Documentation : **Predictive Modeling for Patient Eligibility**

1**. Introduction**

  Problem 1 - The development of drugs is critical in providing therapeutic options for patients suffering from chronic and terminal illnesses. "Target Drug", in particular, is designed to enhance the patient's health and well-being without causing dependence on other medications that could potentially lead to severe and life-threatening side effects. These drugs are specifically tailored to treat a particular disease or condition, offering a more focused and effective approach to treatment, while minimising the risk of harmful reactions.

The objective in this assignment is to develop a predictive model which will predict whether a patient will be eligible*** for "Target Drug" or not in next 30 days. Knowing if the patient is eligible or not will help physician treating the patient make informed decision on the which treatments to give.

  - the task involves predicting patient eligibility for a specific drug based on various features.

2. **Data Preprocessing:**

  data preprocessing in building accurate predictive models.

  steps involved in data preprocessing, including:

  Loading and exploring the dataset.

  Handling missing values using imputation.

  Feature scaling using StandardScaler.

3. Feature Engineering:

  process of feature engineering, which involves creating new features or modifying existing ones to improve model performance.

feature engineering techniques applied to your dataset, such as:

   - Calculating time-related features (TimeSinceLastEvent, TimeSinceFirstEvent, etc.).

   - One-hot encoding categorical variables.

4. Model Training

 model training and evaluation.

Describe the machine learning algorithm used (Logistic Regression) and – because target is binary classification

steps for model training:

 Splitting the dataset into training and validation sets.

 Fitting the model on the training data.

 Evaluating the model's performance using F1 score on the validation set.

5.Predictions on Test Data

 process of making predictions on the test dataset.

 using the same preprocessing steps on the test data as on the training data.

 Applying scaling and imputation to the test data.

 Making predictions using the trained model.