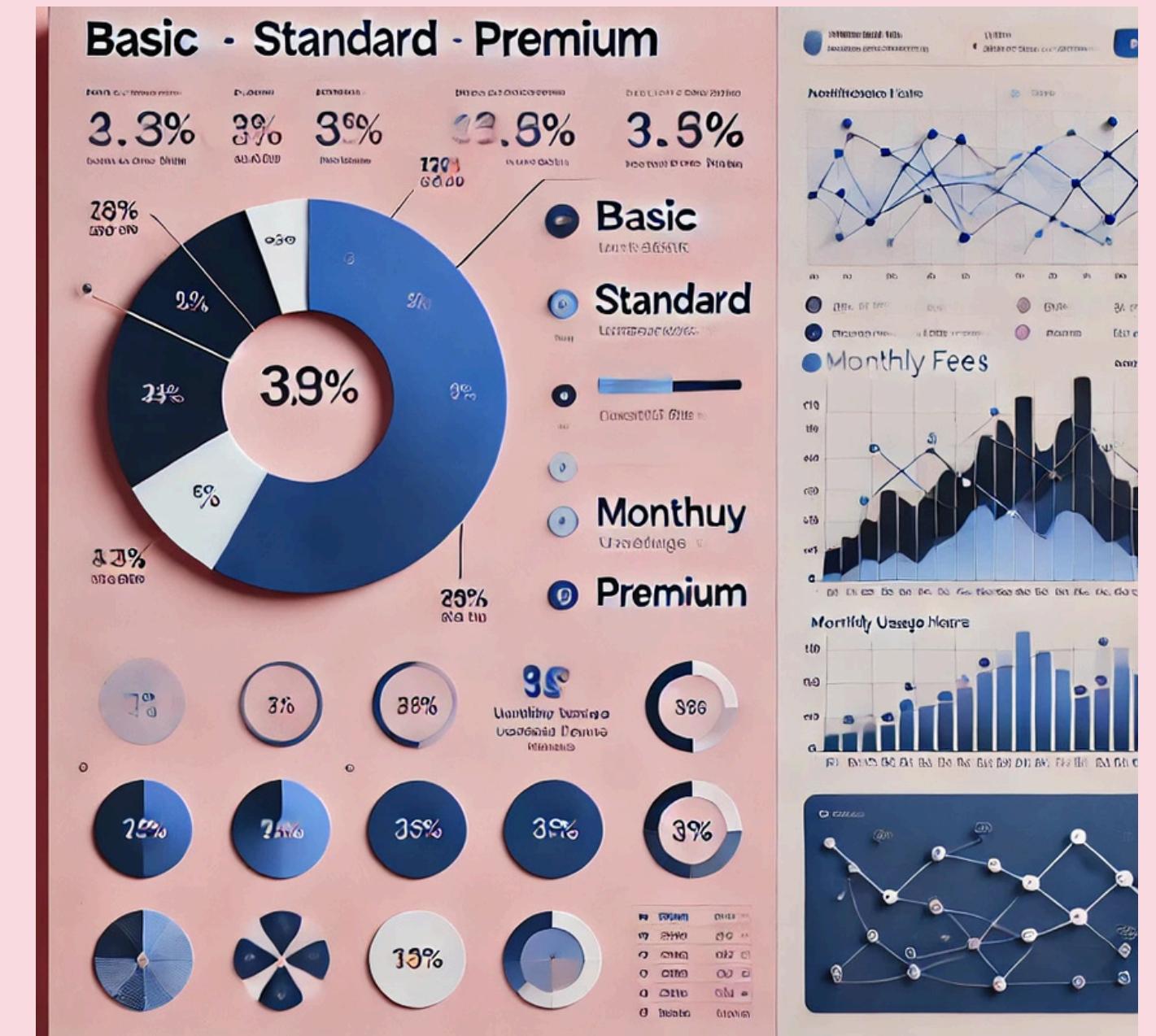
Data Exploration & Preprocessing

Exploratory Data Analysis (EDA):

- Analyze distributions of features like age, subscription plans, and tenure.
- Identify patterns in churned vs. active users.

Preprocessing Steps:

- Handle missing data.
- Encode categorical variables (Gender, Subscription Plan).
- Normalize numerical features (Monthly Usage Hours, Monthly Fee).



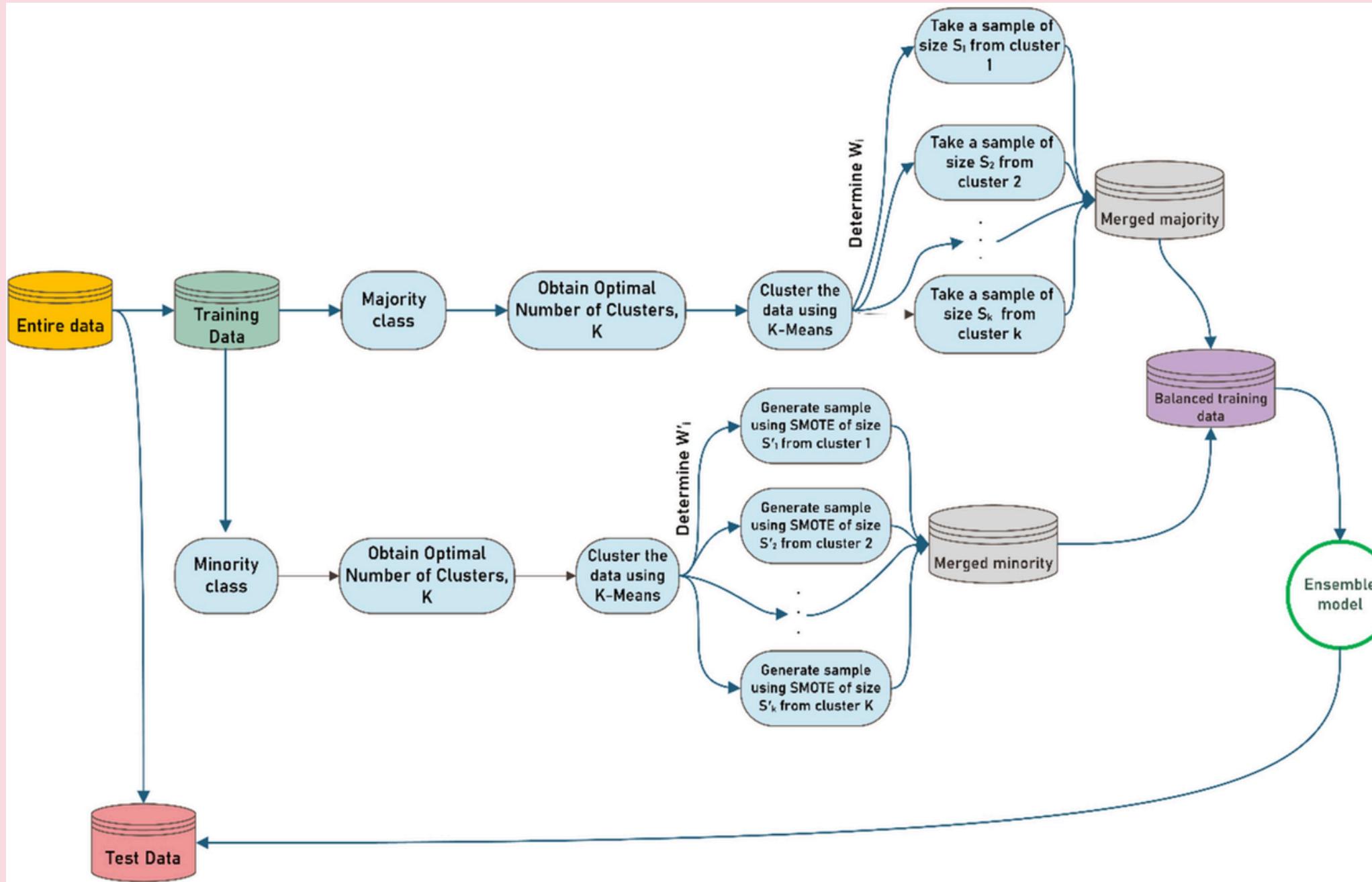
Handling Imbalanced Data

Problem:

- Fewer customers churn compared to active ones, creating class imbalance.

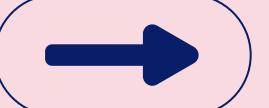
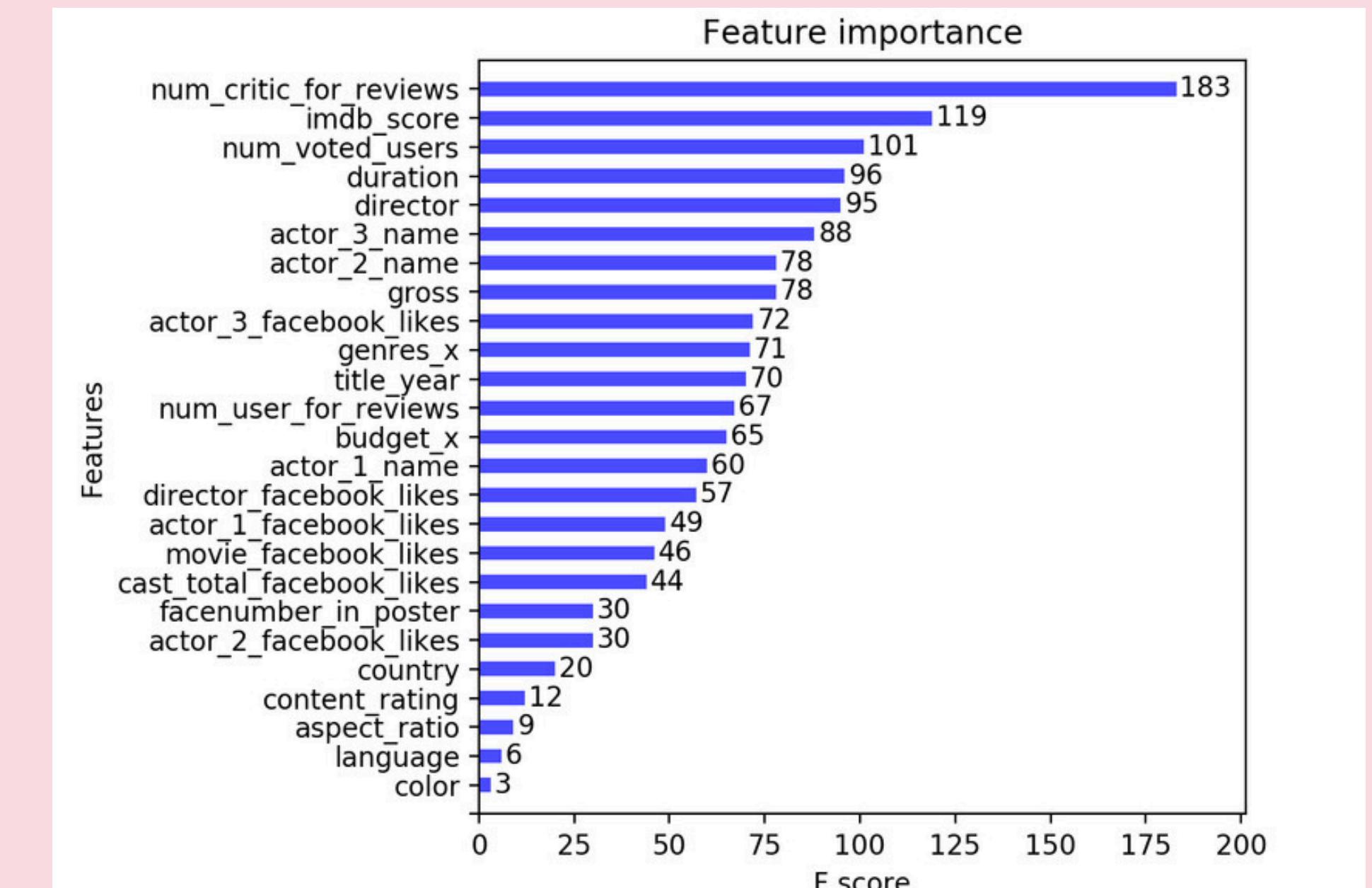
Solution:

- **SMOTE (Synthetic Minority Over-sampling Technique):** Generates synthetic samples for the minority class.
- **Class Weights:** Adjust model training to give more importance to churned customers.
- **Oversampling / Undersampling:** Balance classes by duplicating minority samples or reducing majority samples.



Feature Importance & Model Selection

- **Approach:** Use models like Random Forest, XGBoost, and Logistic Regression.
- **Feature Importance:** Top predictors include Tenure, Monthly Usage Hours, and Subscription Plan.
- **Evaluation Metrics:** Focus on Recall and F1-Score due to class imbalance.



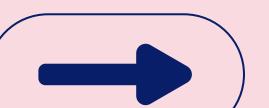
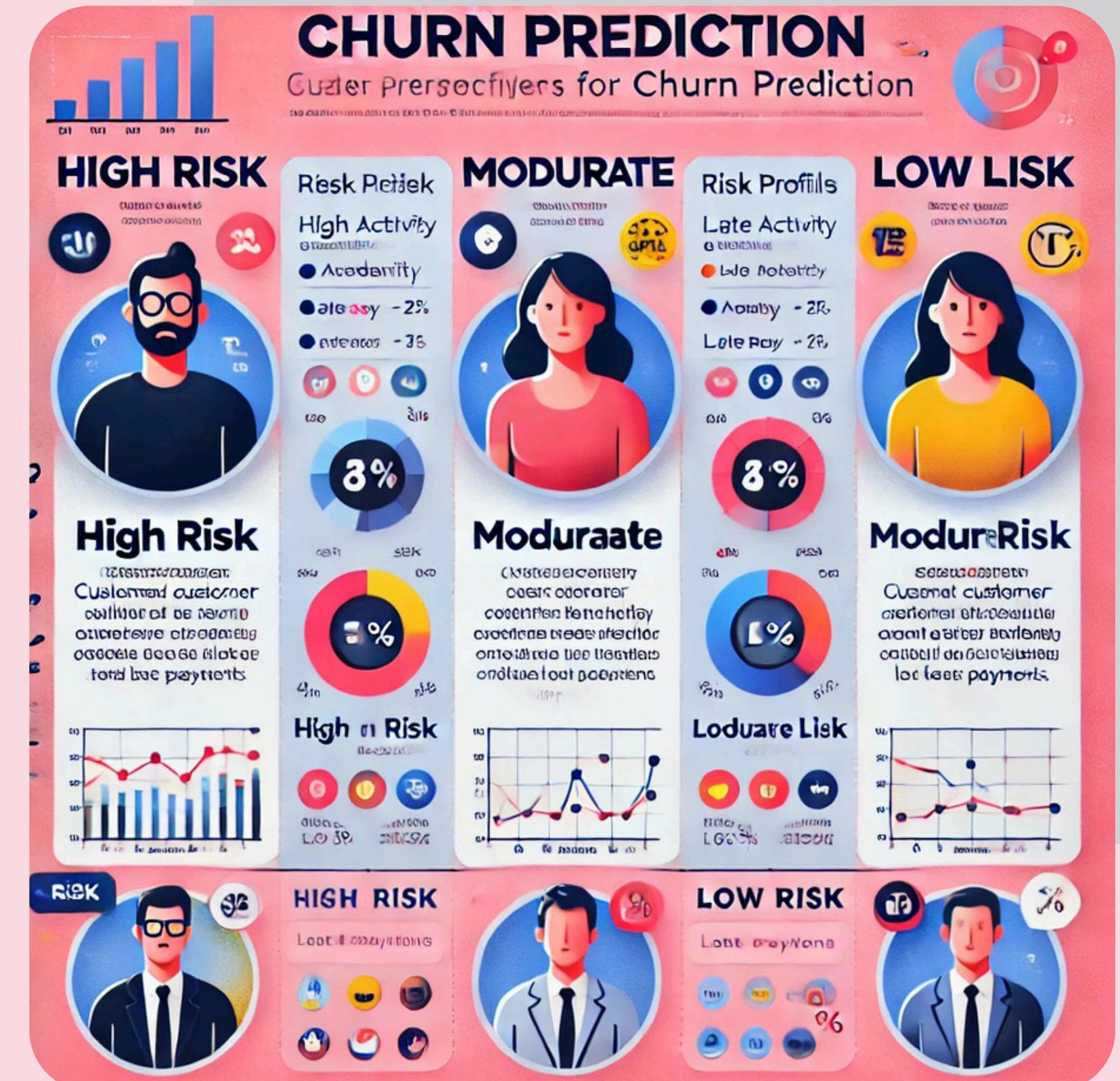
Explaining Model Predictions to Non-Technical Teams

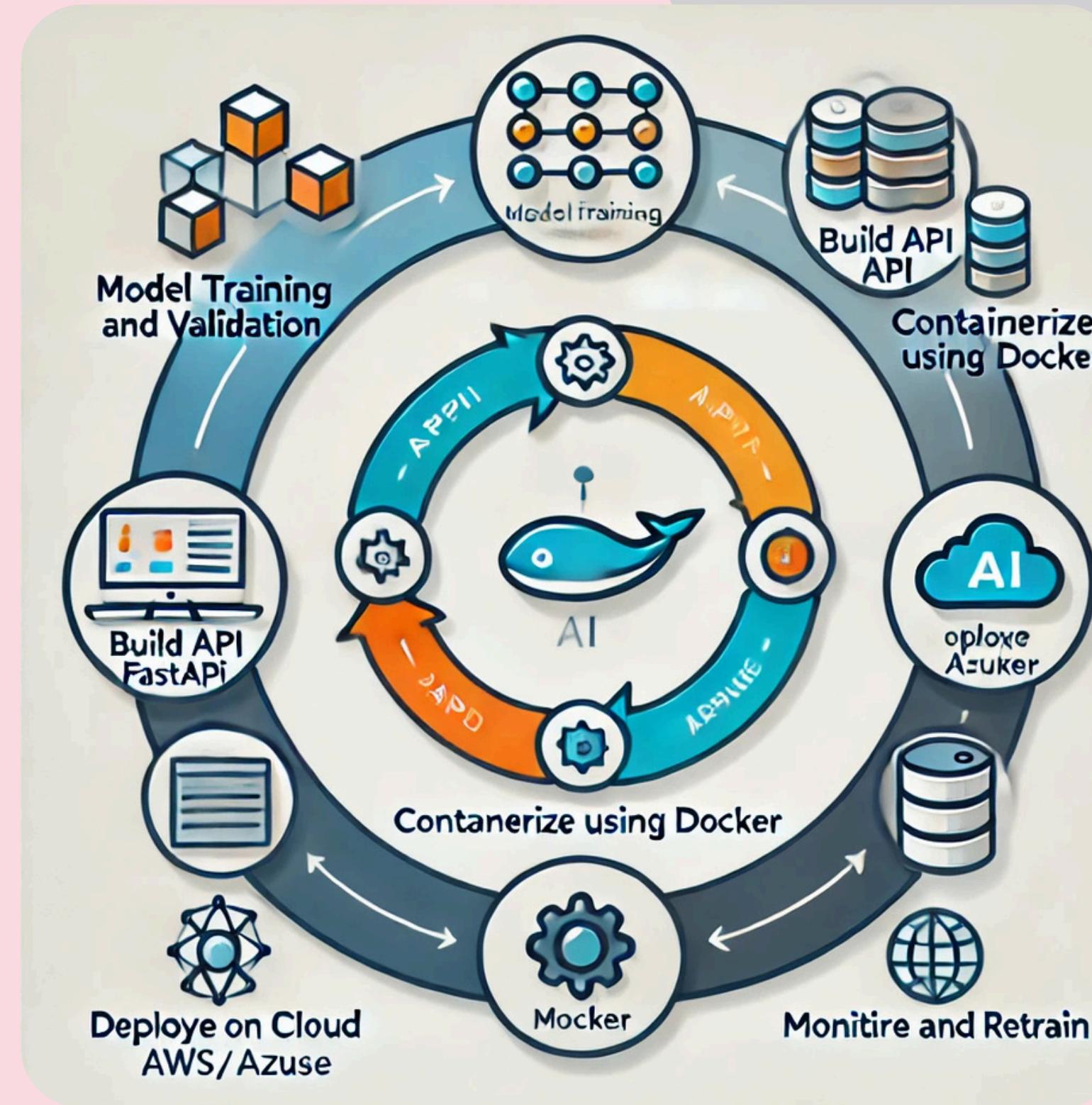
Simplified Explanation:

- The model predicts churn based on customer activity, payment behavior, and subscription history.
- Example: Customers with low usage and short tenure are flagged as at risk.

Business Impact:

- Helps tailor retention offers and improve customer engagement.





Deployment Strategy

Steps to Deploy:

- Model training and validation.
- Build an API (Flask/FastAPI) to serve predictions.
- Containerize using Docker for portability.
- Deploy on cloud platforms (AWS, Azure).
- Monitor model performance and retrain periodically.



Conclusion & Recommendations

Summary:

This customer churn prediction project used machine learning to identify at-risk customers for a streaming service. By analyzing historical data, we developed a model predicting churn based on features like age, subscription plan, and monthly usage. The model provides actionable insights, allowing the business to implement preventative measures to reduce churn and retain customers.

Key findings include:

- Important Predictors
- Handling Imbalanced Data
- Model Performance



Recommendations:

- Retention: Use the model to target at-risk customers with personalized offers.
- Engagement: Improve onboarding for high-usage users and content for low-usage users.
- Model: Retrain regularly and add new features for better accuracy.
- Marketing: Personalize campaigns using segmentation.
- Collaboration: Share insights with the product team to improve engagement.
- Long-Term: Use feedback to refine the model and predict customer lifetime value.

THANK YOU

[Github](#) | [Linkedin](#)

10

