

Aim:

To implement, train, and compare the performance of various fundamental machine learning algorithms — namely **Least Squares Regression**, **Perceptron**, **Support Vector Machine (SVM)**, **Logistic Regression**, and **PCA-based Perceptron** — on the **Iris dataset**, and visualize their decision boundaries and accuracies.

Algorithm / Methodology:

Step 1: Load and Preprocess Data

1. Load the **Iris dataset** using `sklearn.datasets.load_iris()`.
2. Split data into **training (70%)** and **testing (30%)** sets (stratified).
3. Apply **StandardScaler** for normalization where required.

Step 2: Implement Models

1. Manual Least Squares Regression (Single Feature)

- Use one feature (sepal length) to fit a simple regression line:

$$y = mx + c \quad y = mx + c$$

- Predict target classes by rounding and clipping continuous outputs.
- Evaluate accuracy and classification report.

2. Perceptron (Full Features)

- Apply Perceptron on all features using standardized data.
- Train and predict using `Perceptron()` from `sklearn.linear_model`.
- Measure performance metrics.

3. Support Vector Machine (SVM)

- Use `SVC(kernel='rbf')` with **GridSearchCV** to find best hyperparameters (`C`, `gamma`).
- Evaluate accuracy and best parameter combination.

4. PCA + Perceptron

- Perform **Principal Component Analysis (PCA)** to reduce features to **2 and 3 dimensions**.
- Visualize 2D and 3D scatter plots for class separation.
- Train Perceptron using the top-2 PCA components and evaluate accuracy.

5. Logistic Regression

- Train Logistic Regression (`LogisticRegression(max_iter=5000)`).
- Evaluate accuracy and classification report.

Step 3: Visualization

- **PCA Visualization:** 2D and 3D projections of data after PCA.
- **Decision Boundaries:** Plot Perceptron and SVM decision regions using first two original features.
- **Accuracy Comparison:** Bar chart comparing all models' accuracies.

Result:

SVM (RBF kernel) and Logistic Regression provide the best classification performance. SVM and Logistic Regression achieved the highest accuracy (91.11%).