



## Experiment No: 03

**Code 01 :-**

```
import pandas as pd

from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score
from sklearn.neighbors import KNeighborsClassifier

# Load the dataset

file = "diabetes.csv"

df = pd.read_csv(file)

# Separate features and target variable

data = df.values

X, y = data[:, :-1], data[:, -1]

print(f"Feature shape: {X.shape}, Target shape: {y.shape}")

# Split the dataset into training and testing sets

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=1)

print(f"Training feature shape: {X_train.shape}, Testing feature shape: {X_test.shape}")

print(f"Training target shape: {y_train.shape}, Testing target shape: {y_test.shape}")

# Initialize and fit the KNN model

model = KNeighborsClassifier()

model.fit(X_train, y_train)
```

---

```
# Make predictions
```

```
y_predict = model.predict(X_test)
```

```
# Calculate and print accuracy
```

```
accuracy = accuracy_score(y_test, y_predict)
```

```
print("Dataframe:\n", df.head()) # Print only the first few rows for clarity
```

```
print("Accuracy: {:.2f}%".format(accuracy * 100))
```

### Code 01 Output :-

```
Feature shape: (768, 8), Target shape: (768,)
Training feature shape: (537, 8), Testing feature shape: (231, 8)
Training target shape: (537,), Testing target shape: (231,)
Dataframe:
   Pregnancies  Glucose  BloodPressure  SkinThickness  Insulin   BMI  \
0             6     148             72             35         0  33.6
1             1      85              66             29         0  26.6
2             8     183              64              0         0  23.3
3             1      89              66             23        94  28.1
4             0     137              40             35       168  43.1

   DiabetesPedigreeFunction  Age  Outcome
0                   0.627    50         1
1                   0.351    31         0
2                   0.672    32         1
3                   0.167    21         0
4                   2.288    33         1
Accuracy: 77.06%
```

### Code 02 :-

```
import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split
```

```
# Define arrays for X and y
```

```
X = np.array([
    [1, 2, 3],
    [4, 5, 6],
    [7, 8, 9],
    [10, 11, 12],
    [13, 14, 15],
```



```
[16, 17, 18],
[19, 20, 21],
[22, 23, 24],
[25, 26, 27],
[28, 29, 30]
])
y = np.array([1, 2, 3, 4, 5, 6, 7, 8, 9, 10])

# Converting to pandas DataFrame for consistency
df = pd.DataFrame(X, columns=['Feature1', 'Feature2', 'Feature3'])
df['Target'] = y

# Separating features and target variable
X = df.iloc[:, :-1] # Features
y = df.iloc[:, -1] # Target variable

# Splitting the dataset into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=0)

# Print the training and testing sets
print("X_train:")
print(X_train)
print("\nX_test:")
print(X_test)
print("\ny_train:")
print(y_train)
print("\ny_test:")
print(y_test)
```

```
X_train:
  Feature1  Feature2  Feature3
9         28         29         30
1          4          5          6
6         19         20         21
7         22         23         24
3         10         11         12
0          1          2          3
5         16         17         18
```

```
X_test:
  Feature1  Feature2  Feature3
2          7          8          9
8         25         26         27
4         13         14         15
```

```
y_train:
9      10
1       2
6       7
7       8
3       4
0       1
5       6
Name: Target, dtype: int64
```

```
y_test:
2       3
8       9
4       5
Name: Target, dtype: int64
```