

[illegible]

07-08-25

### 3. Study of the classifiers with respect to statistical parameters

Aim: to study and compare the performance of different classifiers on an open source data set

#### description

① Accuracy : fraction of predicted observations

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

② Precision : Ratio of correctly predicted positive to the total predicted positives

$$\text{Precision} = \frac{TP}{TP + FP}$$

③ Recall =  $\frac{TP}{TP + FN}$

④ F1-score =  $\frac{2 \times \text{precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$

⑤ Confusion matrix :- A 2D table showing values of TP, FP, FN

#### Algorithm

- ① Import the libraries
- ② Load the open source dataset
- ③ Split the dataset into training (80%) and testing 20%.
- ④ Select and apply model
- ⑤ Train the model with training data
- ⑥ Predict the output on the test data

- ⑦ Compute evaluation metrics and confusion matrix
- ⑧ Repeat 4-5 for all classifiers
- ⑨ compare the results.

code :-

Step: 1 - Import libraries

Import numpy as np

Import pandas as pd

from sklearn.model\_selection import train\_test\_split

from sklearn.neighbors import KNeighborsClassifier

from sklearn.svm import SVC

Import matplotlib.pyplot as plt

Step: 2

iris = load\_iris()

X = iris.data

Y = iris.target

X\_train, X\_test, Y\_train, Y\_test = train\_test\_split(X, Y, test\_size=0.3, random\_state=42)

for name, cls in classifiers.items():

cls.fit(X\_train, Y\_train)

Y\_pred = cls.predict(X\_test)

print(f" {name} == ")

print(" Accuracy", round(accuracy\_score(Y\_test, Y\_pred), 4))

print(" Precision", round(precision\_score(Y\_test, Y\_pred,

print(" f1-score", round(f1\_score(Y\_test, Y\_pred, average='macro'), 4))

Output :-

Accuracy : 1.00

Precision : 1.00

Recall : 1.00

F1-Score : 1.00

Confusion matrix :-

Setosa	19	0	0
versicolour	0	13	0
Virginica	0	0	13
	Setosa	versicolour	Virginica

cm = Confusion\_matrix(y\_test, y\_pred)

disp = Confusion\_matrix\_display(Confusion\_matrix = cm,

display\_labels = Poisson\_target\_names)

des.plot = Cmap = plt.cm.Blues)

plt.title ("Confusion matrix - Poisson")

plt.show

Result:- Implemented different types of ~~data~~ classifiers

On the Poisson dataset successfully. And the accuracy

and precision is 1.

~~off~~



```
[ ]
▶ # Step 1: Import Libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.datasets import load_iris
from sklearn.linear_model import LogisticRegression
from sklearn.model_selection import train_test_split
from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

# Step 2: Load the Iris Dataset
iris = load_iris()
X = pd.DataFrame(iris.data, columns=iris.feature_names)
y = pd.Series(iris.target, name='species')

# Step 3: Split the Dataset
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=42)

# Step 4: Train the Logistic Regression Model
model = LogisticRegression(max_iter=200)
model.fit(X_train, y_train)

# Step 5: Make Predictions
y_pred = model.predict(X_test)

# Step 6: Evaluate the Model
print("Accuracy:", accuracy_score(y_test, y_pred))
print("\nClassification Report:\n", classification_report(y_test, y_pred))
print("\nConfusion Matrix:\n", confusion_matrix(y_test, y_pred))
```

### # Step 7: Visualize the Confusion Matrix

```
plt.figure(figsize=(6,4))
sns.heatmap(confusion_matrix(y_test, y_pred), annot=True, fmt='d', cmap='Blues',
xticklabels=iris.target_names, yticklabels=iris.target_names)
plt.title("Confusion Matrix")
plt.xlabel("Predicted")
plt.ylabel("Actual")
plt.show()
```



Accuracy: 1.0

#### Classification Report:

	precision	recall	f1-score	support
0	1.00	1.00	1.00	19
1	1.00	1.00	1.00	13
2	1.00	1.00	1.00	13
accuracy			1.00	45
macro avg	1.00	1.00	1.00	45
weighted avg	1.00	1.00	1.00	45

Confusion Matrix

