

31/10/25

LAB-2

Implement a Classifier using Open-Source Dataset.

→ AIM:-

To implement a machine learning classifier using Open Source dataset (Iris) and evaluate its performance.

→ OBJECTIVE:-

- To understand the process of building a classifier using Scikit-learn.
- To load and explore an open-source dataset.
- To preprocess data using normalization or standardization.
- To apply the K-nearest neighbors (KNN) classification algorithm.
- To evaluate the performance of the classifier using accuracy and confusion matrix.

→ PSEUDO CODE:-

START

1. Import required libraries
2. Load the Iris dataset
3. Split the dataset into training and testing sets.
4. Standardize the features using StandardScaler.
5. Initialize KNN classifier with $K=3$
6. Train the classifier on the training data.
7. Make prediction on test data.
8. Evaluate the model.

OBSERVATIONS:-

- Dataset : This dataset has 150 samples and 4 features with 3 classes.
- Data Split : 10% Training, 30% Testing
- Preprocessing : Data is scaled using StandardScaler
- Classifiers Used : KNN $k=3$
- Accuracy : ~ 0.95 to 1.00
- Confusion Matrix : Mostly correct predictions
- Visualization : Heatmap shows correct classification in diagonal entries

BETTER CODE :-

- a. calculate Accuracy
- b. generate classification Report
- c. Create confusion Matrix
- d. Visualize confusion matrix using Seaborn heatmap.

END

RESULT:-

- The KNN classifier was successfully implemented on Iris dataset
- The model achieved an accuracy of ~96% - 100% depending on the data split.
- The confusion matrix confirms that the model performs well in classifying the flower types.