07/08/25

LAB-3

Study the Classifiers with respect to Stastical Parameters

→ AIM:-

To implement and compare the performance of different machine learning classification algorithms on Iris dataset using confusion matrices and accuracy metrics.

→ OBJECTIVES :-

> To apply preprocessing on Iris dataset.
> To train multiple classifiers.
> To evaluate each model's performance
> To compare models visually.

→ PSEUDOCODE :-

1. Import necessary libraries
2. Load the dataset.
3. Split dataset to training and testing.
4. Standardize using StandardScalar
5. Initialize Classifier:
   > Logistic Regression, KNN, Decision tree, SVM
6. For each classifier:
   > Train, predict, calculate accuracy & generate confusion matrix
7. Plot confusion matrix Seaborn heatmap
8. Bar chart to compare accuracies.

OBSERVATIONS:

Logistic Regression
Accuracy: 0.9556
Confusion Matrix: [[11 0 0]
[0 15 0]
[0 2 17]]

KNN
Accuracy: 0.911
Confusion Matrix: [[11 0 0]
[0 15 0]
[0 4 15]]

Decision tree
Accuracy: 0.911
Confusion Matrix: [[11 0 0]
[0 15 0]
[0 4 15]]

SVM
Accuracy: 0.9556
Confusion Matrix: [[11 0 0]
[0 15 0]
[0 2 17]]

* Plotted confusion matrix for each
* Bar chart for accuracies.

Confusion Matrix: A table that summarize the performance of classification model by showing correct and incorrect predictions across all classes.

Accuracy: The ratio of correctly predicted observations to the total observations.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

TP: True positives
TN: True negatives
FP: False positive
FN: False negative

RESULT:

> All models performed well, with Logistic Regression and SVM achieving the highest accuracy. 95.56%
> KNN and Decision Tree achieved 91.11%, showed misclassification in class 2.
> Confusion matrix helped identify that most errors occured when predicting the virginica class.