

Assignment -7

- 1) The **Iris dataset** is a classic example for demonstrating classification algorithms. It consists of 150 samples of iris flowers belonging to three species: *Setosa*, *Versicolor*, and *Virginica*, with four input features (sepal and petal length/width). Use SVC from sklearn.svm on the Iris dataset and follow the steps below:
 - a) Load the dataset and perform train–test split (80:20).
 - b) Train three different SVM models using the following kernels: Linear, Polynomial (degree=3), RBF
 - c) Evaluate each model using:
 - Accuracy
 - Precision
 - Recall
 - F1-Score
 - d) Display the confusion matrix for each kernel.
 - e) Identify which kernel performs the best and why.
- 2) SVM models are highly sensitive to the scale of input features. When features have different ranges, the algorithm may incorrectly assign higher importance to variables with larger magnitudes, affecting the placement of the separating hyperplane. **Feature scaling** ensures that all attributes contribute equally to distance-based computations, which is especially crucial for kernels like RBF or polynomial.
 - a) Use the **Breast Cancer dataset** from sklearn.datasets.load_breast_cancer.
 - b) Train an SVM (RBF kernel) model **with** and **without** feature scaling (StandardScaler). Compare both results using:
 - Training accuracy
 - Testing accuracy
 - c) Discuss the effect of feature scaling on SVM performance.