**Final Year B. Tech. (CSE) – I: 2022-23**

**5CS462: PE5 - Data Mining Lab**

**Assignment No. 5**

**PRN: 2019BTECS00077 Date:19 Sep 2022**

**Full name: Biradar Avinash Vishnu**

**Title: Design, implementation and performance evaluation of Various Classifiers**

**Objective: Design and implementation of following classifiers**

* **Regression classifier.**
* **Naïve Bayesian Classifier.**
* **k-NN classifier**
* **Three-layer Artificial Neural Network (ANN) classifier**

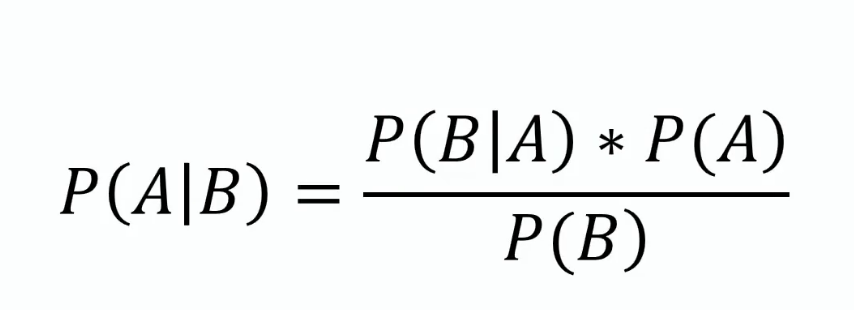
**Dataset Use: Car evaluation, Brest Cancer, Balance Scale**

**Introduction & Theory:**

1. **Regression classifier:**

 It is a supervised learning classification algorithm which is used to predict observations to a discrete set of classes. Practically, it is used to classify observations into different categories. Hence, its output is discrete in nature. **Logistic Regression** is also called **Logit Regression**. It is one of the most simple, straightforward and versatile classification algorithms which is used to solve classification problems.

**2.Naive Bayes** is a classification algorithm that works based on the Bayes theorem. Before explaining about Naive Bayes, first, we should discuss Bayes Theorem. Bayes theorem is used to find the probability of a hypothesis with given evidence.



In this, using Bayes theorem we can find the probability of A, given that B occurred. A is the hypothesis and B is the evidence.

P(B|A) is the probability of B given that A is True.

P(A) and P(B) is the independent probabilities of A and B.

**3.KNN**

K-nearest neighbors (KNN) algorithm is a type of supervised ML algorithm which can be used for both classification as well as regression predictive problems. However, it is mainly used for classification predictive problems in industry. The following two properties would define KNN well −

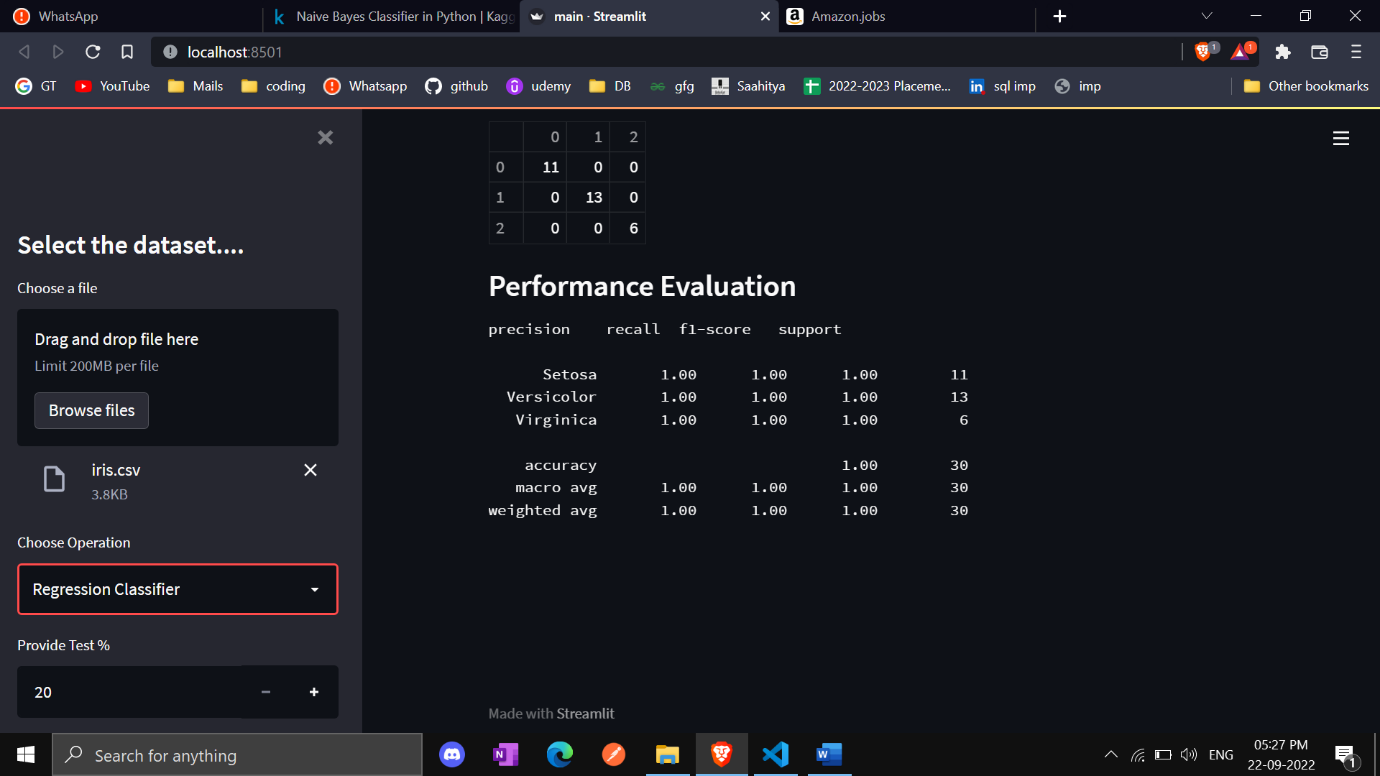
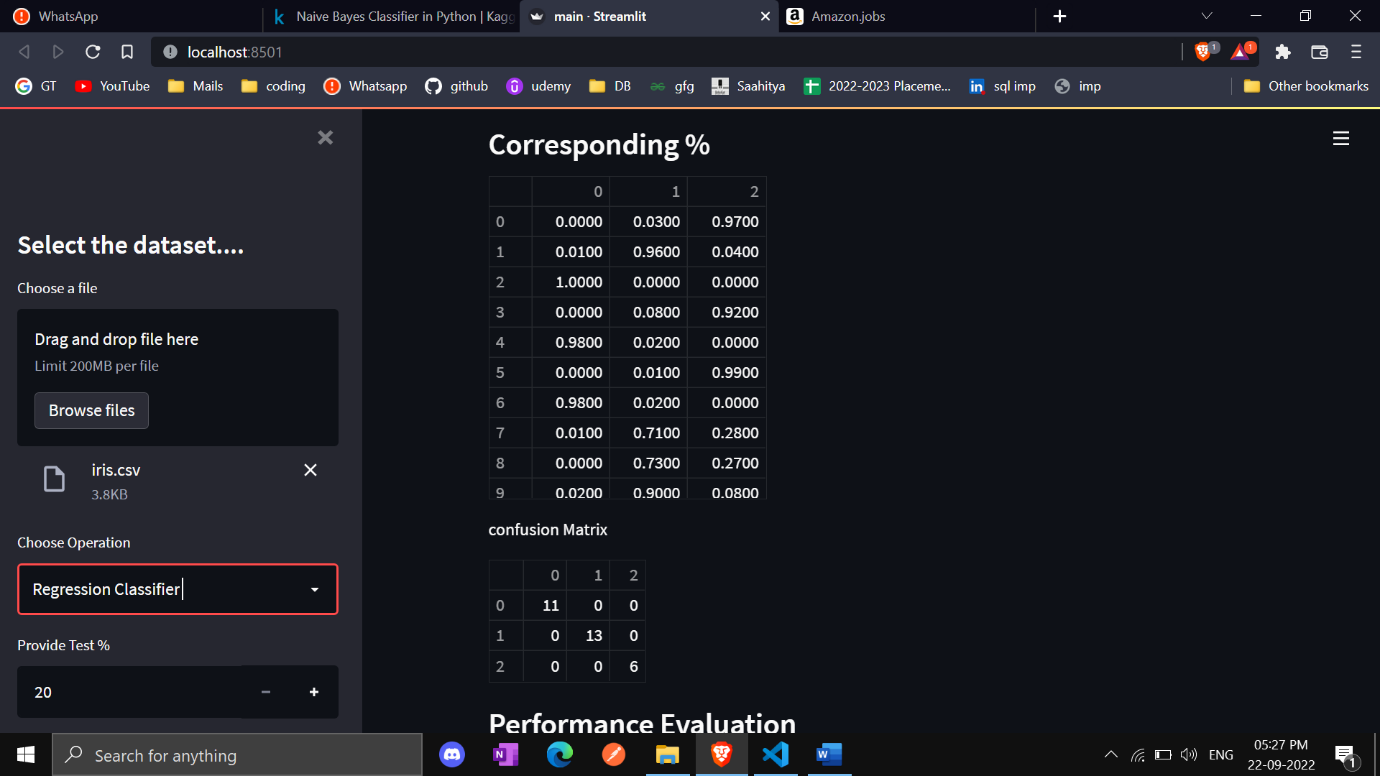
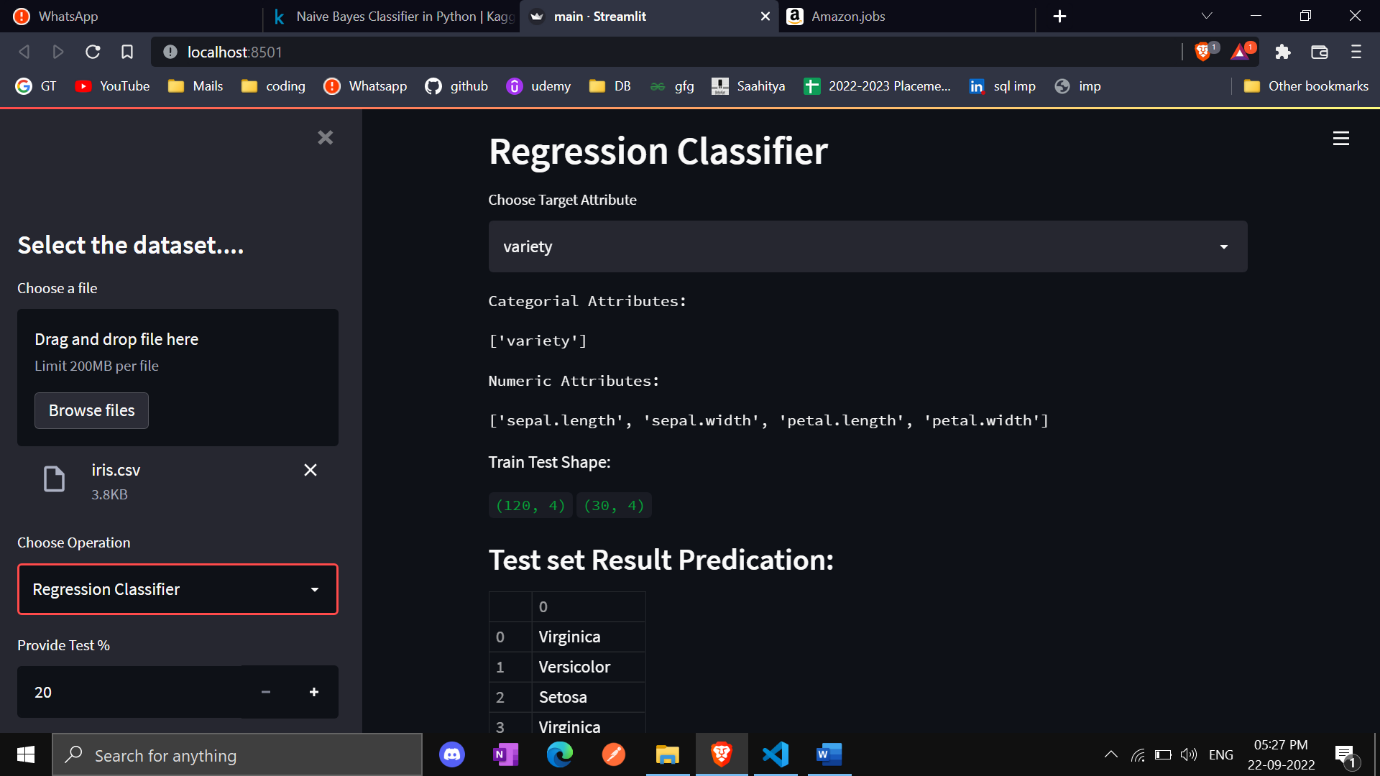
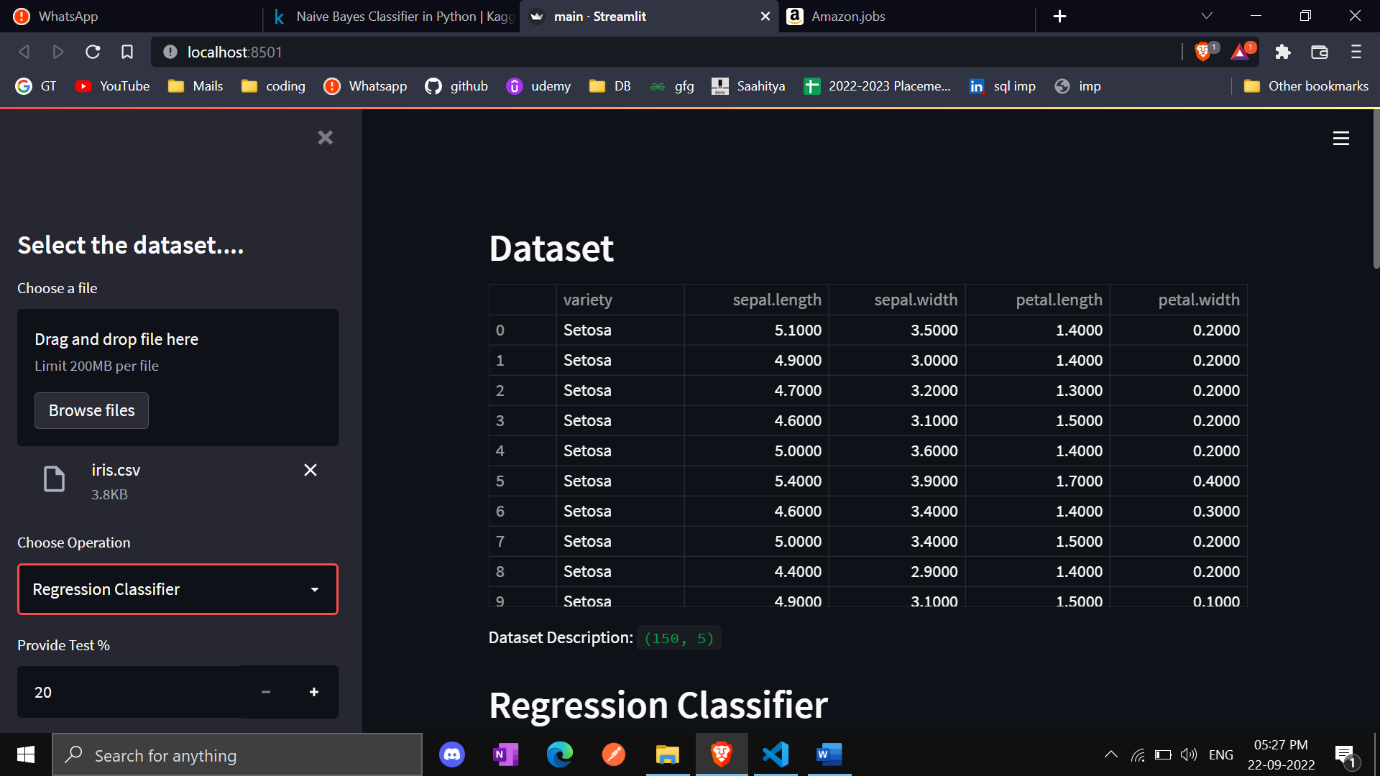
* **Lazy learning algorithm** − KNN is a lazy learning algorithm because it does not have a specialized training phase and uses all the data for training while classification.
* **Non-parametric learning algorithm** − KNN is also a non-parametric learning algorithm because it doesn’t assume anything about the underlying data.

**4 ANN**

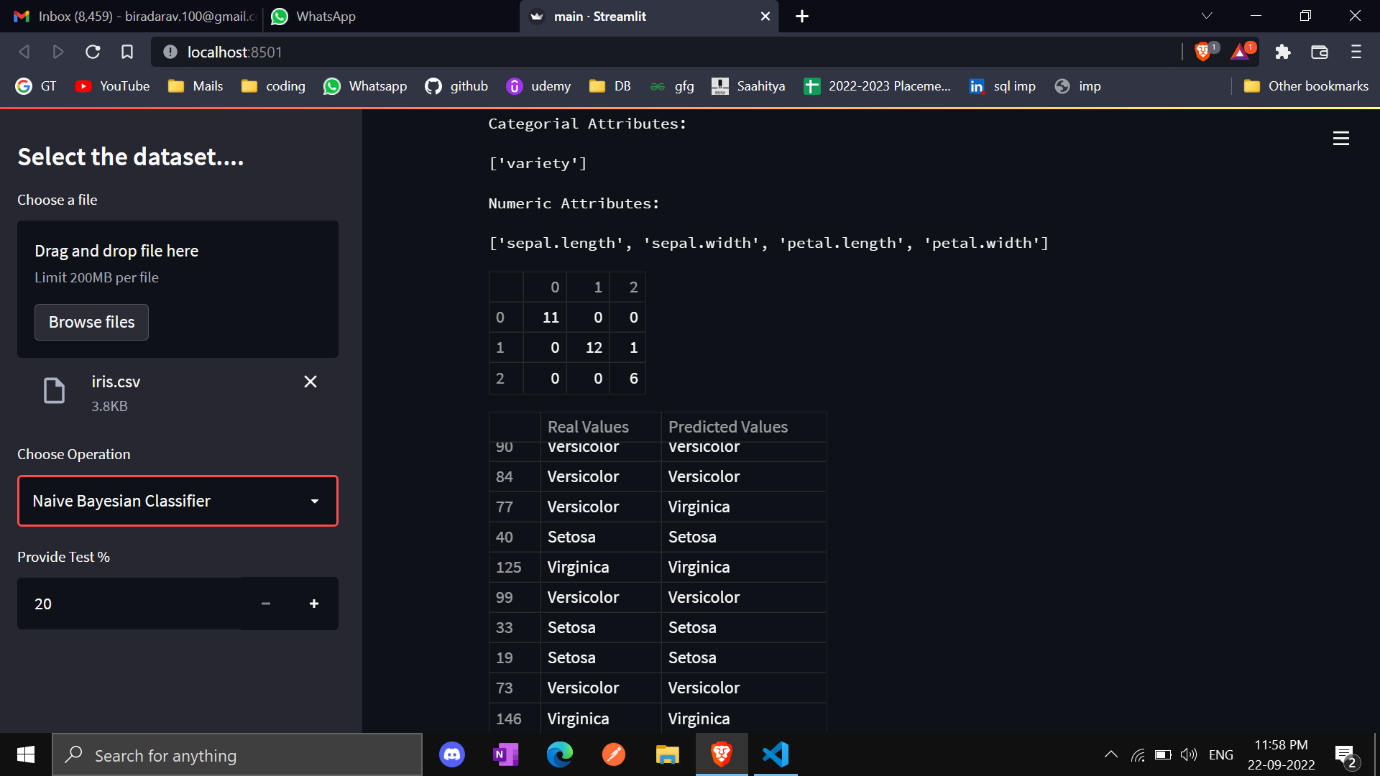
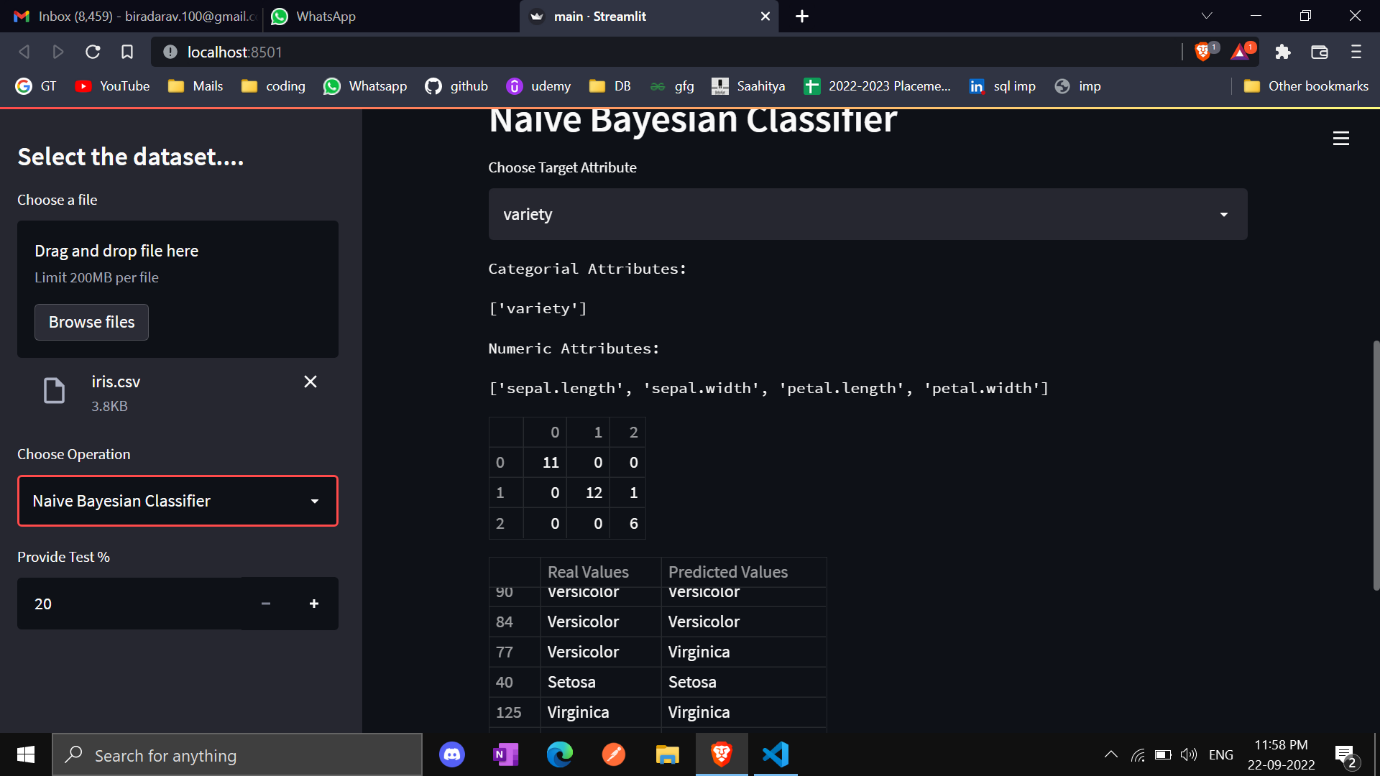
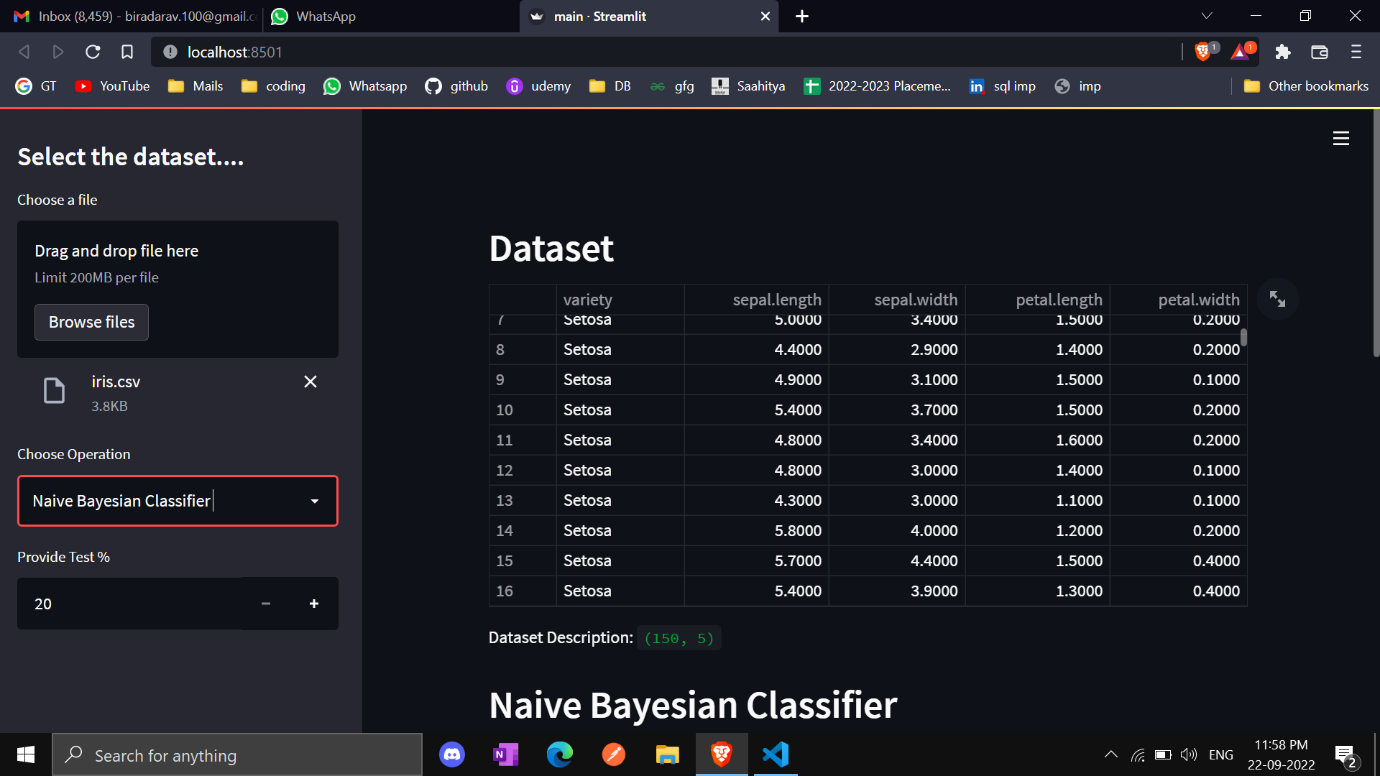
Classification ANNs seek to classify an observation as belonging to some discrete class as a function of the inputs. The input features (independent variables) can be categorical or numeric types, however, we require a categorical feature as the dependent variable.

**Implementation / Code:**

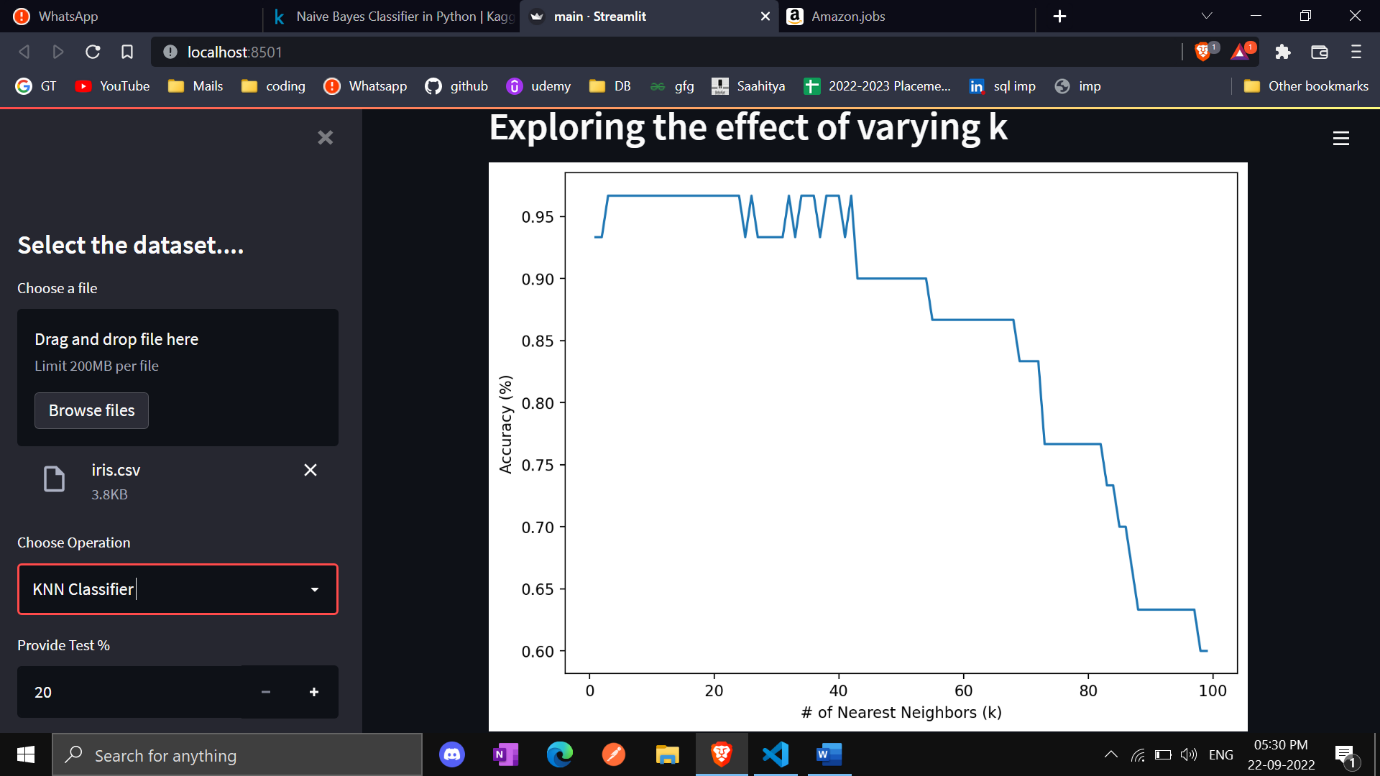
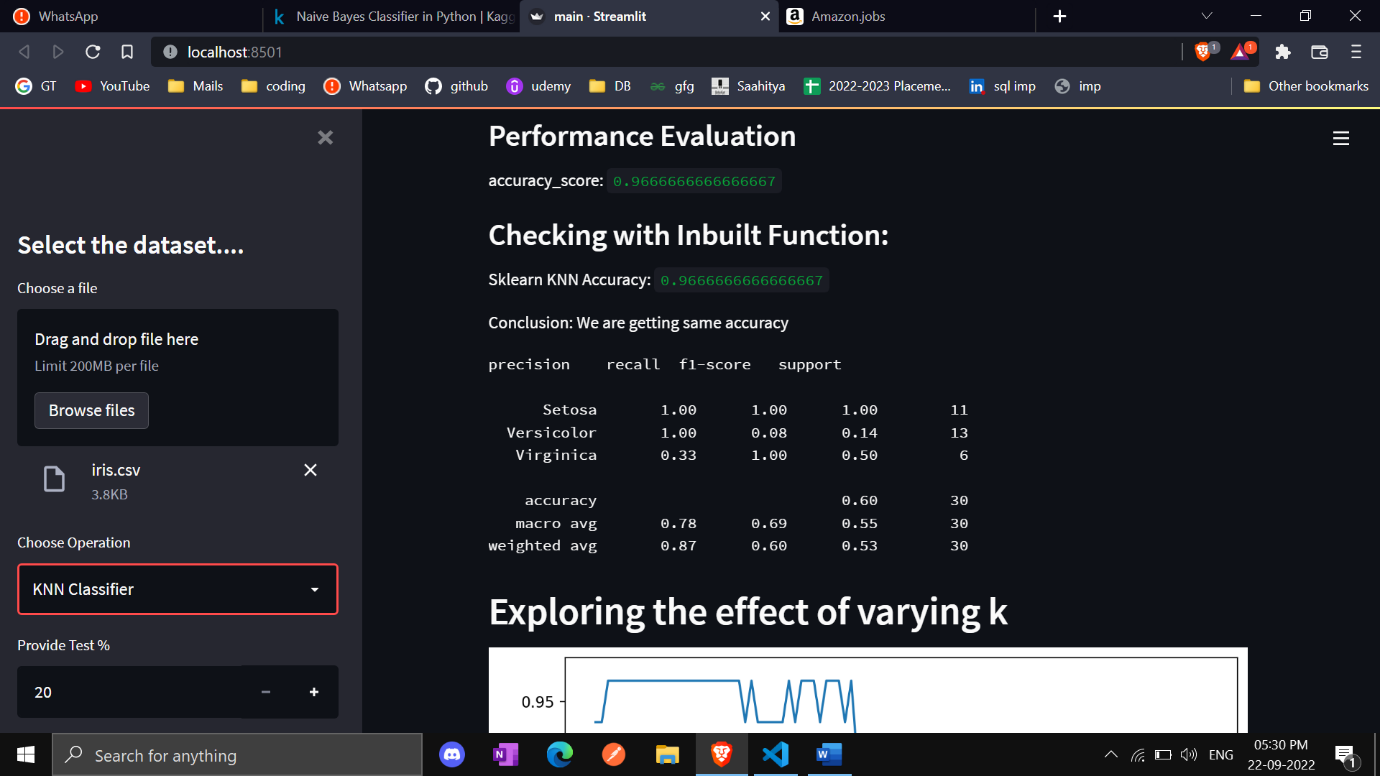
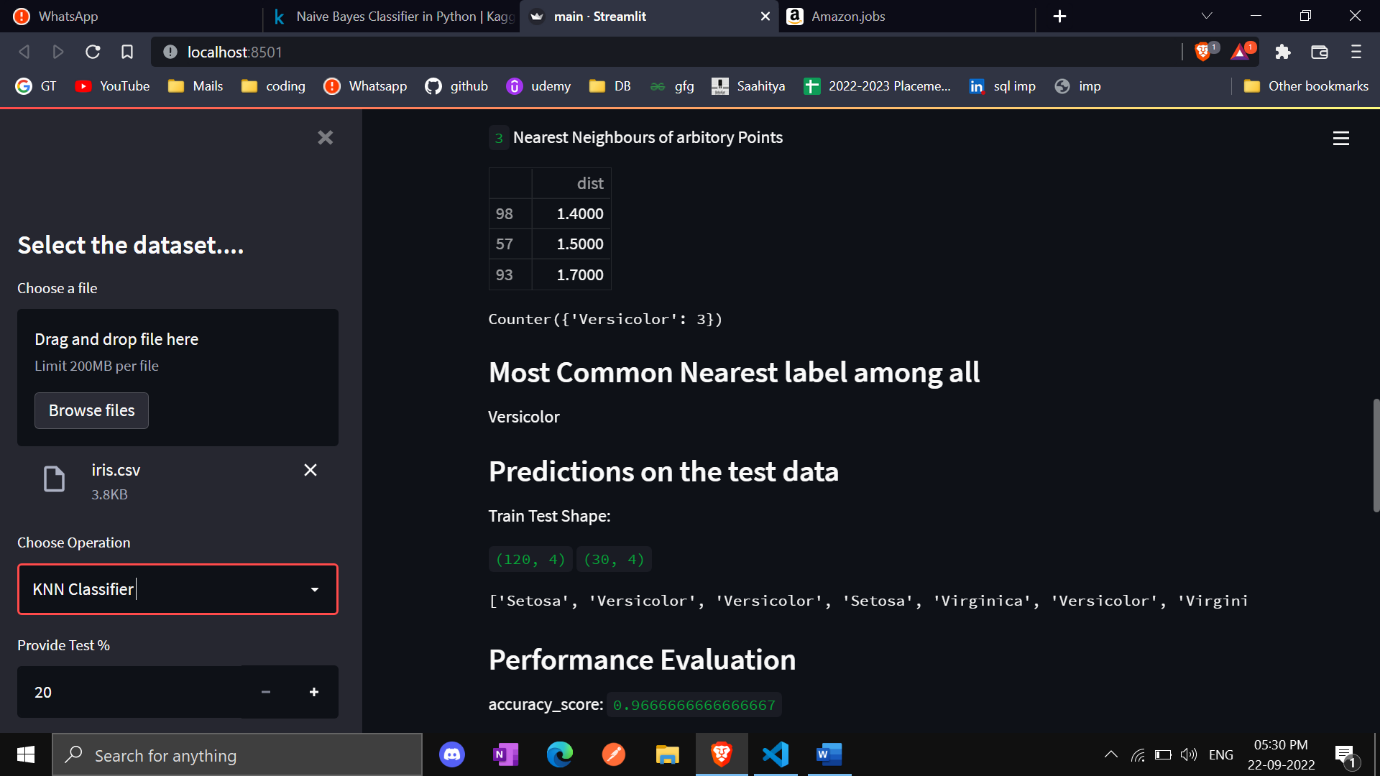
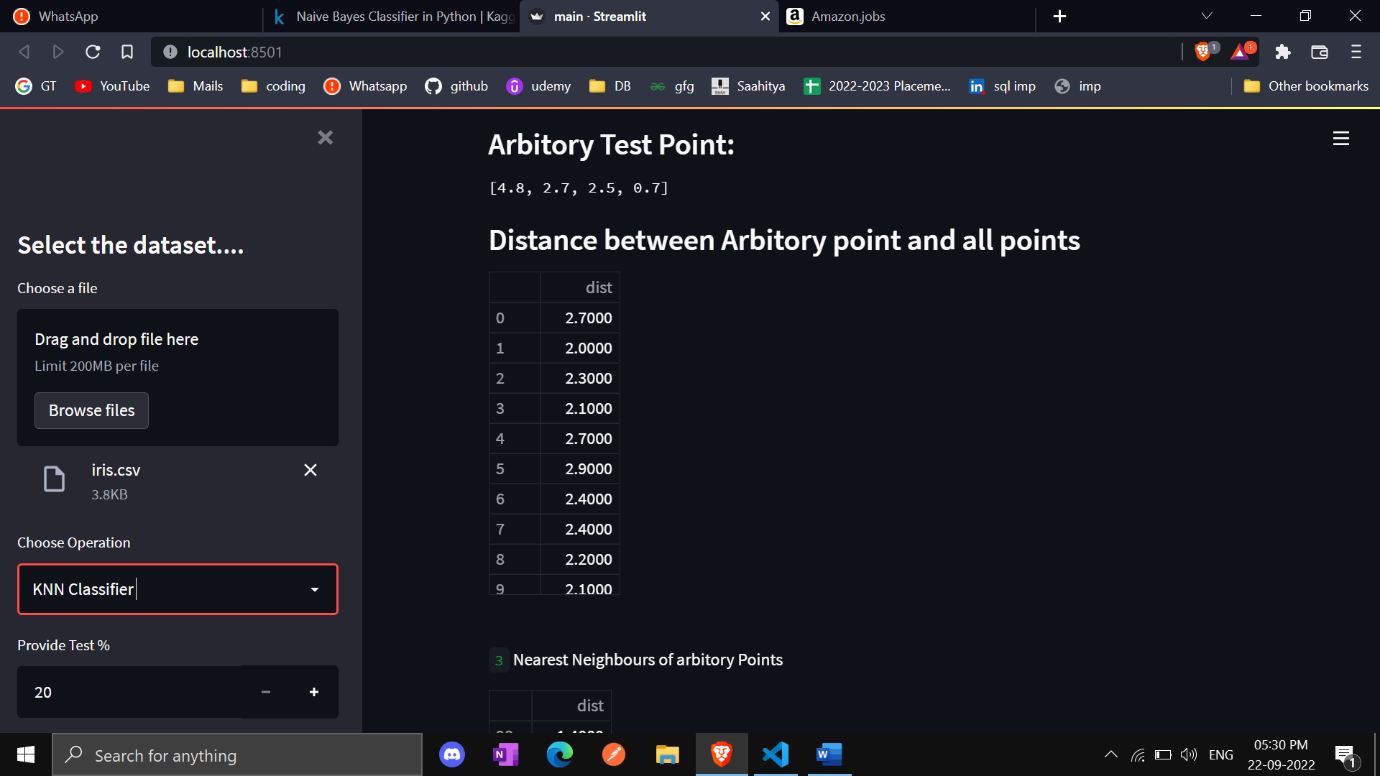
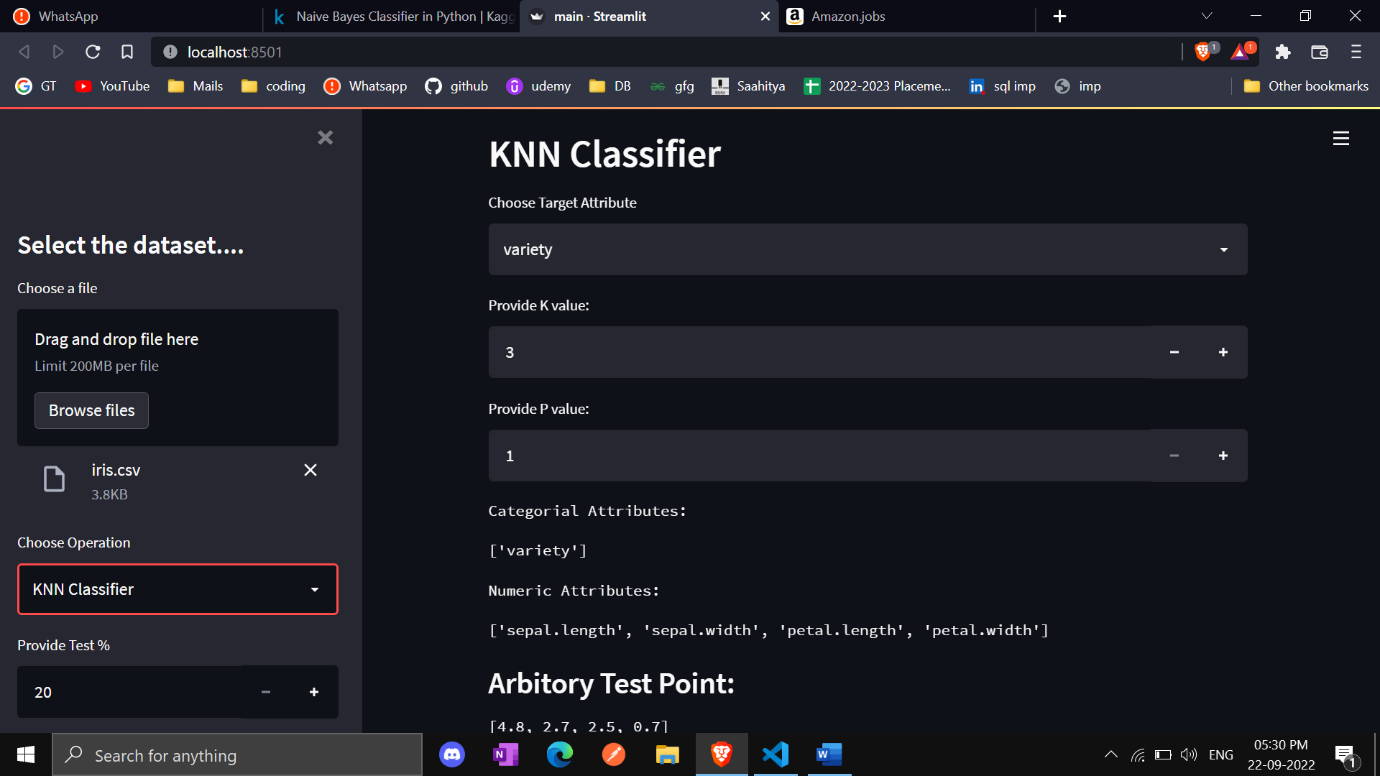
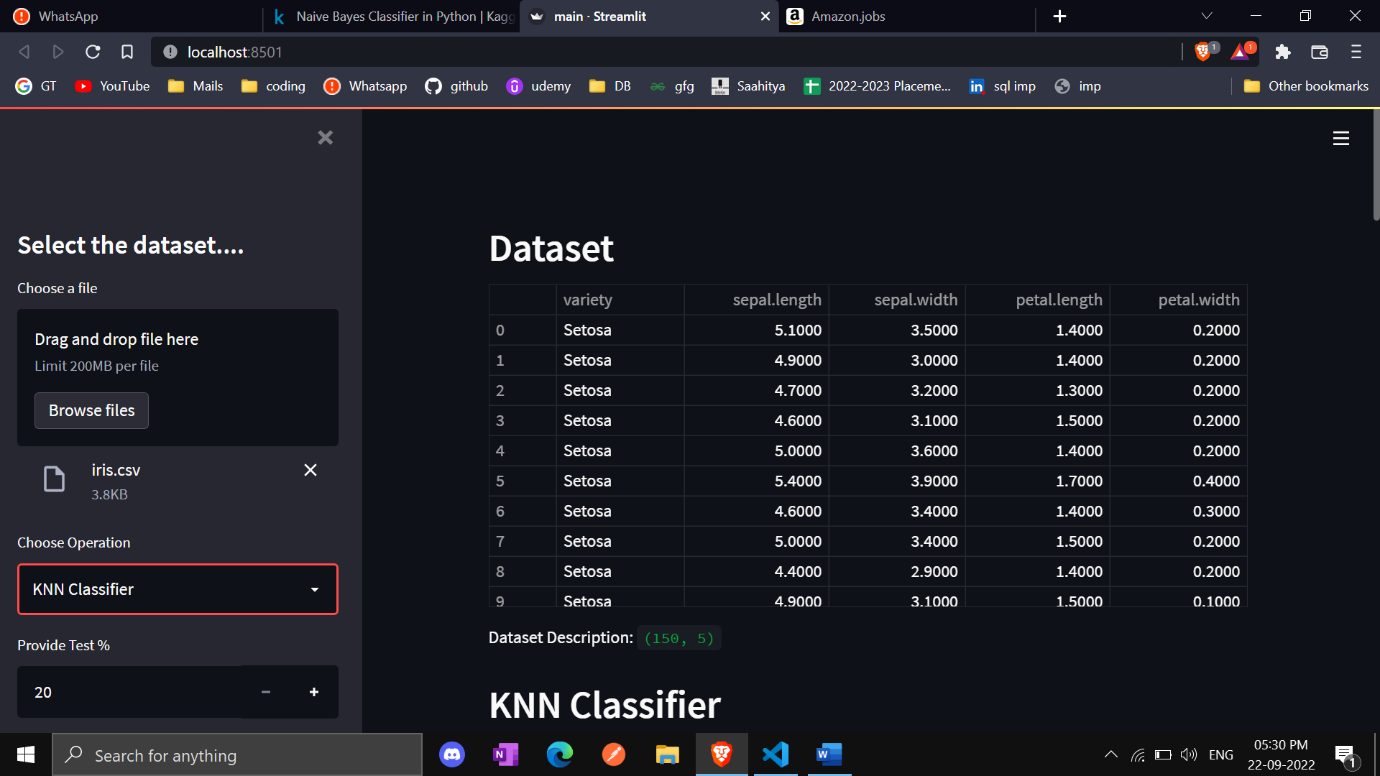
**1.Regression Classifier**

****

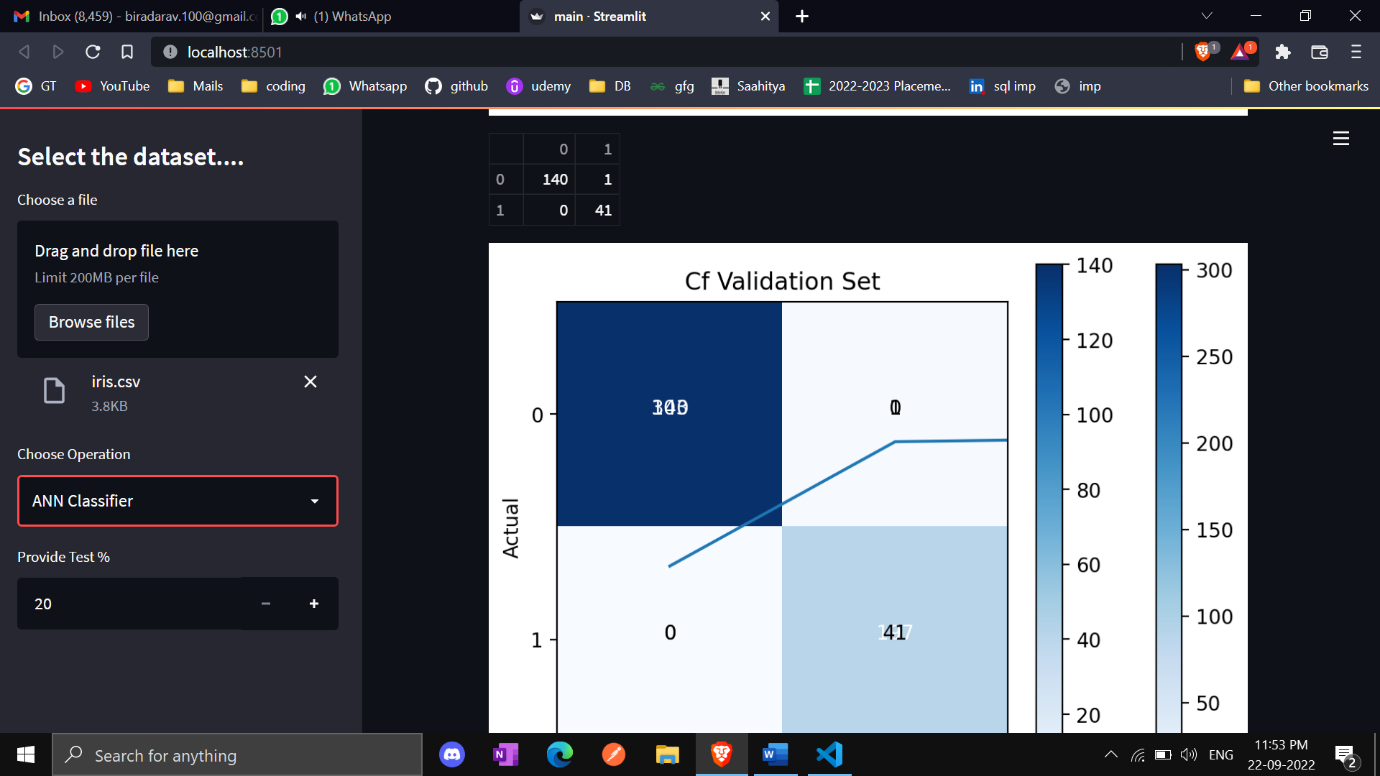
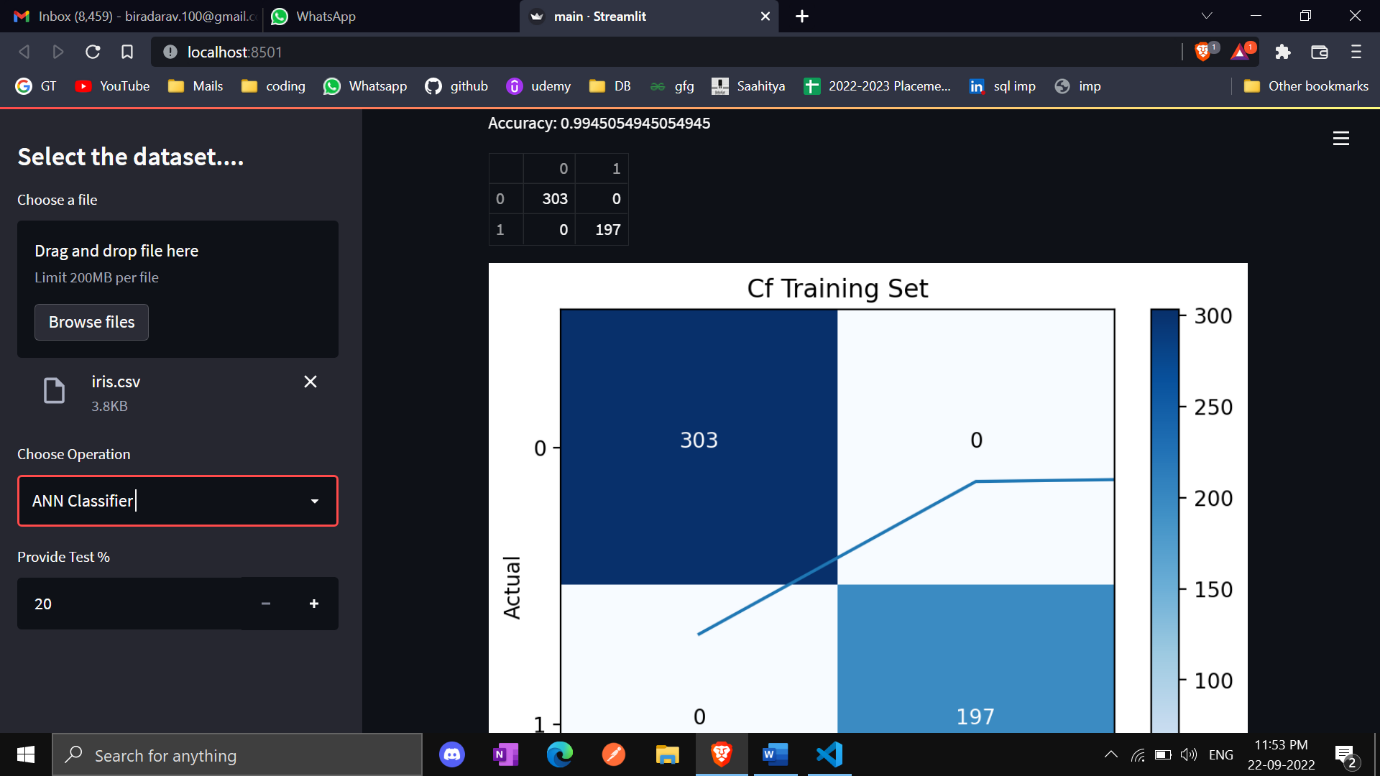
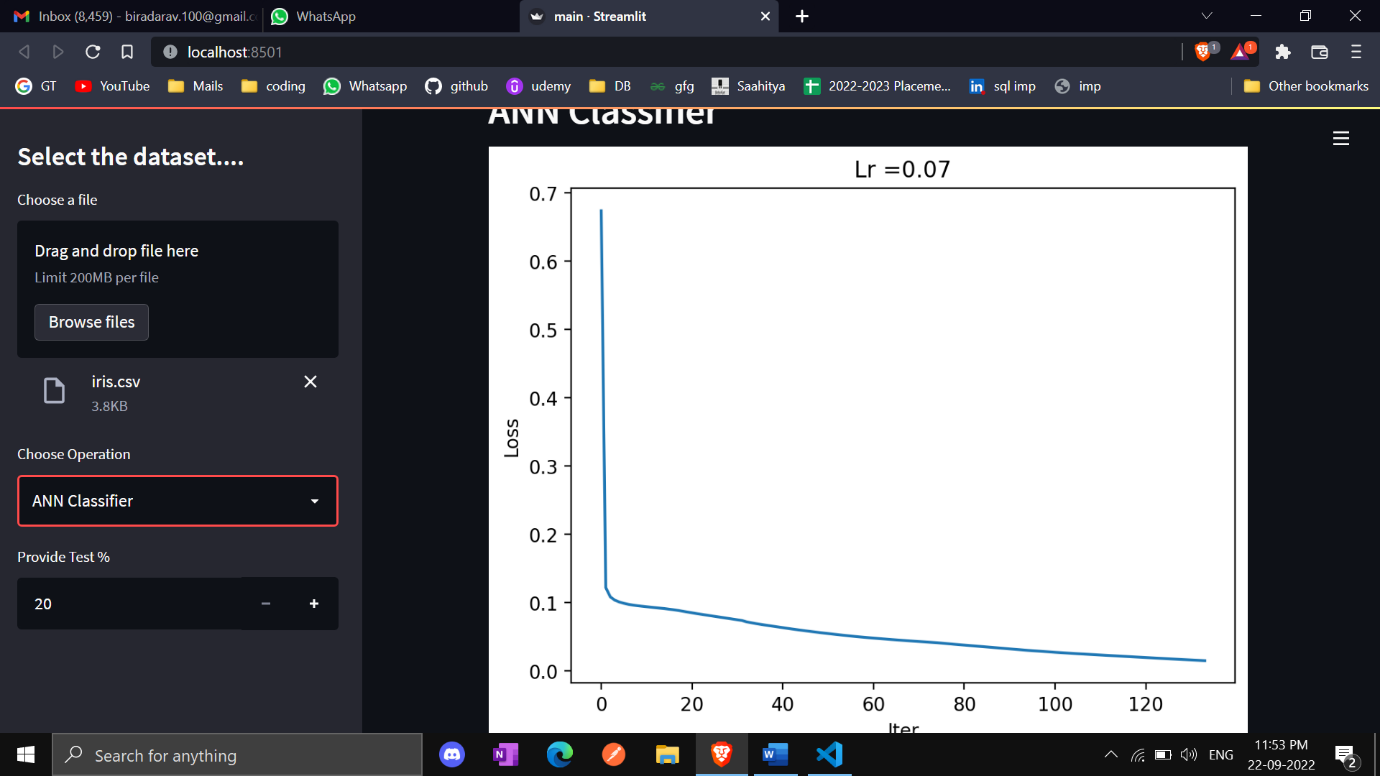
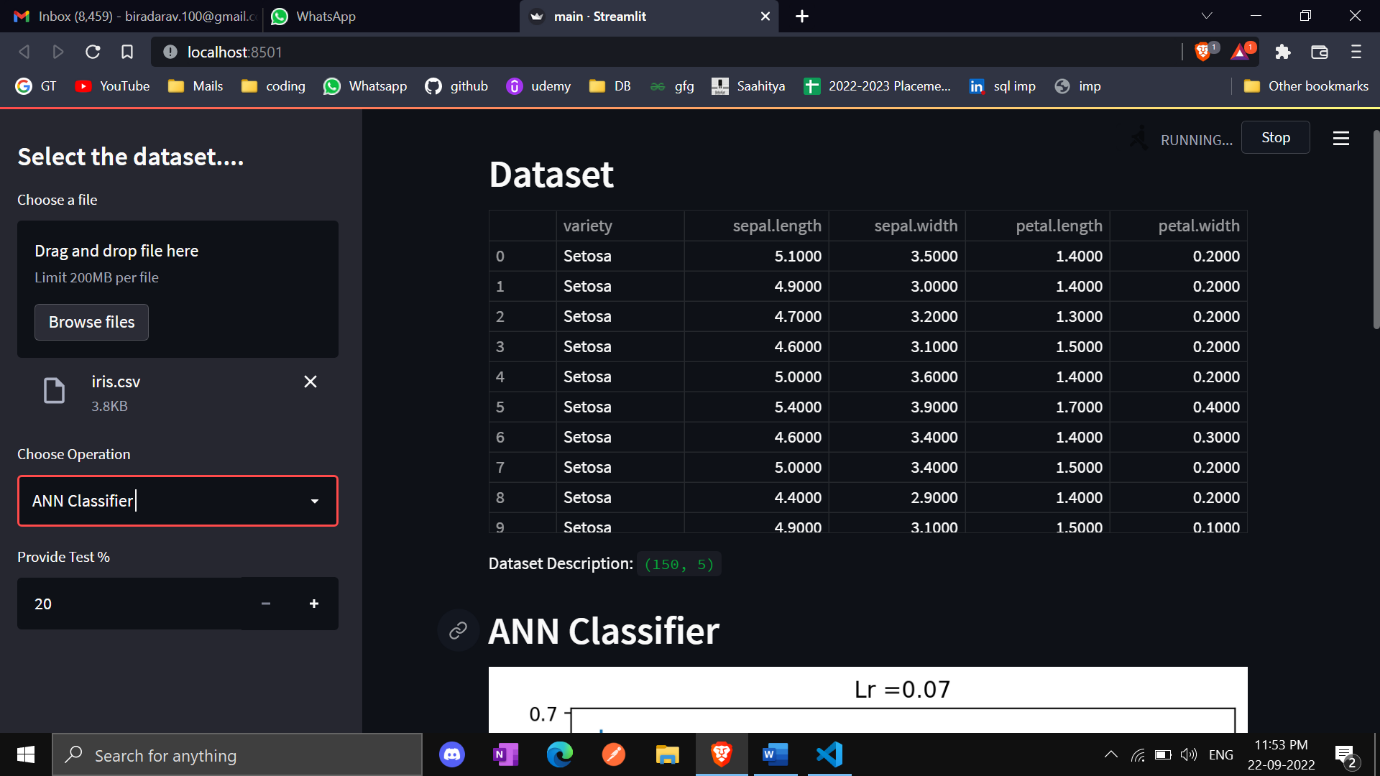
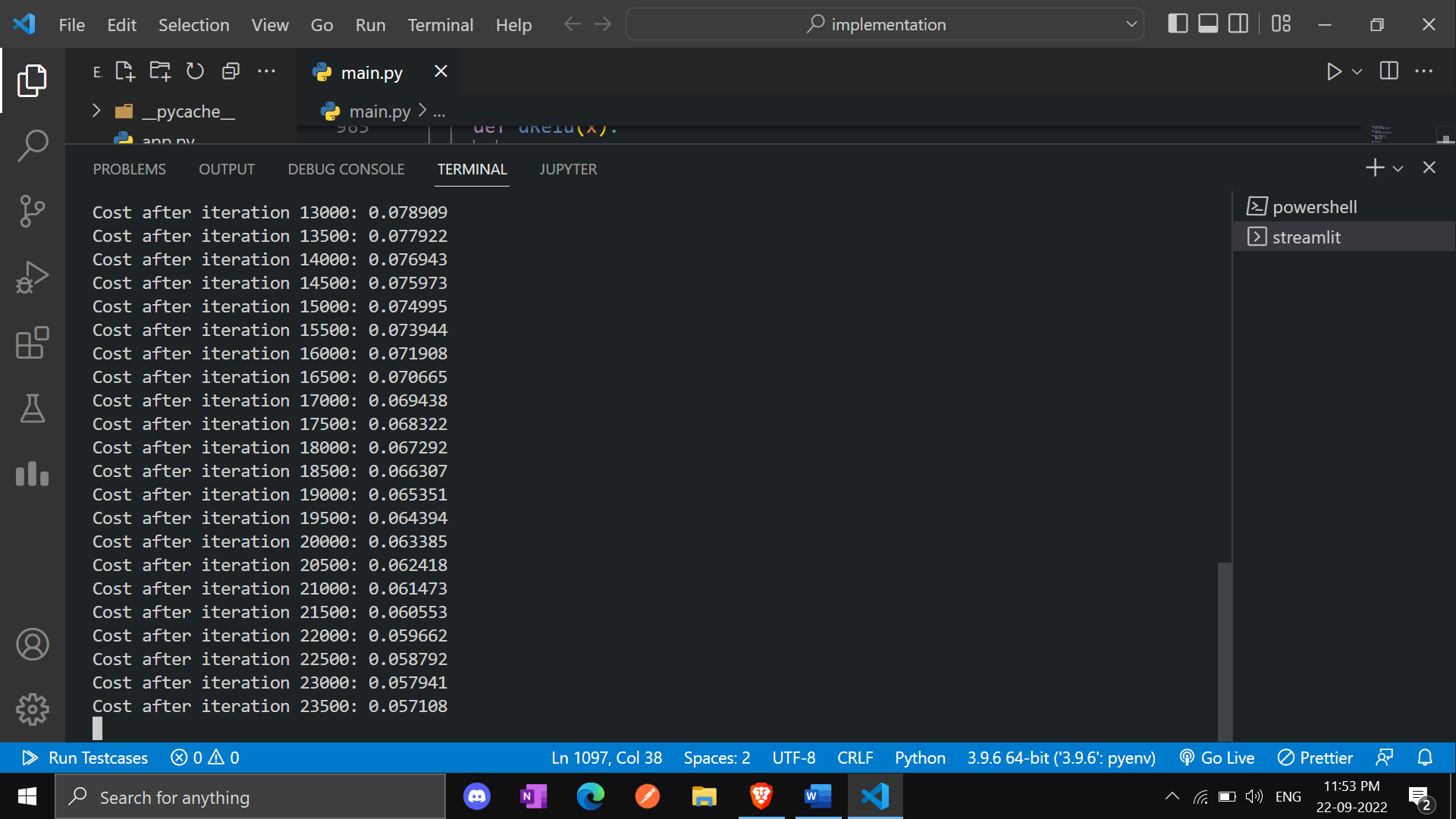
**2.Naive Bayes Classifier**

****

**3.KNN Classifier**

****

**4.ANN Classifier**

****