Logistic Regression Project

AIML with Python Program

Project Overview

Students are required to complete a Logistic Regression project using a real-world dataset. This project was discussed in the 14th live session and must be completed as per expert instructions.

Dataset Selection

Choose any dataset from Kaggle using the link below:

🔗 **https://www.kaggle.com/datasets**

Dataset Guidelines

- **Dataset should have a clear target variable**
- **Prefer datasets with more numerical columns**
- **Avoid extremely large datasets**
- **Dataset must be suitable for classification**

Project Steps (Mandatory)

1. Data Loading & Understanding

- **Load the dataset**
- **Use `head()`, `info()`, `describe()` for data understanding**

2. Data Preprocessing

- **Handle missing values**
- **Perform encoding (if required)**
- **Apply scaling (if required)**

3. Exploratory Data Analysis (EDA)

- **Perform basic analysis and visualization**
- **Understand feature relationships**

4. Model Building

- **Split dataset into train and test**
- **Apply Logistic Regression**
- **Train the model**

5. Model Evaluation

- **Calculate accuracy**
- **Generate confusion matrix**
- **Interpret results**

6. Conclusion

- **Summarize insights**
- **Mention model performance**
- **Explain learnings**

Submission Format

- **File type: .ipynb**
- **Clear headings and comments required**
- **Outputs must be visible**
- **Final conclusion mandatory**

Submission Deadline

30th January 2026