Methodology:

1. Dataset is loaded and split into training and testing.
2. The samples are split by class in training data.
3. Mean and standard deviation are computed for each column of the dataset.
4. For testing data, class posteriors are computed for a single row assuming Gaussian distribution.
5. Class with highest posterior is selected as the prediction.
6. Accuracy is computed

Results:

Multi-class Gaussian Naïve Bayes classifier has been realized.

An average accuracy of 91.5% has been noted on running the code 5 times on different test and training data each time.

Accuracy averaged over 5 iterations:

```
Accuracy: Iteration 0 92.85714285714286
Accuracy: Iteration 1 95.23809523809524
Accuracy: Iteration 2 95.23809523809524
Accuracy: Iteration 3 90.47619047619048
Accuracy: Iteration 4 83.33333333333333
Averaged Accuracy 91.42857142857143
```

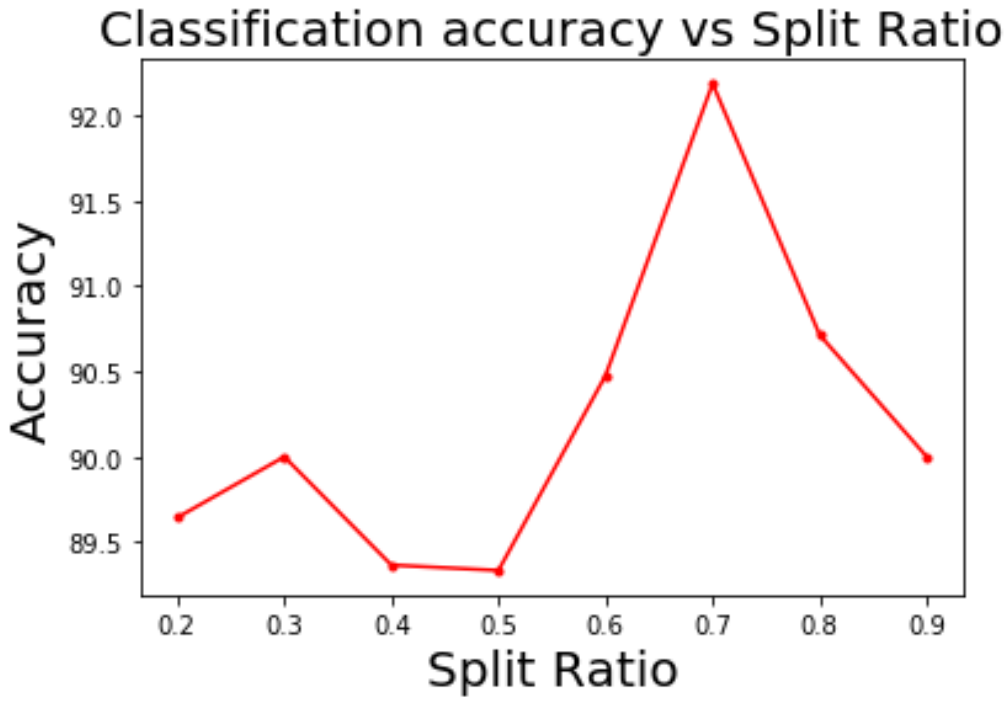

Fig: Split ratio of dataset into x% train and 1-x% test, vs Accuracy

Sample Outputs:

```
42 / 210 in testing
Class: feature-wise [Mean, Standard Deviation, Length]
 1 [(14.435535714285711, 1.230142518931492, 56), (14.345714285714283, 0.5816680000977038, 56),
(0.879942857142857, 0.016531829983320135, 56), (5.530910714285713, 0.23342550119982522, 56),
(3.2543571428571427, 0.17966457925523616, 56), (2.741967857142857, 1.2727911716320222, 56),
(5.107517857142858, 0.27812743914004134, 56)]
Class: feature-wise [Mean, Standard Deviation, Length]
 2 [(18.4338596491228, 1.4343176007948428, 57), (16.172105263157892, 0.6070029914085657, 57),
(0.8843105263157891, 0.01567512456330692, 57), (6.159035087719298, 0.26283814716286913, 57),
(3.6921929824561395, 0.19278980398688955, 57), (3.654298245614035, 1.2465353837868307, 57),
(6.031438596491229, 0.24832460161937958, 57)]
Class: feature-wise [Mean, Standard Deviation, Length]
 3 [(11.839038461538461, 0.7074235484651871, 52), (13.226153846153847, 0.3231168904388592,
52), (0.8497442307692312, 0.023339072796589953, 52), (5.218096153846154, 0.13046376698041862,
52), (2.850673076923077, 0.1489967391607715, 52), (4.820750000000001, 1.3004374509085357, 52),
(5.097846153846153, 0.14859240006630908, 52)]

Final Accuracy: 90.47619047619048
```

```
42 / 210 in testing
Class: feature-wise [Mean, Standard Deviation, Length]
 1 [(14.269482758620688, 1.2586046609956503, 58), (14.271379310344829, 0.5981358417437157,
58), (0.8787499999999999, 0.01594392861402492, 58), (5.503448275862068, 0.23503924065163675,
58), (3.2310344827586204, 0.1807961889752474, 58), (2.772124137931034, 1.2300270666413387,
58), (5.0951551724137945, 0.2690847907665901, 58)]
Class: feature-wise [Mean, Standard Deviation, Length]
 2 [(18.177692307692304, 1.487510903416444, 52), (16.067307692307686, 0.6478262759885837, 52),
(0.883376923076923, 0.01603586195931733, 52), (6.121346153846152, 0.2849765648854949, 52),
(3.6631346153846143, 0.18708150984878266, 52), (3.678499999999999, 1.222911128190017, 52),
(5.997365384615387, 0.2648166238521982, 52)]
Class: feature-wise [Mean, Standard Deviation, Length]
 3 [(11.823454545454547, 0.7062347164700093, 55), (13.216181818181822, 0.3245823008632324,
55), (0.8499509090909093, 0.023311925677384073, 55), (5.2129818181818175, 0.13519553615646118,
55), (2.852618181818182, 0.15138918339328225, 55), (4.790818181818182, 1.4327541735711404,
55), (5.097799999999998, 0.1525275202855027, 55)]

Final Accuracy: 95.23809523809524
```

Assumptions:

1. Training data is taken as 80% of the dataset, rest is testing
2. Gaussian Probability is taken
3. Dataset is present in the directory of the code with the name ‘download.txt’