**TASK STATUS PROGRESS UPDATE**

**Task 1: Heart Disease Dataset Acquisition** ✅ **(Completed).**

* ☑ Government accredited open-source healthcare dataset website.

**Task 2: Heart Disease Dataset Cleaning & Preprocessing** ✅ **(Completed).**

**Objective:** Ensure the dataset is clean and ready for EDA.  
**Key Actions:**

* ☑ Remove duplicates.
* ☑ Handle missing values.
* ☑ Standardize formats.

**Task 3: Data Transformation** ✅ **(Completed).**

**Task 4: Loading Cleaned & Processed Version of Heart Disease Dataset** ✅ **(Completed).**

**Task 5: Exploratory Data Analysis 🔄 (Partially Completed)**

**Objective:** Implementation of EDA before data modeling.  
**Key Actions:**

* ☑ Renaming column header, getting summary information of the dataset.
* ☑ Checking & removing any duplicate values and empty cells.
* ☑ Addition of new column from existing column in the data set.
* ☐ Basic Visualization using matplotlib & seaborn python library.
* ☐ Resetting dataset index using python to save output in Excel & CSV format.
* ☐ Modeling.

**Task 6: Data Modeling 🔘 (In Progress).**  
**Objective:** Choosing the best fit algorithms or model to be used on the dataset.

**Key Actions:**

* ☑ **Evaluate different algorithms** (**📉** Linear Regression, **🌲** Decision Tree & **🌲🌲🌲** Random Forest) — in testing phase
* ☐ **Select and justify the best model** based on dataset characteristics.
* ☐ **Model Implementation** after selecting the best fit model.

**Task 7: Data Visualization & Communication 🔘 (In Progress).**  
**Task 8: Project Review ❌.**

**Task 9: Project Upload to GitHub ❌.**

**Task 10: Project Task Report Submission ❌.**