**🧑‍💻 Modular Machine Learning Pipeline**

**📊 Dataset Used**

* **Dataset**: Breast Cancer Wisconsin (Diagnostic) Dataset
* **Source**: sklearn.datasets.load\_breast\_cancer()
* **Description**: The dataset contains **569 samples** with **30 features**, representing tumor characteristics. Target variable indicates whether the tumor is **malignant (1)** or **benign (0)**.

**⚙️ Steps Included in the Pipeline**

1. **Data Loading Module**
   * Loads dataset from sklearn.
   * Handles missing values (if any).
2. **Preprocessing Module**
   * Encodes categorical features (if present).
   * Scales features using StandardScaler.
   * Splits dataset into training and testing sets.
3. **Model Training Module**
   * Supports supervised ML models (Logistic Regression, Decision Tree, Random Forest).
   * Ensures reproducibility using random\_state.
4. **Evaluation Module**
   * Computes Accuracy, Precision, Recall, and F1-score.
   * Displays **Confusion Matrix**.
   * Prints full **Classification Report**.

**▶️ How to Run the Code**

1. Clone this repository:
2. git clone https://github.com/your-username/ml-pipeline.git
3. cd ml-pipeline
4. Install dependencies:
5. pip install -r requirements.txt
6. Run the pipeline script:
7. python ml\_pipeline.py
8. Results will include:
   * Evaluation metrics (Accuracy, Precision, Recall, F1-score).
   * Confusion Matrix plot.
   * Classification report.