REPORT 2

ALGORITHMS USED: -

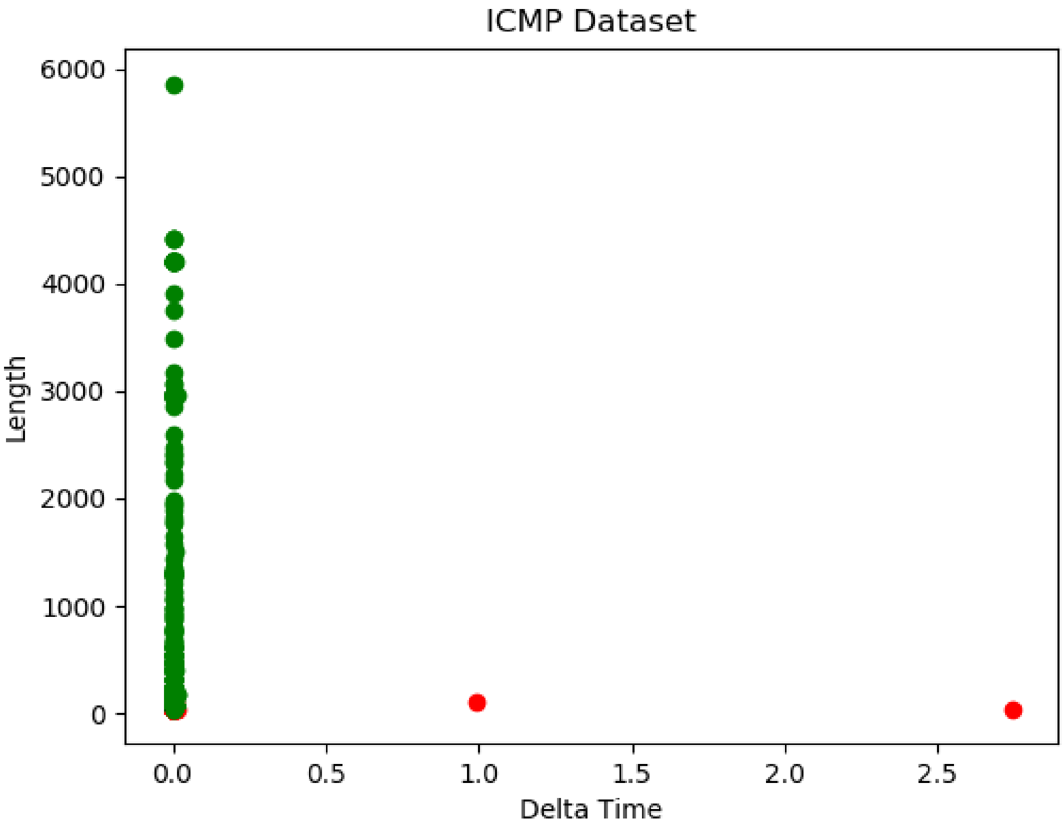
1. **OneVsRestClassifier:** OneVsRestClassifier is used for multiclass classification (usually when number of classes is greater than 2). Based on the number of classes present, an equal number of classifiers should be passed. Each classifier is used for classification of a single class. This way, each classifier can be used for the class it is best suited for. It is suitable for multiclass classification where we have strong classifiers for each class present.

2. **BaggingClassifier:** Bagging Classifier is an ensemble meta-classifier. It takes a number of base classifiers and fits them on random subsets of the original dataset and then aggregates the individual predictions using voting or weighted average. This can reduce variance introduced due to poor classification by classifiers. It can be used when the classifiers used clash in values causing inaccurate measurements. Using this will train the classifiers using different subsets and hence will not generate complementary classifications for the same data.

3. **MultinomialNB:** MultinomialNB is a variant of the Naïve Bayes algorithm. It makes the same naïve assumption that there is conditional independence between every pair of features. It implements the naïve Bayes algorithm for multinomially distributed data. It is used in text classification as the data here is represented by word vector counts. Each class is given a parameterized vector having components equal to the number of features. A component for a given feature i in a vector for a given class y is the probability of that feature appearing in a sample belonging to that class. MultinomialNB usually requires integer feature counts but in practice, fractional counts such as tf-idf may also work.

VISUALIZATION GRAPHS OF THE DATASETS: (red – flood, green – normal)

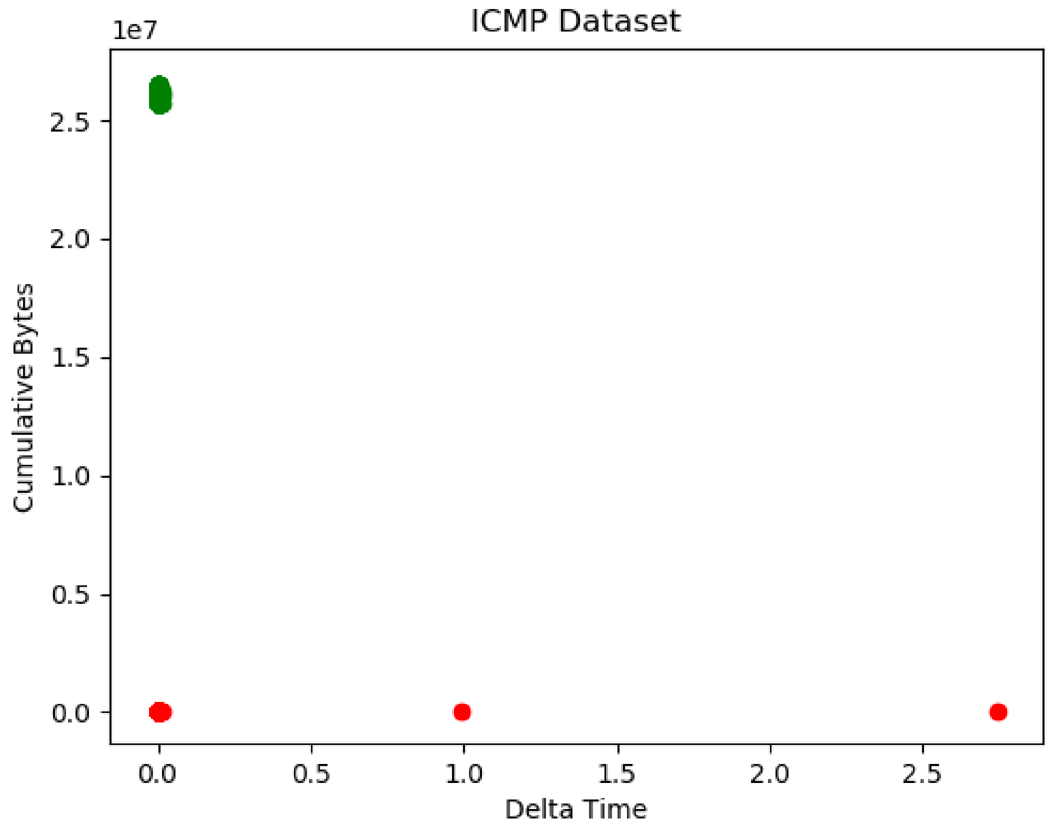
ICMP



We can see from this graph that when the length of the sent packets is 0, then irrespective of delta time values, it is a flood condition

Similarly, when delta time is 0, then irrespective of length values, it is a normal condition.

This shows that packets that take lesser time and are larger in size are of normal condition and packets that take more time and are smaller in size are of flood condition.

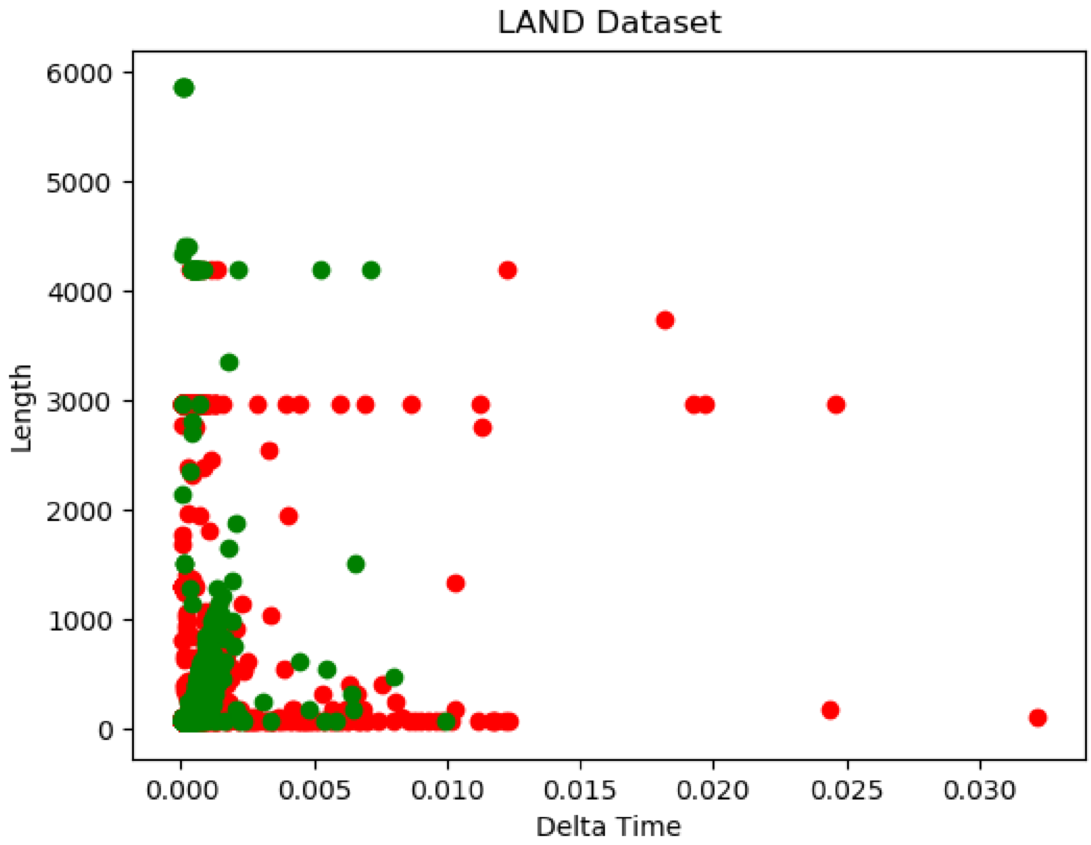


From this graph, we can see that when cumulative bytes is 0, for all Delta times values, it is a flood condition.

Similarly, when Delta time is 0, for all values of Cumulative bytes, it is a normal condition.

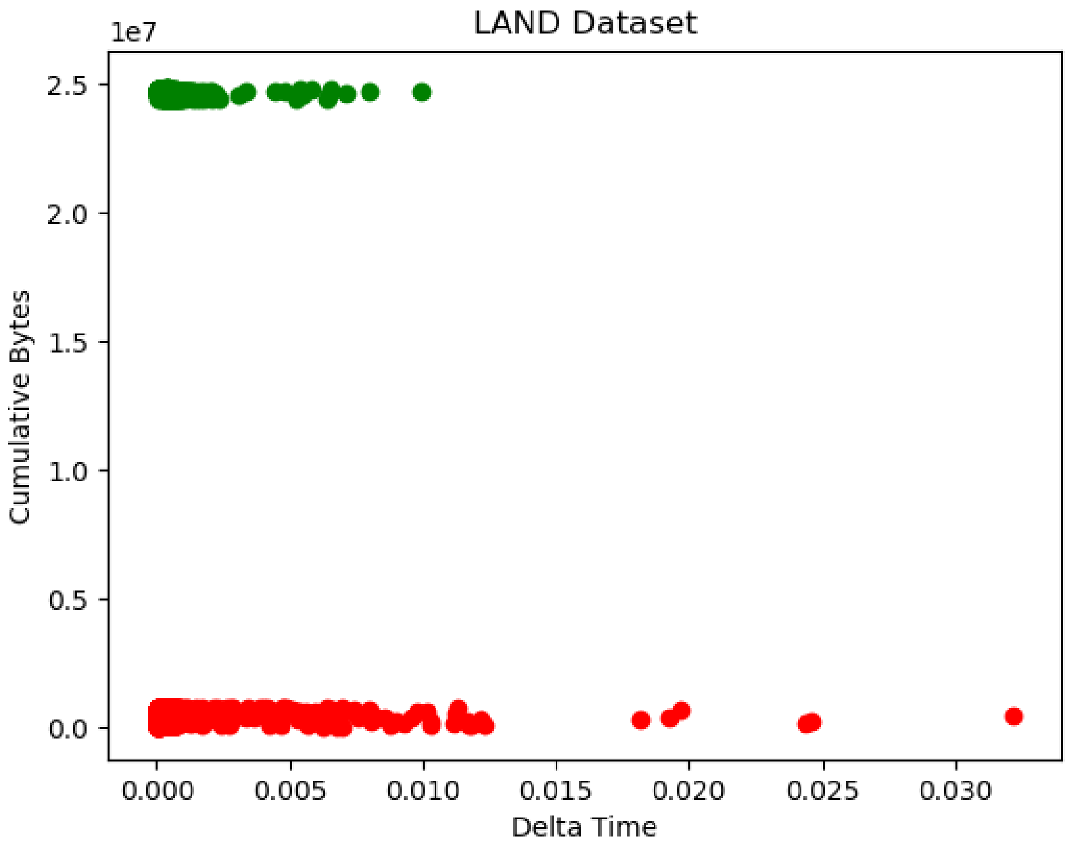
This shows that packets that take lesser time and have cumulative bytes are of normal condition and packets that take more time and do not have cumulative bytes are of flood condition.

LAND



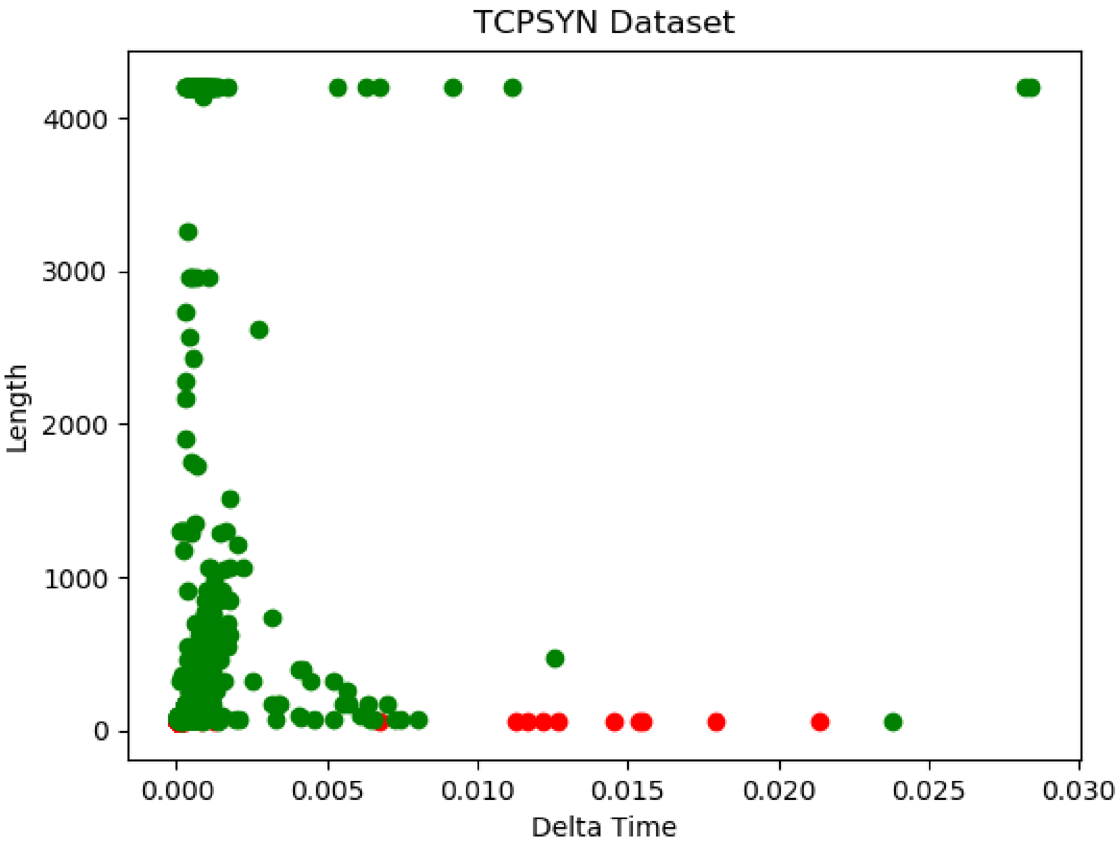
From the graph it is visible that when Delta time is closer to 0 (0 to 0.01), for all value of length, it can be either normal or flood condition thus showing that the following parameters are not enough to differentiate.

Similarly, when delta time is greater than 0.01, for all values of length, it definitely is a flood condition.



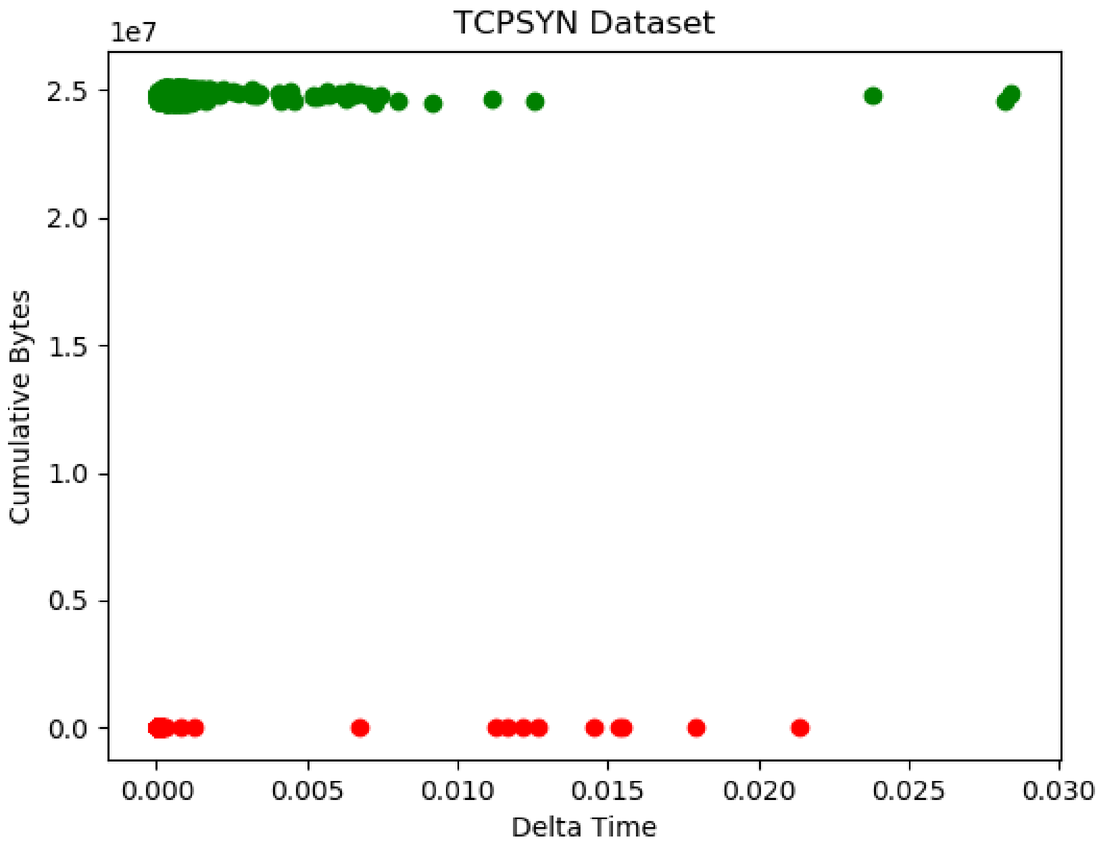
From the graph, it is visible that cumulative bytes alone determines flood or normal conditions. For higher values of cumulative bytes, for all value of delta time, it is a normal condition. Similarly, for lower values of cumulative bytes, for all values of delta time, it is a flood condition.

TCPSYN



From the graph it is visible that flood conditions only arise when length is almost 0 and delta time is in the range [0.070 to 0.022]. Hence flood conditions depend on both parameters.

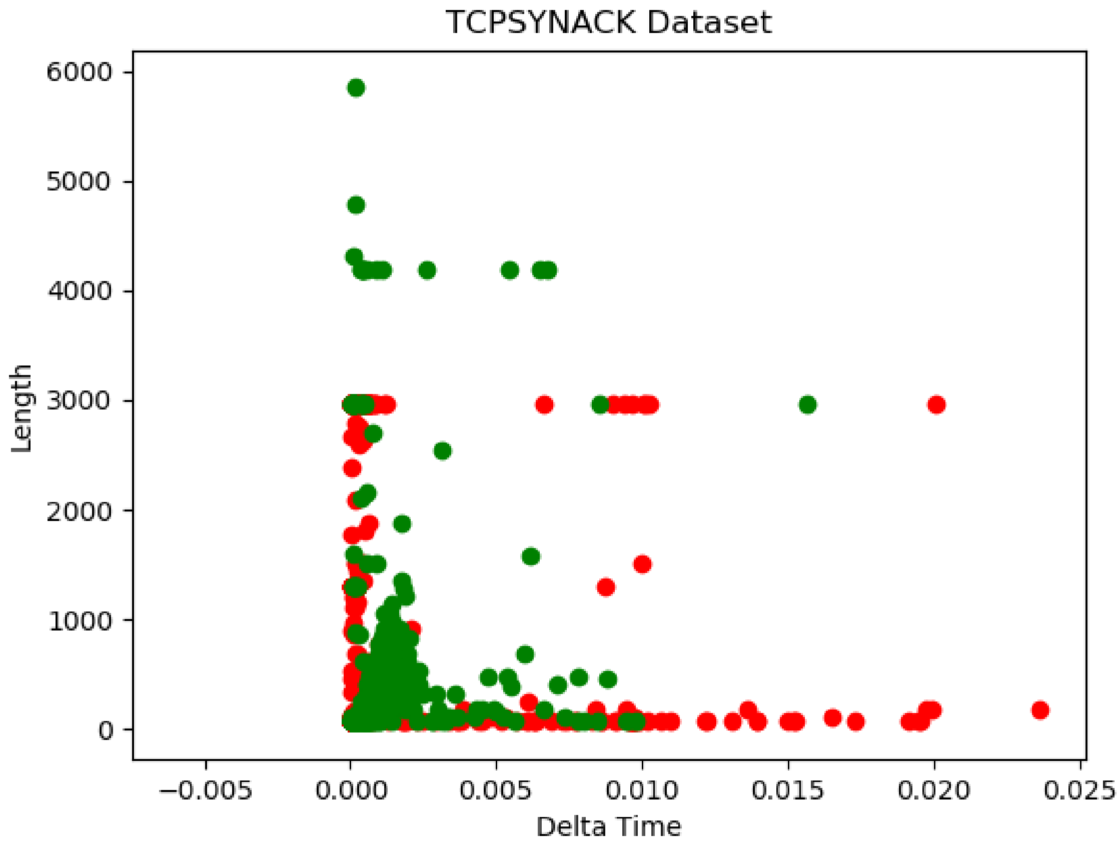
For all other values of Length and Delta Time, there exist normal conditions



From the graph it is visible that for all Cumulative bytes alone determines the conditions. I.e. when cumulative bytes are high, for all values of Delta time, there is normal conditions.

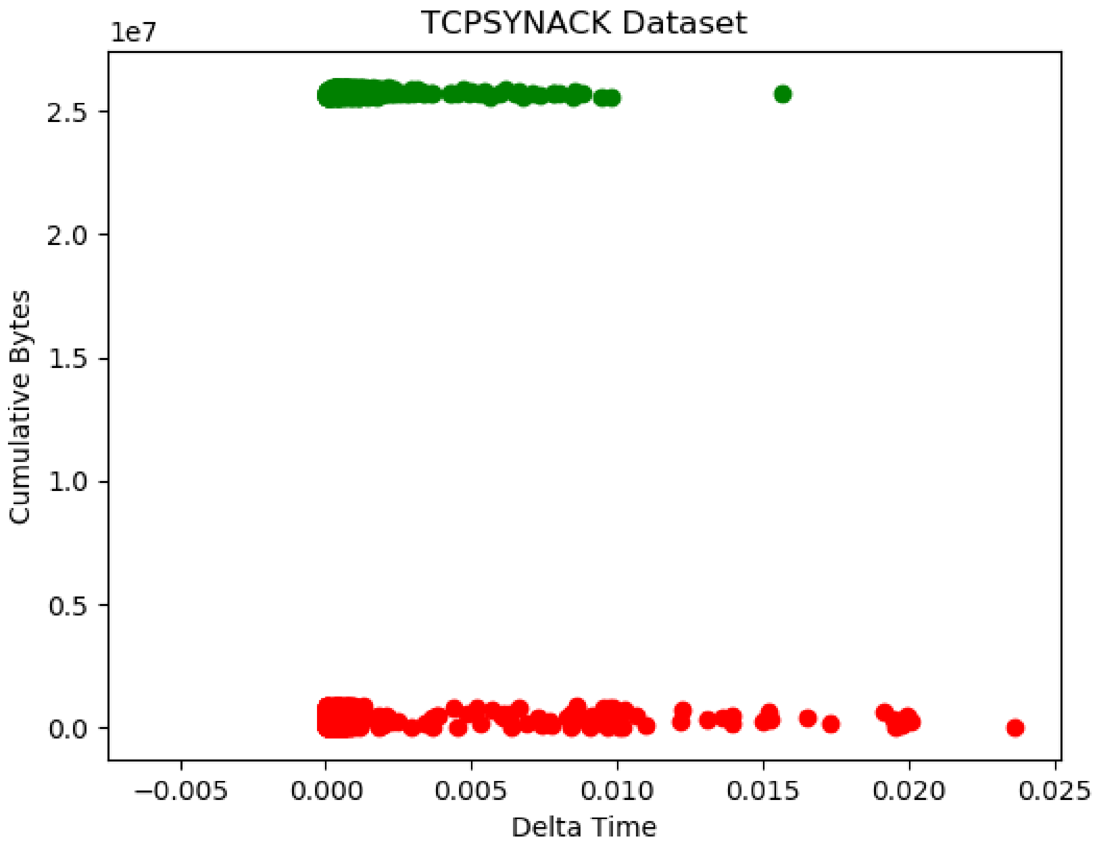
Similarly when cumulative bytes are 0, for all values of Delta Time, there is flood conditions

TCPSYNACK



From the graph we can see that when delta time lies between 0 and 0.01 and length lies between 0 and 3000, it cannot be determined whether there is flood or normal conditions as both conditions occur, hence these two parameters are not enough.
For the same range of delta time (0 to 0.01) and length > 3000, only normal conditions exist.

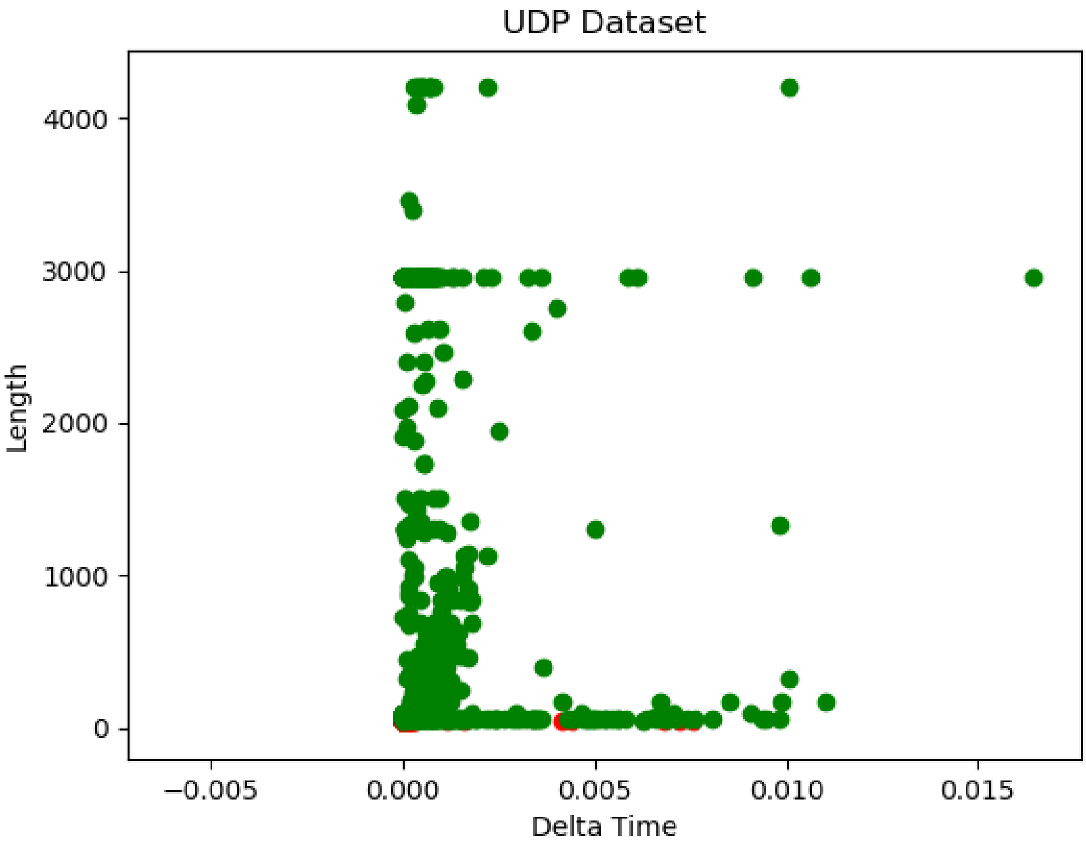
When delta time is greater than 0.01, for all values of length, flood conditions exist.



From the graph, we can see that cumulative bytes solely determine flood or normal conditions. When cumulative bytes is 0 then for all values of Delta time it is flood conditions.

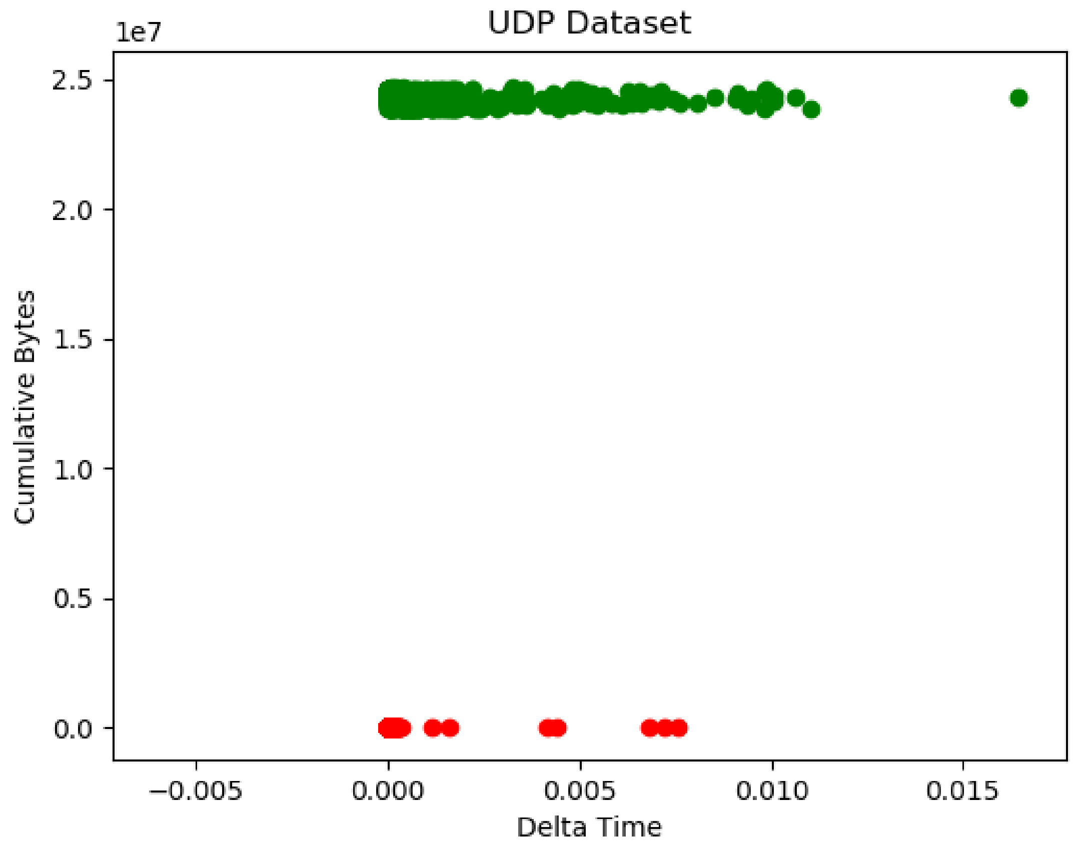
Similarly, when cumulative bytes are greater than 0, then for all values of Delta time, normal conditions exist.

UDP



From the graph it is visible that for length 0 we cannot determine whether flood conditions or normal conditions exist as for all values of delta time, many have both conditions.

When length is not 0 then it solely determines that there is normal conditions as for all values of delta time, there is normal conditions.



From the graph it is visible that cumulative bytes solely determine the condition. For Cumulative bytes equal to 0, for all values of Delta time flood conditions exist.

Similarly for higher values of cumulative bytes, for all values of delta time normal conditions exist.

OUTPUTS BY VARYING ALGORITHM PARAMETERS:

Table1:

1. OneVsRestClassifier(LogisticRegression())
2. BaggingClassifier()
3. MultinomialNB()

Table2:

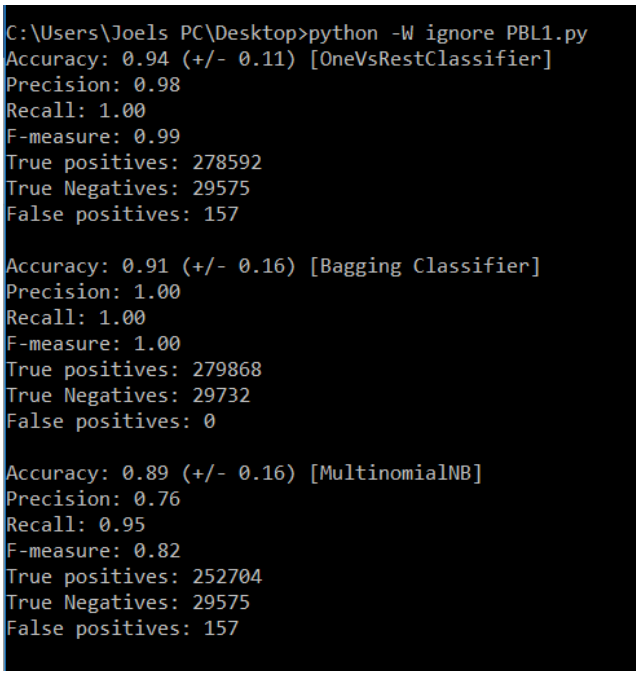
1. OneVsRestClassifier(GaussianNB())
2. BaggingClassifier(max_samples=200)
3. MultinomialNB(alpha=2)

Table3:

1. OneVsRestClassifier(RandomForestClassifier())
2. BaggingClassifier(n_estimators=5,max_samples=200)
3. MultinomialNB(alpha=2,fit_prior=False)

ICMP

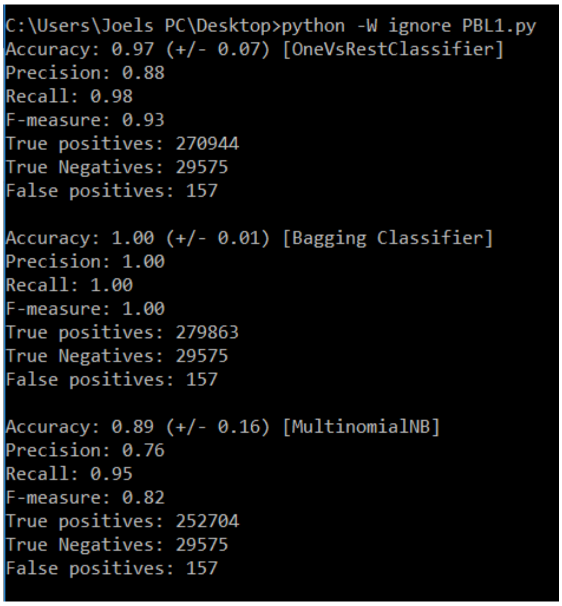
TABLE1



```
C:\Users\Joels PC\Desktop>python -W ignore PBL1.py
Accuracy: 0.94 (+/- 0.11) [OneVsRestClassifier]
Precision: 0.98
Recall: 1.00
F-measure: 0.99
True positives: 278592
True Negatives: 29575
False positives: 157

Accuracy: 0.91 (+/- 0.16) [Bagging Classifier]
Precision: 1.00
Recall: 1.00
F-measure: 1.00
True positives: 279868
True Negatives: 29732
False positives: 0

Accuracy: 0.89 (+/- 0.16) [MultinomialNB]
Precision: 0.76
Recall: 0.95
F-measure: 0.82
True positives: 252704
True Negatives: 29575
False positives: 157
```

S. No	Model	Accuracy (%)
1	OneVsRestClassifier	94
2	BaggingClassifier	91
3	MultinomialNB	89

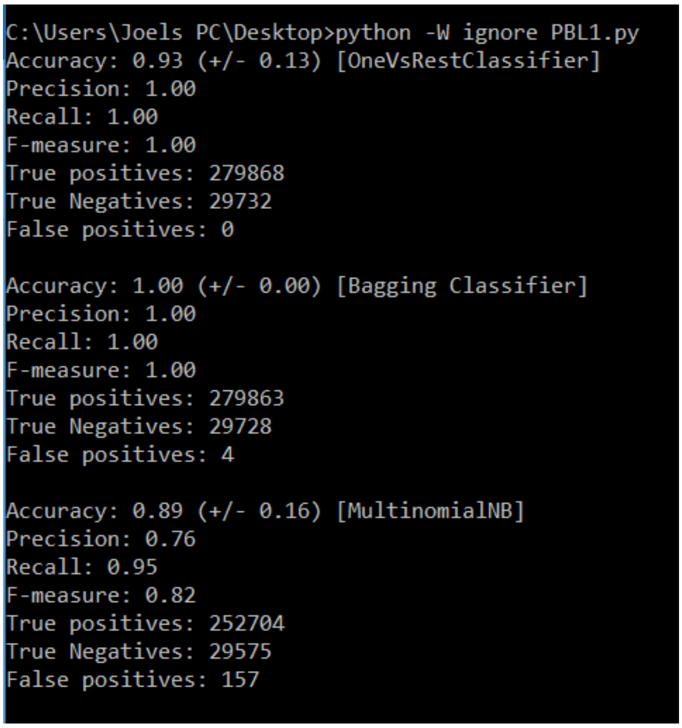
TABLE2



```
C:\Users\Joels PC\Desktop>python -W ignore PBL1.py
Accuracy: 0.97 (+/- 0.07) [OneVsRestClassifier]
Precision: 0.88
Recall: 0.98
F-measure: 0.93
True positives: 270944
True Negatives: 29575
False positives: 157

Accuracy: 1.00 (+/- 0.01) [Bagging Classifier]
Precision: 1.00
Recall: 1.00
F-measure: 1.00
True positives: 279863
True Negatives: 29575
False positives: 157

Accuracy: 0.89 (+/- 0.16) [MultinomialNB]
Precision: 0.76
Recall: 0.95
F-measure: 0.82
True positives: 252704
True Negatives: 29575
False positives: 157
```

S. No	**Model**	**Accuracy (%)**
1	OneVsRestClassifier	97
2	BaggingClassifier	100
3	MultinomialNB	89

TABLE3

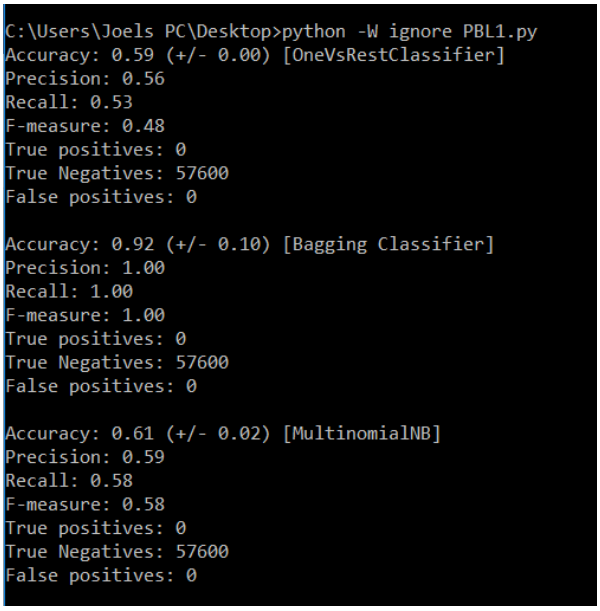


```
C:\Users\Joels PC\Desktop>python -W ignore PBL1.py
Accuracy: 0.93 (+/- 0.13) [OneVsRestClassifier]
Precision: 1.00
Recall: 1.00
F-measure: 1.00
True positives: 279868
True Negatives: 29732
False positives: 0

Accuracy: 1.00 (+/- 0.00) [Bagging Classifier]
Precision: 1.00
Recall: 1.00
F-measure: 1.00
True positives: 279863
True Negatives: 29728
False positives: 4

Accuracy: 0.89 (+/- 0.16) [MultinomialNB]
Precision: 0.76
Recall: 0.95
F-measure: 0.82
True positives: 252704
True Negatives: 29575
False positives: 157
```

S. No	Model	Accuracy (%)
1	OneVsRestClassifier	93
2	BaggingClassifier	100
3	MultinomialNB	89

LAND

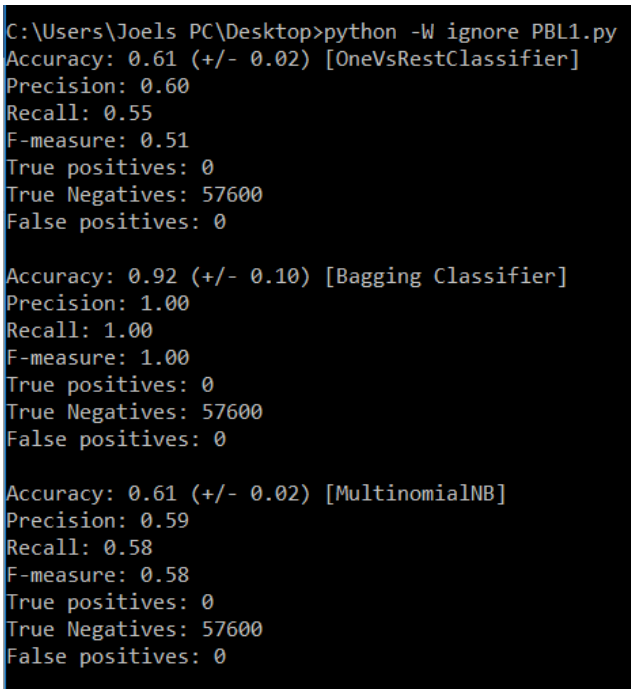
TABLE1



```
C:\Users\Joels PC\Desktop>python -W ignore PBL1.py
Accuracy: 0.59 (+/- 0.00) [OneVsRestClassifier]
Precision: 0.56
Recall: 0.53
F-measure: 0.48
True positives: 0
True Negatives: 57600
False positives: 0

Accuracy: 0.92 (+/- 0.10) [Bagging Classifier]
Precision: 1.00
Recall: 1.00
F-measure: 1.00
True positives: 0
True Negatives: 57600
False positives: 0

Accuracy: 0.61 (+/- 0.02) [MultinomialNB]
Precision: 0.59
Recall: 0.58
F-measure: 0.58
True positives: 0
True Negatives: 57600
False positives: 0
```

S. No	Model	Accuracy (%)
1	OneVsRestClassifier	59
2	BaggingClassifier	92
3	MultinomialNB	61

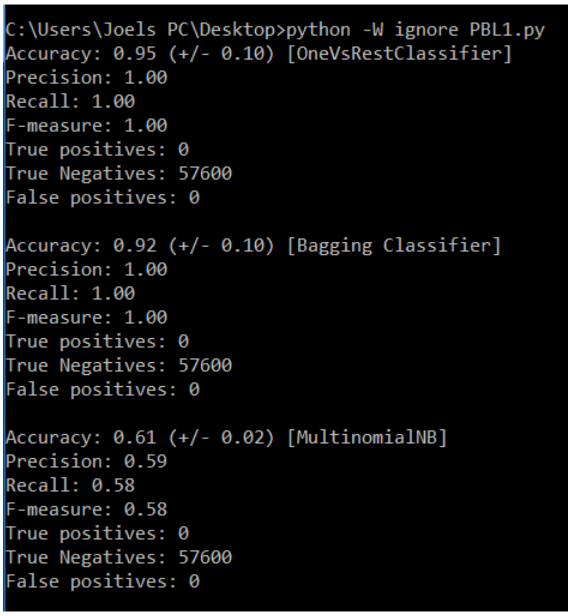
TABLE2



```
C:\Users\Joels PC\Desktop>python -W ignore PBL1.py
Accuracy: 0.61 (+/- 0.02) [OneVsRestClassifier]
Precision: 0.60
Recall: 0.55
F-measure: 0.51
True positives: 0
True Negatives: 57600
False positives: 0

Accuracy: 0.92 (+/- 0.10) [Bagging Classifier]
Precision: 1.00
Recall: 1.00
F-measure: 1.00
True positives: 0
True Negatives: 57600
False positives: 0

Accuracy: 0.61 (+/- 0.02) [MultinomialNB]
Precision: 0.59
Recall: 0.58
F-measure: 0.58
True positives: 0
True Negatives: 57600
False positives: 0
```

S. No	Model	Accuracy (%)
1	OneVsRestClassifier	61
2	BaggingClassifier	92
3	MultinomialNB	61

TABLE3

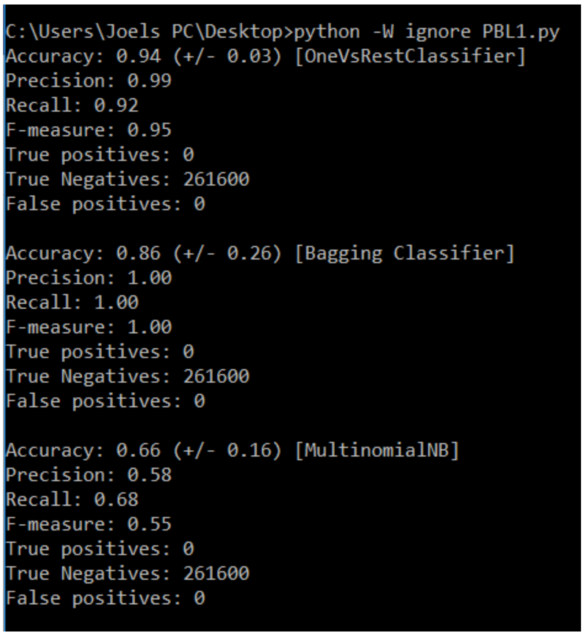


```
C:\Users\Joels PC\Desktop>python -W ignore PBL1.py
Accuracy: 0.95 (+/- 0.10) [OneVsRestClassifier]
Precision: 1.00
Recall: 1.00
F-measure: 1.00
True positives: 0
True Negatives: 57600
False positives: 0

Accuracy: 0.92 (+/- 0.10) [Bagging Classifier]
Precision: 1.00
Recall: 1.00
F-measure: 1.00
True positives: 0
True Negatives: 57600
False positives: 0

Accuracy: 0.61 (+/- 0.02) [MultinomialNB]
Precision: 0.59
Recall: 0.58
F-measure: 0.58
True positives: 0
True Negatives: 57600
False positives: 0
```

S. No	**Model**	**Accuracy (%)**
1	OneVsRestClassifier	95
2	BaggingClassifier	92
3	MultinomialNB	61

TCPSYN

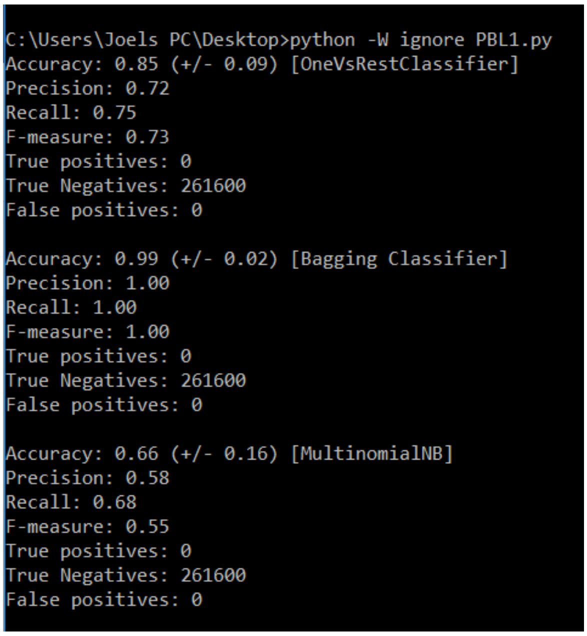
TABLE1



```
C:\Users\Joels PC\Desktop>python -W ignore PBL1.py
Accuracy: 0.94 (+/- 0.03) [OneVsRestClassifier]
Precision: 0.99
Recall: 0.92
F-measure: 0.95
True positives: 0
True Negatives: 261600
False positives: 0

Accuracy: 0.86 (+/- 0.26) [Bagging Classifier]
Precision: 1.00
Recall: 1.00
F-measure: 1.00
True positives: 0
True Negatives: 261600
False positives: 0

Accuracy: 0.66 (+/- 0.16) [MultinomialNB]
Precision: 0.58
Recall: 0.68
F-measure: 0.55
True positives: 0
True Negatives: 261600
False positives: 0
```

S. No	Model	Accuracy (%)
1	OneVsRestClassifier	94
2	BaggingClassifier	86
3	MultinomialNB	66

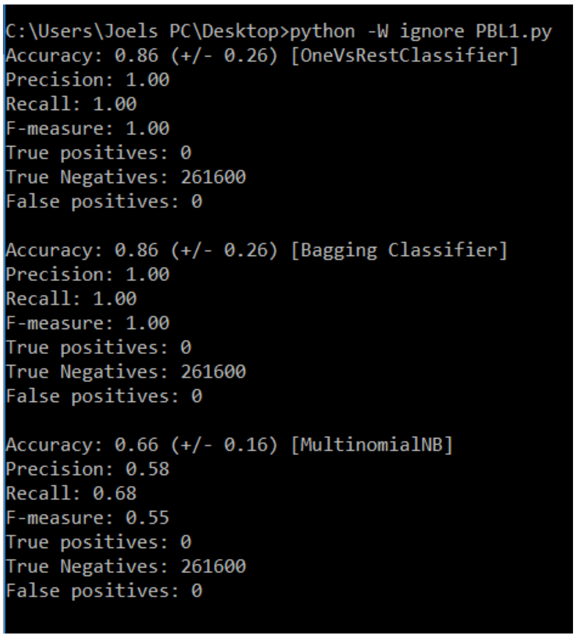
TABLE2



```
C:\Users\Joels PC\Desktop>python -W ignore PBL1.py
Accuracy: 0.85 (+/- 0.09) [OneVsRestClassifier]
Precision: 0.72
Recall: 0.75
F-measure: 0.73
True positives: 0
True Negatives: 261600
False positives: 0

Accuracy: 0.99 (+/- 0.02) [Bagging Classifier]
Precision: 1.00
Recall: 1.00
F-measure: 1.00
True positives: 0
True Negatives: 261600
False positives: 0

Accuracy: 0.66 (+/- 0.16) [MultinomialNB]
Precision: 0.58
Recall: 0.68
F-measure: 0.55
True positives: 0
True Negatives: 261600
False positives: 0
```

S. No	Model	Accuracy (%)
1	OneVsRestClassifier	85
2	BaggingClassifier	99
3	MultinomialNB	66

TABLE3

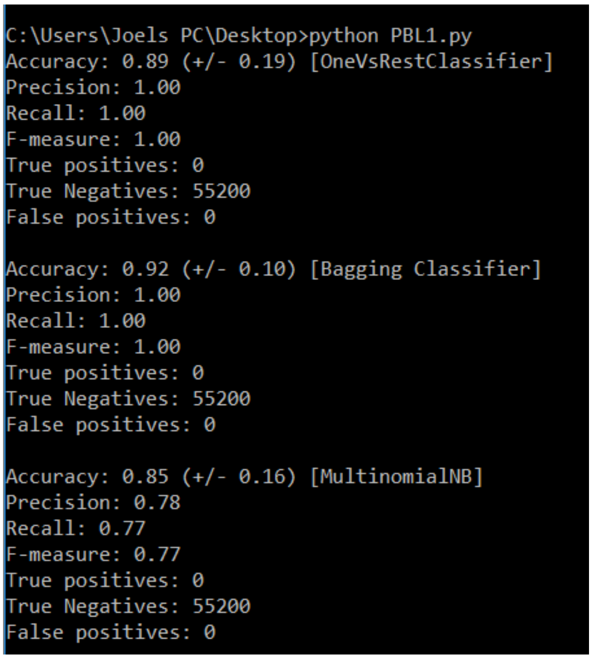


```
C:\Users\Joels PC\Desktop>python -W ignore PBL1.py
Accuracy: 0.86 (+/- 0.26) [OneVsRestClassifier]
Precision: 1.00
Recall: 1.00
F-measure: 1.00
True positives: 0
True Negatives: 261600
False positives: 0

Accuracy: 0.86 (+/- 0.26) [Bagging Classifier]
Precision: 1.00
Recall: 1.00
F-measure: 1.00
True positives: 0
True Negatives: 261600
False positives: 0

Accuracy: 0.66 (+/- 0.16) [MultinomialNB]
Precision: 0.58
Recall: 0.68
F-measure: 0.55
True positives: 0
True Negatives: 261600
False positives: 0
```

S. No	Model	Accuracy (%)
1	OneVsRestClassifier	86
2	BaggingClassifier	86
3	MultinomialNB	66

TCPSYNACK

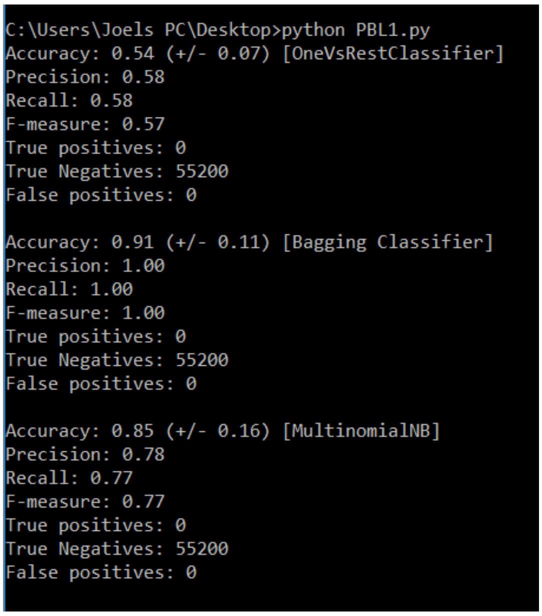
TABLE1



```
C:\Users\Joels PC\Desktop>python PBL1.py
Accuracy: 0.89 (+/- 0.19) [OneVsRestClassifier]
Precision: 1.00
Recall: 1.00
F-measure: 1.00
True positives: 0
True Negatives: 55200
False positives: 0

Accuracy: 0.92 (+/- 0.10) [Bagging Classifier]
Precision: 1.00
Recall: 1.00
F-measure: 1.00
True positives: 0
True Negatives: 55200
False positives: 0

Accuracy: 0.85 (+/- 0.16) [MultinomialNB]
Precision: 0.78
Recall: 0.77
F-measure: 0.77
True positives: 0
True Negatives: 55200
False positives: 0
```

S. No	Model	Accuracy (%)
1	OneVsRestClassifier	89
2	BaggingClassifier	92
3	MultinomialNB	85

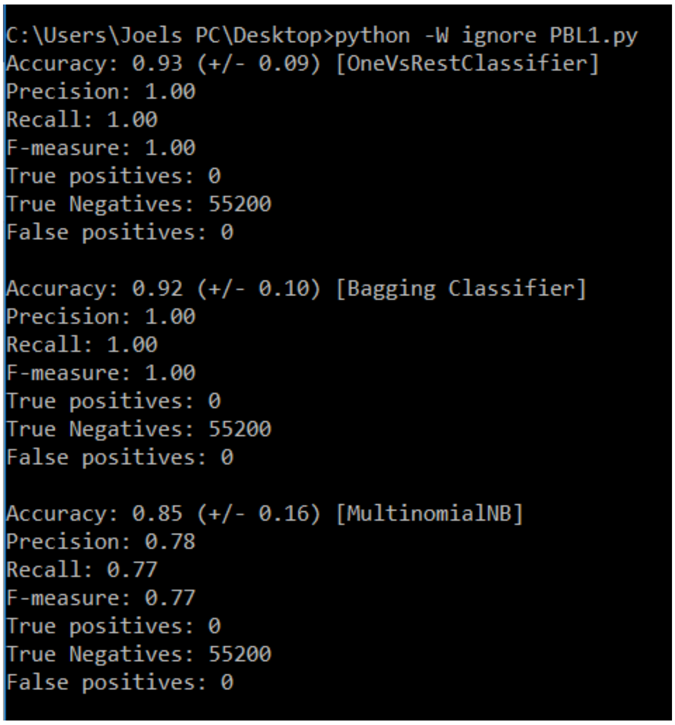
TABLE2



```
C:\Users\Joels PC\Desktop>python PBL1.py
Accuracy: 0.54 (+/- 0.07) [OneVsRestClassifier]
Precision: 0.58
Recall: 0.58
F-measure: 0.57
True positives: 0
True Negatives: 55200
False positives: 0

Accuracy: 0.91 (+/- 0.11) [Bagging Classifier]
Precision: 1.00
Recall: 1.00
F-measure: 1.00
True positives: 0
True Negatives: 55200
False positives: 0

Accuracy: 0.85 (+/- 0.16) [MultinomialNB]
Precision: 0.78
Recall: 0.77
F-measure: 0.77
True positives: 0
True Negatives: 55200
False positives: 0
```

S. No	Model	Accuracy (%)
1	OneVsRestClassifier	54
2	BaggingClassifier	91
3	MultinomialNB	85

TABLE3

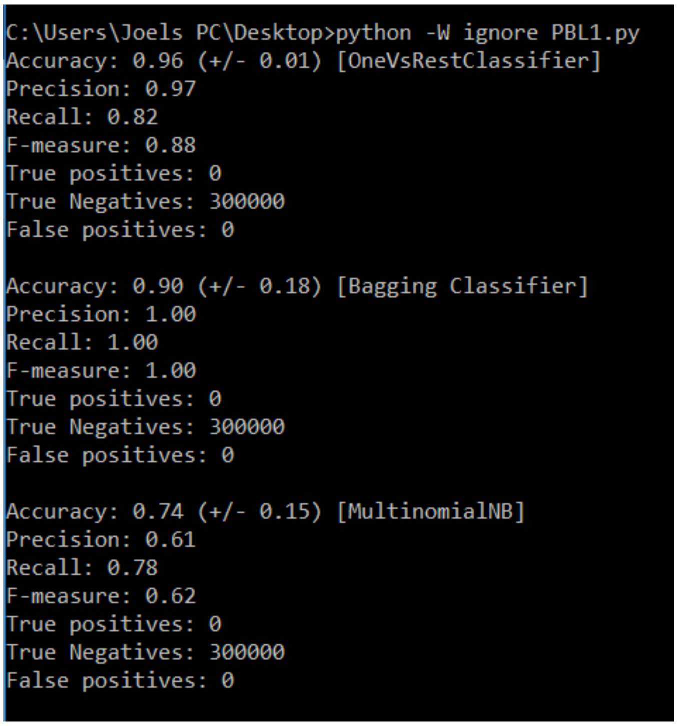


```
C:\Users\Joels PC\Desktop>python -W ignore PBL1.py
Accuracy: 0.93 (+/- 0.09) [OneVsRestClassifier]
Precision: 1.00
Recall: 1.00
F-measure: 1.00
True positives: 0
True Negatives: 55200
False positives: 0

Accuracy: 0.92 (+/- 0.10) [Bagging Classifier]
Precision: 1.00
Recall: 1.00
F-measure: 1.00
True positives: 0
True Negatives: 55200
False positives: 0

Accuracy: 0.85 (+/- 0.16) [MultinomialNB]
Precision: 0.78
Recall: 0.77
F-measure: 0.77
True positives: 0
True Negatives: 55200
False positives: 0
```

S. No	Model	Accuracy (%)
1	OneVsRestClassifier	93
2	BaggingClassifier	92
3	MultinomialNB	85

UDP

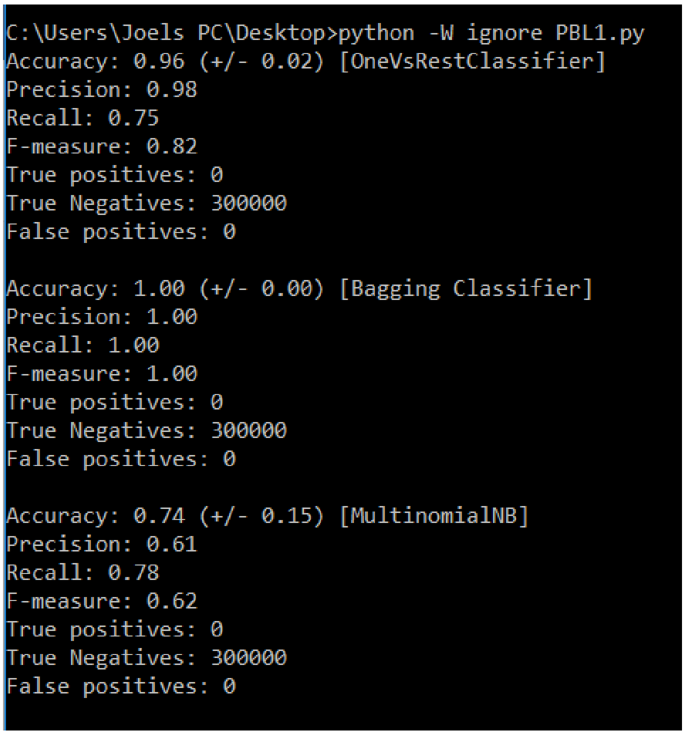
TABLE1



```
C:\Users\Joels PC\Desktop>python -W ignore PBL1.py
Accuracy: 0.96 (+/- 0.01) [OneVsRestClassifier]
Precision: 0.97
Recall: 0.82
F-measure: 0.88
True positives: 0
True Negatives: 300000
False positives: 0

Accuracy: 0.90 (+/- 0.18) [Bagging Classifier]
Precision: 1.00
Recall: 1.00
F-measure: 1.00
True positives: 0
True Negatives: 300000
False positives: 0

Accuracy: 0.74 (+/- 0.15) [MultinomialNB]
Precision: 0.61
Recall: 0.78
F-measure: 0.62
True positives: 0
True Negatives: 300000
False positives: 0
```

S. No	Model	Accuracy (%)
1	OneVsRestClassifier	96
2	BaggingClassifier	90
3	MultinomialNB	74

TABLE2

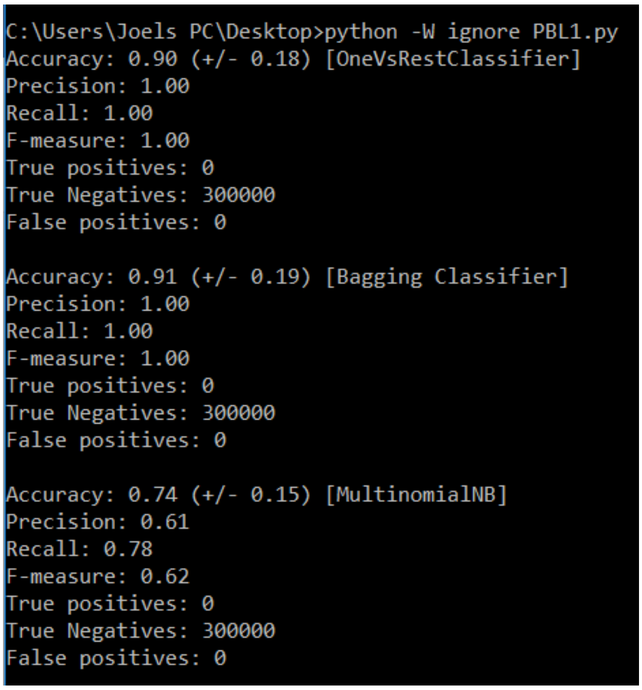


```
C:\Users\Joels PC\Desktop>python -W ignore PBL1.py
Accuracy: 0.96 (+/- 0.02) [OneVsRestClassifier]
Precision: 0.98
Recall: 0.75
F-measure: 0.82
True positives: 0
True Negatives: 300000
False positives: 0

Accuracy: 1.00 (+/- 0.00) [Bagging Classifier]
Precision: 1.00
Recall: 1.00
F-measure: 1.00
True positives: 0
True Negatives: 300000
False positives: 0

Accuracy: 0.74 (+/- 0.15) [MultinomialNB]
Precision: 0.61
Recall: 0.78
F-measure: 0.62
True positives: 0
True Negatives: 300000
False positives: 0
```

S. No	Model	Accuracy (%)
1	OneVsRestClassifier	96
2	BaggingClassifier	100

3	MultinomialNB	74

TABLE3



```
C:\Users\Joels PC\Desktop>python -W ignore PBL1.py
Accuracy: 0.90 (+/- 0.18) [OneVsRestClassifier]
Precision: 1.00
Recall: 1.00
F-measure: 1.00
True positives: 0
True Negatives: 300000
False positives: 0

Accuracy: 0.91 (+/- 0.19) [Bagging Classifier]
Precision: 1.00
Recall: 1.00
F-measure: 1.00
True positives: 0
True Negatives: 300000
False positives: 0

Accuracy: 0.74 (+/- 0.15) [MultinomialNB]
Precision: 0.61
Recall: 0.78
F-measure: 0.62
True positives: 0
True Negatives: 300000
False positives: 0
```

S. No	**Model**	**Accuracy (%)**
1	OneVsRestClassifier	90
2	BaggingClassifier	91
3	MultinomialNB	74