Assignment No : 5

```
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.neighbors import KNeighborsClassifier
from sklearn.metrics import confusion_matrix, accuracy_score, precision_score, recall_score
df = pd.read_csv("diabetes.csv")
df = df.apply(pd.to_numeric, errors='coerce') df.fillna(df.mean(),
inplace=True)
X = df.drop('Outcome', axis=1)
y = df['Outcome'].astype(int)
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
scalar = StandardScalar()
X_train = scaler.fit_transform(X_train)
X_test = scaler.transform(X_test)

k = 5
knn = KNeighborsClassifier(n_neighbors=k)
knn.fit(X_train, y_train)
y_pred = knn.predict(X_test)
cm = confusion_matrix(y_test, y_pred)
print("Confusion Matrix:\n", cm)
```

output:

```
Confusion Matrix:
 [[79 20]
 [27 28]]
```

```
accuracy = accuracy_score(y_test, y_pred)
error_rate = 1 - accuracy
precision = precision_score(y_test, y_pred)
recall = recall_score(y_test, y_pred)

print(f'Accuracy: {accuracy * 100:.2f}%')
print(f'Error Rate: {error_rate * 100:.2f}%')
print(f'Precision: {precision:.2f}')
print(f'Recall: {recall:.2f}')
```

output:

Accuracy: 69.48%

Error Rate: 30.52%

Precision: 0.58

Recall: 0.51