

Expt No: 8

Date: 22-04-25

KNN: K-Nearest Neighbours Algorithm

Aim:

To implement the K-Nearest Neighbours (K) algorithm for classification using Python to classify Iris flower species based on sepal and petal measurements.

Algorithm:

1. Start the program
2. Import the required libraries (numpy, pandas, sklearn).
3. Load the Iris dataset (Iris.csv) file
4. Display dataset information and check for null values.
5. Separate the dataset into features and labels.
6. Split the data into training and testing sets.
7. Create a K NeighboursClassifier model
8. Train the model using `fit()` on training data.
9. Predict flower classes for test data using `predict()`
10. Evaluate model performance using ~~score~~ score(), confusion matrix(), classification report().
11. Stop the program.

Program :

```
import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.neighbors import KNeighborsClassifier
from sklearn.metrics import confusion_matrix,
classification_report.

df = pd.read_csv('Iris.csv')
df.info()
print(df['variety'].value_counts())
print(df.head())

features = df.iloc[:, :-1].values
label = df.iloc[:, 4].values

xtrain, xtest, ytrain, ytest = train_test_split(
    features, label, test_size = 0.2, random_state = 42)

model_knn = KNeighborsClassifier(n_neighbors = 3)
model_knn.fit(xtrain, ytrain)

print("Training Accuracy : ", model_knn.score(xtrain,
    ytrain))

print("Testing Accuracy : ", model_knn.score(xtest, ytest))

y_pred = model_knn.predict(xtest)
cm = confusion_matrix(ytest, y_pred)
print("\n Confusion Matrix : \n", cm)

print("\n Classification Report : \n", classification_report(
    ytest, y_pred))
```

Output :

Training Accuracy : 0.9583

Testing Accuracy : 1.0

Confusion Matrix :

$\begin{bmatrix} 10 & 0 & 0 \end{bmatrix}$

$\begin{bmatrix} 0 & 9 & 1 \end{bmatrix}$

$\begin{bmatrix} 0 & 0 & 10 \end{bmatrix}$

Classification Report :

	precision	recall	f1 - score	support
Setosa	1.00	1.00	1.00	10
Versicolour	1.00	0.90	0.95	10
Virginica	0.91	1.00	0.95	10
accuracy			0.97	30
macro avg	0.97	0.97	0.97	30
mi weighted avg	0.97	0.97	0.97	30

Result:

The KNN Classifier was successfully trained and tested on the Iris dataset. The model achieved 97% accuracy, effectively classifying Iris flowers into three species: Setosa, Versicolour and Virginica.