**Assignment No. 5**

PRN: 2019BTECS00056

Name: Rohan Bondre

**Aim:**

Use / extend the data analysis tool (menu driven GUI) developed in Assignment No. 1 to perform the following classification task :

1. Design and implement the following classifiers:

a) Regression classifier.

b) Naïve Bayesian Classifier.

c) *k*-NN classifier (Take *k* = 1,3,5,7)

d) Three-layer Artificial Neural Network (ANN) classifier (use back propagation). Plot error graph (iteration vs error).

2. Tabulate the results in confusion matrix and evaluate the performance of above classifier using following metrics:

a) Recognition rate

b) Misclassification rate

c) Sensitivity

d) Specificity

e) Precision & Recall

3. Use the following data sets from UCI machine learning repository :

a) IRIS

b) Breast Cancer

Run commands:

Go to folder location of main.py file and run command :

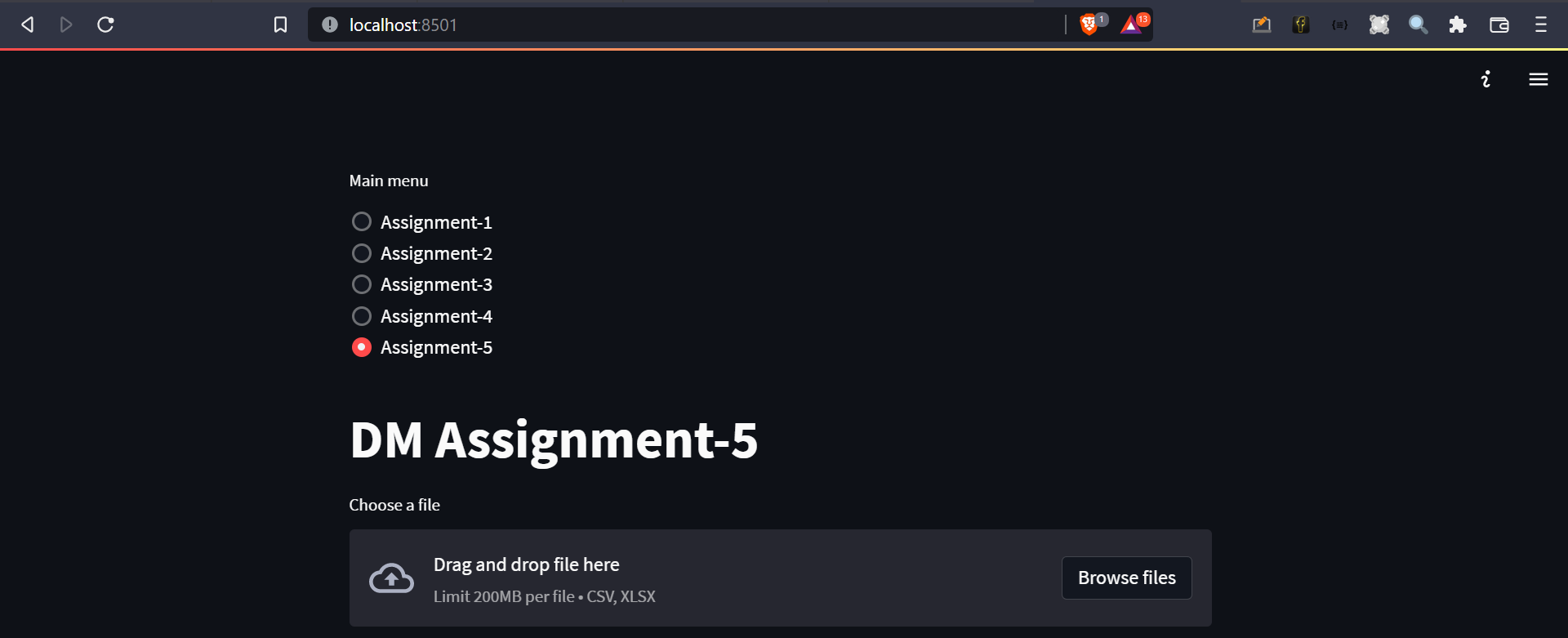
**streamlit run main.py**

**Procedure:**

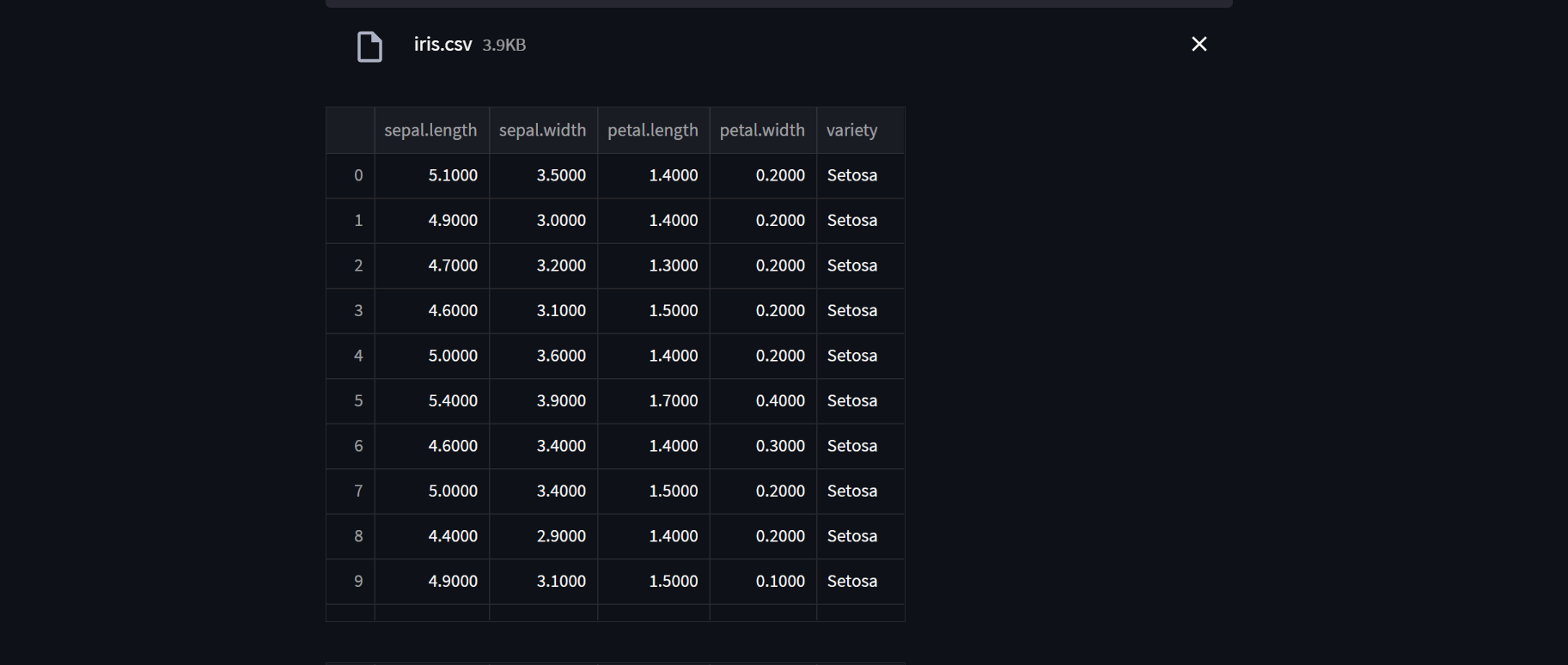
1. First divided dataset into x-train(70%) and x-test(30%).
2. Calculated Euclidian distance for x-train dataset.
3. Added column “distance” in dataset.
4. Sorted dataset by distance column.
5. Selected top k tuples for query.

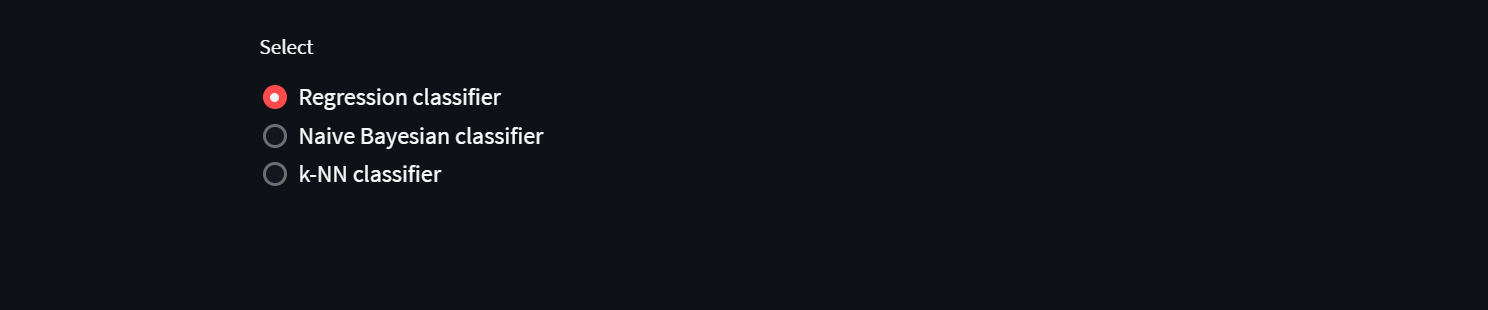
**Results/Observation(ScreenShots):**

**snapshots**



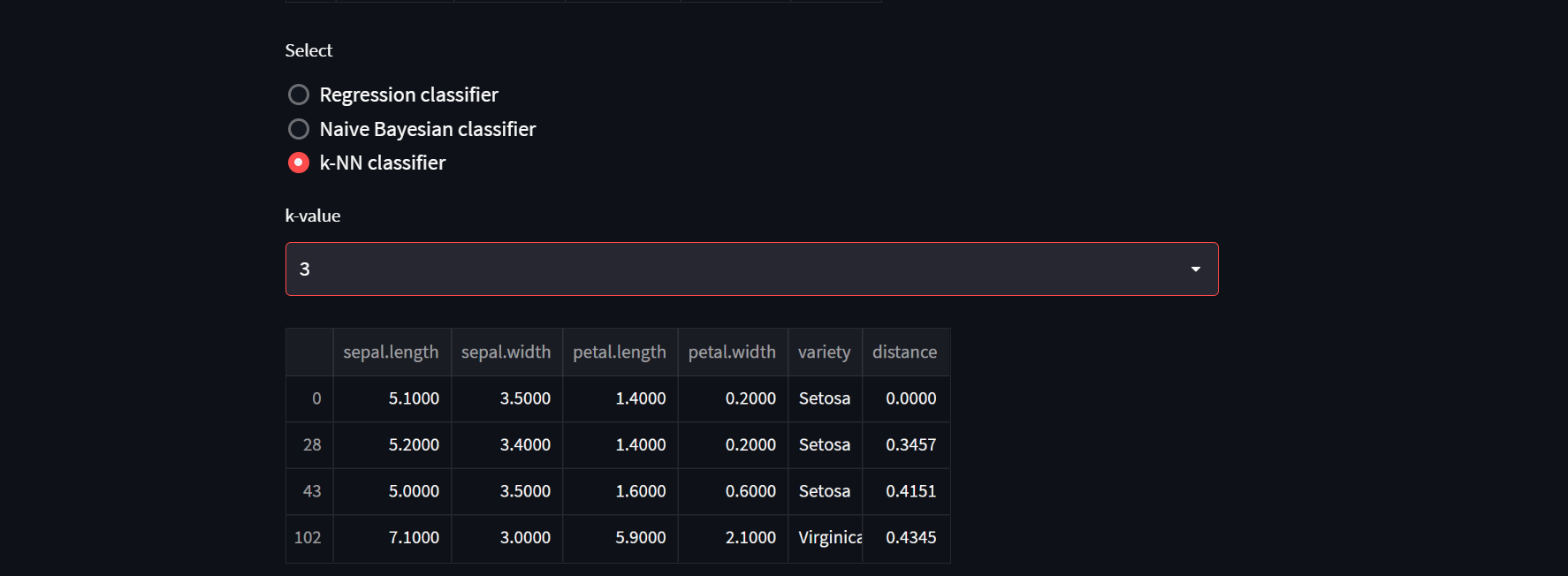
Loaded Iris Dataset:



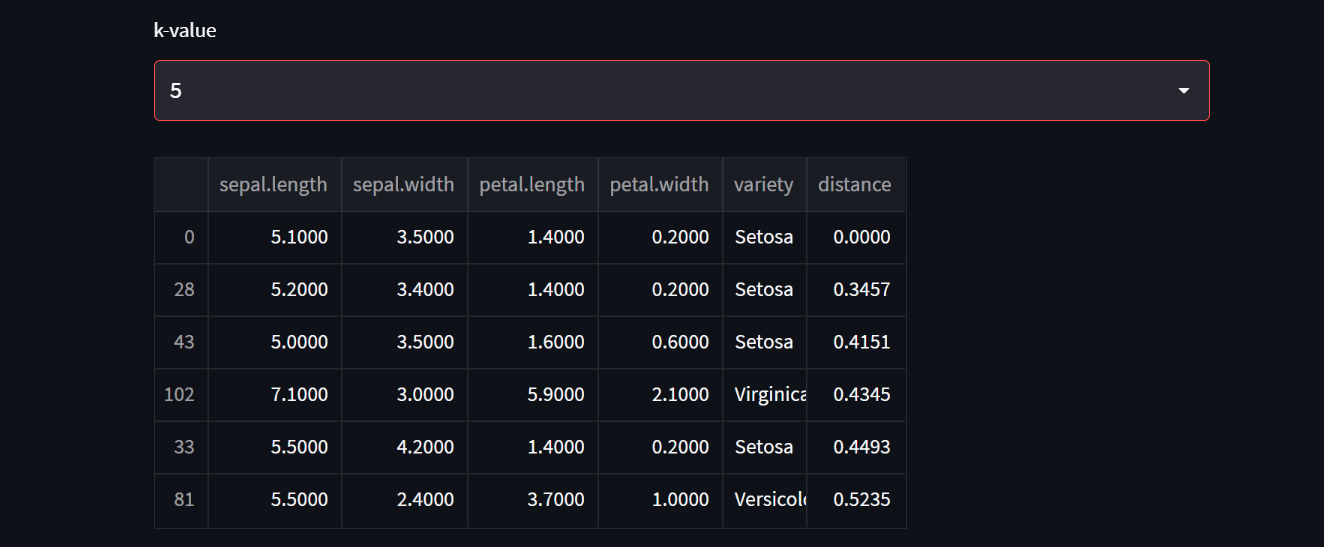


K=1

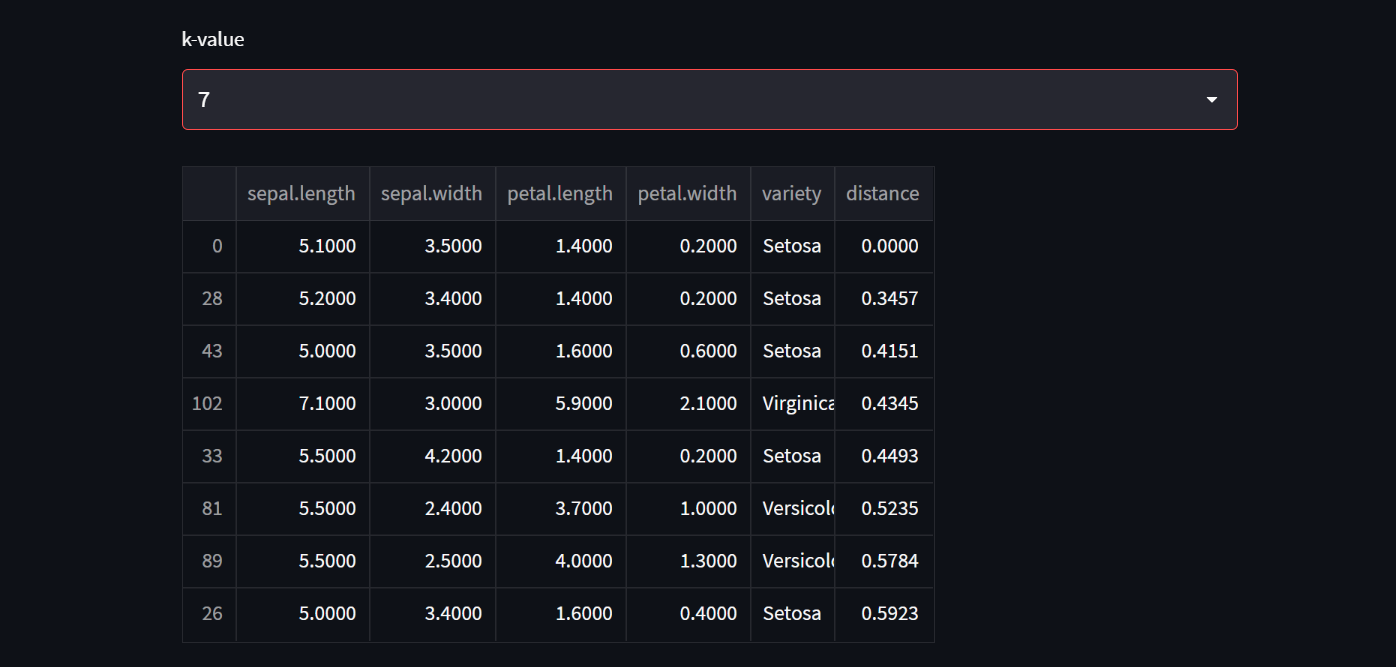


K = 3

K=5



K = 7



**Conclusion:**

Using the stream lit framework for UI and using python language dataset is loaded and displayed in grid form. Calculated Euclidian distance for all x-train dataset and selected top k tuples based on k value. Build knn classifier algorithm to pick top k nearest neighbour tuples.

Stream lit is an open-source app framework for Machine Learning and Data Science teams.