**Assignment 2: Logistic Regression Implementation**

Problem Statement

Your task is to implement a logistic regression model to predict the probability of disease occurrence based on health metrics.

Approach

* **Step 1:** Understand the dataset and the problem. Identify input features and the target variable.
* **Step 2:** Preprocess the data: normalize the features, split the dataset into training and testing sets.
* **Step 3:** Implement the logistic regression model. Use the sigmoid function for the hypothesis.
* **Step 4:** Compute the cost function using cross-entropy loss.
* **Step 5:** Apply gradient descent to minimize the cost function and find the best parameters.
* **Step 6:** Use the trained model to predict disease occurrence and evaluate its performance.

Hints

* Dataset is provided with this Assignment.
* Feature normalization can significantly improve the convergence of gradient descent.
* Consider using sklearn for model evaluation metrics like accuracy, precision, and recall.

Grading Rubric (Total: 100 Points)

* **Data Preprocessing (15 points):** Effective preprocessing and feature normalization.
* **Model Implementation (20 points):** Correct implementation of the logistic regression model and sigmoid function.
* **Cost Function and Optimization (25 points):** Accurate implementation of the cost function and effective optimization using gradient descent.
* **Prediction Accuracy (20 points):** Accuracy of the model in predicting disease occurrence on the test set.
* **Model Evaluation (20 points):** Comprehensive evaluation of the model using relevant metrics.