



```
In [ ]: # Import Libraries
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.neighbors import KNeighborsClassifier
from sklearn.metrics import accuracy_score, classification_report

# Load Dataset
df = pd.read_csv("Dataset of Diabetes.csv")

# Independent Variable (BMI - second last column)
X = df.iloc[:, -2].values.reshape(-1, 1)

# Dependent Variable (Class - last column, Y/N)
y = df.iloc[:, -1].values

# Train-Test Split
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.2, random_state=42)

# Create KNN Model (k = 5)
model = KNeighborsClassifier(n_neighbors=6)

# Train Model
model.fit(X_train, y_train)

# Predict
y_pred = model.predict(X_test)
prediction = pd.DataFrame({'Actual': y_test, 'Predicted': y_pred})
print(prediction)

# Accuracy
accuracy = accuracy_score(y_test, y_pred)

print("Accuracy:", accuracy * 100, "%")
```

	Actual	Predicted
0	Y	Y
1	Y	Y
2	Y	Y
3	Y	Y
4	Y	Y
..
195	Y	Y
196	Y	Y
197	Y	Y
198	Y	Y
199	N	N

[200 rows x 2 columns]

Accuracy: 91.5 %