

1. Gradient Boosting Regression on the California Housing Dataset

- **Dataset:** Use the *California Housing Dataset* from `sklearn.datasets`.
- **Tasks:**
 1. Load and split the dataset into training and testing sets (70% train, 30% test).
 2. Initialize a Gradient Boosting Regressor with `n_estimators=200` and experiment with `learning_rate` values of 0.01, 0.1, and 0.2.
 3. For each learning rate, train the model and evaluate it using Mean Squared Error (MSE) and R^2 score.
 4. Plot MSE as a function of the number of boosting iterations for each learning rate.
 5. Summarize how learning rate impacts model performance, explaining underfitting or overfitting at different rates.

2. Gradient Boosting Classification on the Wine Quality Dataset

- **Dataset:** Use the *Wine Quality Dataset* from the UCI repository or Kaggle.
- **Tasks:**
 1. Load the Wine Quality dataset, and convert the `quality` column into a binary classification: “high” (≥ 7) and “low” (< 7).
 2. Split the data into training and testing sets (75% train, 25% test).
 3. Initialize a Gradient Boosting Classifier with `n_estimators=150`.
 4. Experiment with different `max_depth` values (`max_depth=2`, `max_depth=4`, and `max_depth=6`) and evaluate each model using accuracy and F1-score.
 5. Display feature importance and create a bar plot of the most influential features.
 6. Explain how tree depth affects model performance and discuss which features are most predictive of wine quality.

3. Gradient Boosting with Early Stopping on Breast Cancer Dataset

- **Dataset:** Use the *Breast Cancer Dataset* from `sklearn.datasets`.
- **Tasks:**
 1. Load the Breast Cancer dataset and split it into training, validation, and testing sets (60% train, 20% validation, 20% test).
 2. Train a Gradient Boosting Classifier with `n_estimators=500` and `learning_rate=0.1`, using the validation set for early stopping.
 3. Use early stopping criteria based on validation accuracy with a patience of 20 rounds.
 4. Compare the test accuracy, F1-score, and number of iterations with and without early stopping.
 5. Summarize the effect of early stopping on performance and training time.