

Project Title: GHW Heart Failure Readmission Prediction

Domain: Healthcare Analytics

Difficulty Level: Intermediate to Advanced

Use the **Hospital Readmissions Data** from Kaggle: [Dataset Link](#)

Dataset Introduction: GHW Heart Failure Readmission Prediction

The dataset consists of heart failure patient records collected from various hospitals. It includes demographic details, medical history, lab test results, and hospital visit records. The primary objective is to analyze this data and build a model to predict whether a patient will be readmitted after being discharged.

Dataset Overview

- **Total Records:** [Varies, check dataset]
- **Target Variable:** Readmission Status (1 = Readmitted, 0 = Not Readmitted)
- **Features:**
 - **Demographics:** Age, Gender, Race
 - **Medical History:** Hypertension, Diabetes, Chronic Diseases
 - **Lab Results & Vitals:** Blood Pressure, Creatinine Levels, Sodium Levels
 - **Hospitalization Details:** Length of Stay, Number of Prior Admissions, Discharge Type
 - **Medications & Treatment:** Prescribed Drugs, Procedures

Key Challenges in the Data

- **Missing or incomplete values** in some patient records.
 - **Potential class imbalance** (fewer patients readmitted vs. not readmitted).
 - **Influence of multiple factors** like lifestyle, comorbidities, and prior hospital visits.
-

Scenario

Hospitals and healthcare providers face challenges in reducing patient readmission rates, especially for heart failure patients. High readmission rates lead to increased healthcare costs, resource strain, and lower patient outcomes. Predicting which patients are at risk of readmission can help hospitals provide better post-discharge care and prevent unnecessary hospital visits.

Problem Statement

You are working as a Data Scientist in a healthcare analytics firm. Your task is to develop a **predictive model** that identifies patients who are at high risk of being readmitted after a heart failure-related hospitalization. The goal is to provide insights that can help healthcare professionals take preventive measures.

Expectations

- Analyze and preprocess the given dataset to extract meaningful patterns.
- Perform **Exploratory Data Analysis (EDA)** to understand key factors influencing readmission.
- Build and evaluate machine learning models to predict patient readmission.
- Provide a report with insights, model performance, and recommendations.

Final Deliverables:

1. A well-documented Jupyter Notebook or Python script.
2. A short report (ppt) summarizing key findings, methodology, and model performance.

Evaluation will be based on:

- Data handling and preprocessing quality
- Depth of analysis and feature selection
- Model performance and justification
- Clarity of insights and recommendations