# Introduction

This file is a comparative analysis on Cycon’s ability to perform Gaussian Naiive Bayes. This serves as proof that the Cycon page is able to perform Gaussian Naiive Bayes. The following shows Gaussian Naiive Bayes results for various datasets.

## Iris.csv

|  |  |
| --- | --- |
| **Dataset:** | |
| Shape: 150 x 5  Samples: 50 samples for 3 classes  Classes: Iris-setosa, Iris-versicolor, Iris-virginica  Purpose: Identify class of iris flowers given petal information. | |
| **Comparative Work:**  [**https://www.kaggle.com/code/vinayshaw/iris-species-100-accuracy-using-naive-bayes**](https://www.kaggle.com/code/vinayshaw/iris-species-100-accuracy-using-naive-bayes) | **Cycon Work:** |
| **Settings:** | |
|  |  |
| **Results:** | |
|  |  |
| **Any Additional Information:** | |
|  | |

## Gender\_Classification\_v7.csv

|  |  |
| --- | --- |
| **Dataset:** | |
| Shape: 5001x8  Samples: 2500 Male samples and 2501 Female samples  Classes: Male and Female  Purpose: Identification of the gender based on the characteristics of a person. | |
| **Comparative Work:**  [**https://www.kaggle.com/code/ihsncnkz/classification-with-naive-bayes-classification/notebook**](https://www.kaggle.com/code/ihsncnkz/classification-with-naive-bayes-classification/notebook) | **Cycon Work:** |
| **Settings:** | |
|  |  |
| **Results:** | |
|  |  |
| **Any Additional Information:** | |
| Note that while the comparative work does not specifically stated that they shuffled during the validation split, the results are exactly the same when shuffling when using a random\_state of 1. | |

## Diabetes.csv

|  |  |
| --- | --- |
| **Dataset:** | |
| A picture containing text, screenshot, number, font  Description automatically generated  Shape: 768 x 9  Samples: 384 for yes and 384 for no  Classes: 0 (no) and 1 (yes)  Purpose: Determine if the person has diabetes. 0 means the person does not have diabetes while 1 means they have diabetes. | |
| **Comparative Work:**  [**https://www.kaggle.com/code/berkayalan/classification-with-naive-bayes-beginner**](https://www.kaggle.com/code/berkayalan/classification-with-naive-bayes-beginner) | **Cycon Work:** |
| **Settings:** | |
|  |  |
| **Results:** | |
|  |  |
| **Any Additional Information:** | |
| Note that the split was done manually in the comparative work. However, we can simply calculate the percentage to split the training and test set by simple math (I.E. 168/768) to get the same training and testing sets. Also note that the comparative work calculates the precision and f1 for the class of (1) only, not the micro or macro. | |