

Machine Learning Project 2

Q1

a) Bagging Routine implemented

b) Results:

1. single classifier

True pos count: 45

Total data count: 60

Accuracy : 75%

2. 10 Bags

True pos count: 53

Total data count: 60

Accuracy : 88.33%

3. 50 Bags

True pos count: 53

Total data count: 60

Accuracy : 88.33%

4. 100 Bags

~~True~~ True pos count: 55

Total data count: 60

Accuracy : 91.66%

A logistic regression model already has low variance. Implementing a bagging routine does not make much difference in accuracy.

~~Aggregate~~

Bootstrap aggregation requires the data points to be independent (have uncorrelated errors) working with a small data set, the probability of the model noise related errors is high.

∴ In this question, the bagging routine is not working as it should

Q2

Single classifier:	85%
10 Iterations:	88.3%
50 Iterations:	88.3%
100 Iterations:	91.6%