High Level Design (HLD)

NoMoreChurn – Telco Risk Intelligence

Date: 25/5/2015

Submitted By :-

Soumen Baidya

Gourisankar Maity

Disha Jaiswal

Document Version Control

Date Issued	Version	Description	Author
25/5/2025	1.0	First Version of Complete HLD	Soumen Baidya
			Gourisankar Maity
			Disha Jasiswal

Contain

Abstract

The telecommunication industry is highly competitive, with companies continuously seeking ways to retain their customers. One of the significant challenges faced by telecom providers is *customer churn*, where customers discontinue their service subscriptions. Identifying customers likely to churn, understanding the reasons behind their decisions, and taking strategic actions can reduce revenue loss significantly.

This project, titled "NoMoreChurn – Telco Risk Intelligence,"aims to predict customer churn using machine learning techniques and visualize insights through a comprehensive Power BI dashboard. The insights derived from the analysis help business stakeholders develop proactive strategies for customer retention.

To analyze telecom customer data to identify factors influencing churn.

* To build a machine learning-based prediction model to forecast customer churn probability.

* To evaluate and compare various ML algorithms for optimal performance.

* To present key insights through an interactive and user-friendly Power BI dashboard.

* To support data-driven decisions for reducing customer churn and improving business strategies.

Introduction

Why this High-Level Design Document?

- Purpose of HLD: to outline architectural and functional overview for non-technical and semi-technical audiences.

Scope

- What the system will and won't cover.
- Only test dataset in Power BI.
- Focus on customer churn modelling, predictions, insights.

General Description

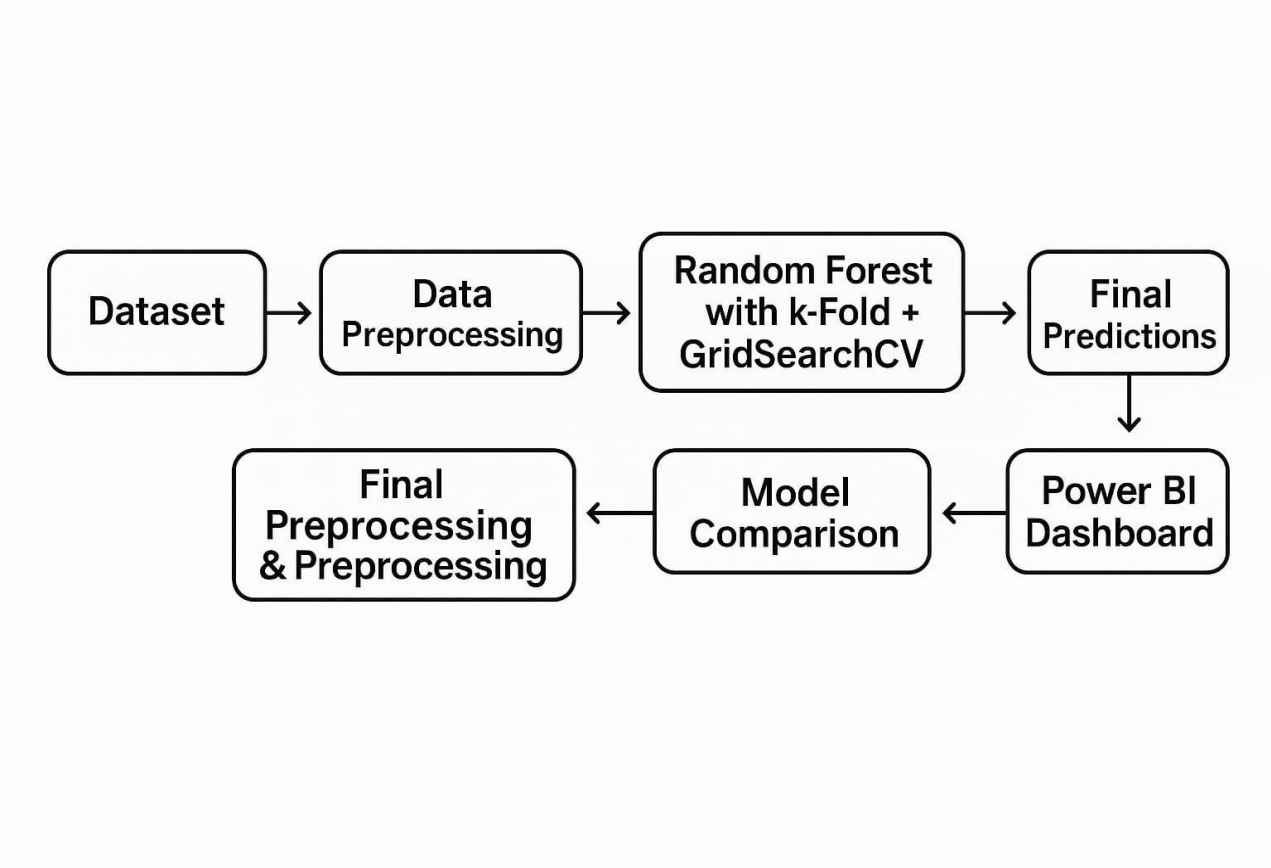
Product Perspective & Problem Statement

- Overview of the solution in business context.
- Problem: High customer churn in telecom.
- Solution: Use ML to predict churn, analyze causes, and visualize it via Power BI.

- **Tools Used**
- Python (EDA, modelling)
- XG Boost, scikit-learn
- Power BI (dashboard visualization)
- pandas, seaborn, matplotlib,

Design Details

Functional Architecture



Optimization

- Feature engineering (e.g., binning, encoding)
- Hyperparameter tuning using GridSearchCV
- Stratified sampling for test split
- Churn probability included for deeper analysis

KPIs

Key Performance Indicators

- Model Accuracy (from Python)
- Precision, Recall (shown in Power BI)
- Actual Churn Rate vs Predicted
- Risk Bands and high-risk customer identification
- Prediction vs Actual Confusion Matrix

Deployment

- Final CSV exported from Python
- Power BI dashboard built on top of it
- Local Power BI file, future plan for Power BI Cloud or embedding into web app (optional mention)