Predicting Hospital Readmission Using a Simple Machine Learning Model

Context: You are tasked with predicting whether a patient is likely to be readmitted to the hospital within 30 days based on historical patient data.

Task Overview:

1. **Dataset Selection & Preprocessing:**
    - Use platforms like Kaggle or Hugging Face to find a suitable dataset.
    - Preprocess the data by handling missing values, encoding categorical variables, and performing basic feature engineering.

2. **Model Building:**
    - Build a Logistic Regression model using Python (libraries like Scikit-learn) to classify whether a patient will be readmitted.
    - Split the data into 70% training and 30% testing.

3. **Model Evaluation:**
    - Evaluate the model using metrics like Accuracy, Precision, Recall, and F1-score.
    - Briefly explain why you chose Logistic Regression and how it performs on this task.

4. **Theoretical Task:**
    - Convert the Logistic Regression mathematical formula into code (in Python) and explain how each term of the formula maps to your implementation.

Deliverables:

- **Code:**
    - Submit a Python notebook (.ipynb) with all your code, from data preprocessing to model evaluation.
    - Make sure the notebook is well-documented, with comments explaining each step.

- **Report (Max 2 Pages):**

    - A brief report covering:
        - Key preprocessing steps taken.
        - Model choice and the rationale behind it.
        - Performance metrics of the model.
        - Theoretical explanation of the chosen model and how it translates mathematical formulas into algorithms.
        - Suggested improvements to your model and why they might work.

Bonus (Optional, but encouraged):

- If you're familiar with cloud platforms (AWS, Google Cloud, etc.), try deploying your model on a cloud service and provide a URL where we can see it in action. A simple API endpoint is sufficient.