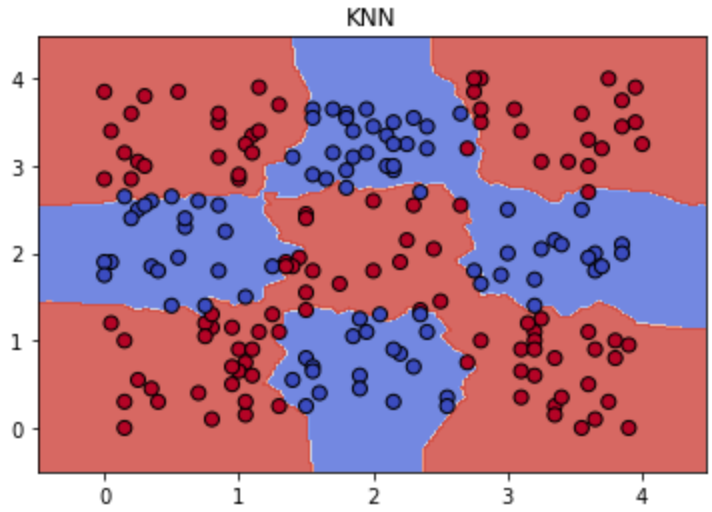
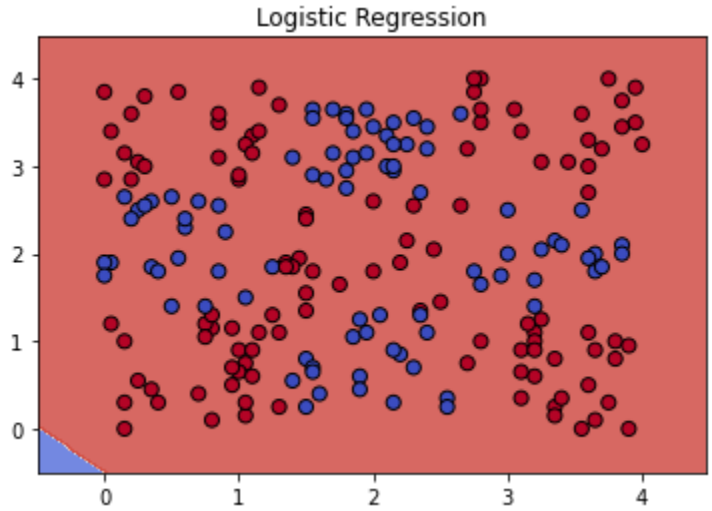
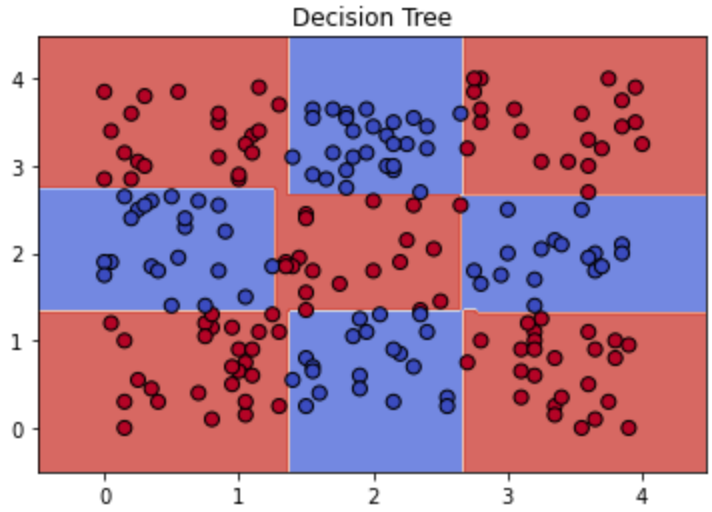


Classifier: KNN
Best Score: 0.93
Training Score: 0.95
Test Score: 0.925

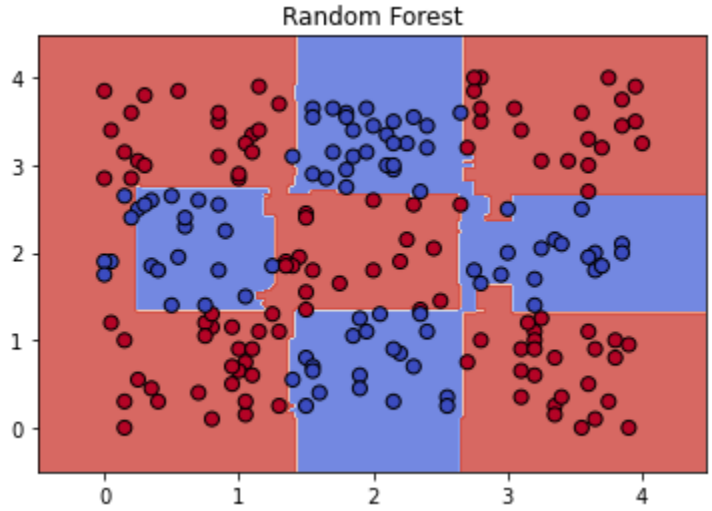
KNN classifier did a good job with a best score of 0.93 and test score of 0.925. The above plot confirms these scores as we see the large portion of red and blue points within the appropriate color scheme. This means that the KNN plot was about 92 to 93 % accurate in classification.



Classifier: Logistic Regression
Best Score: 0.6066666666666667
Training Score: 0.6066666666666667
Test Score: 0.565
Logistic regression did not fare well at all with a test score of about 56% and best score of about 60%. This type of regression model is not at all good at achieving high classification accuracy.

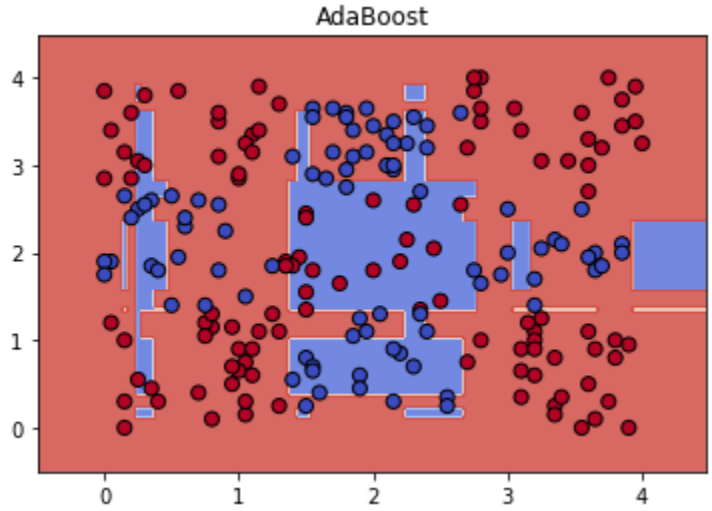


Classifier: Decision Tree
Best Score: 0.9733333333333334
Training Score: 1.0
Test Score: 0.995
The decision tree classifier was the best by far with a best score of 97% accuracy and test score of 99% and training score of 100%. This classification method did a phenomenal job at achieving high classification accuracy.



Classifier: Random Forest
Best Score: 0.9400000000000001
Training Score: 0.9966666666666667
Test Score: 0.95

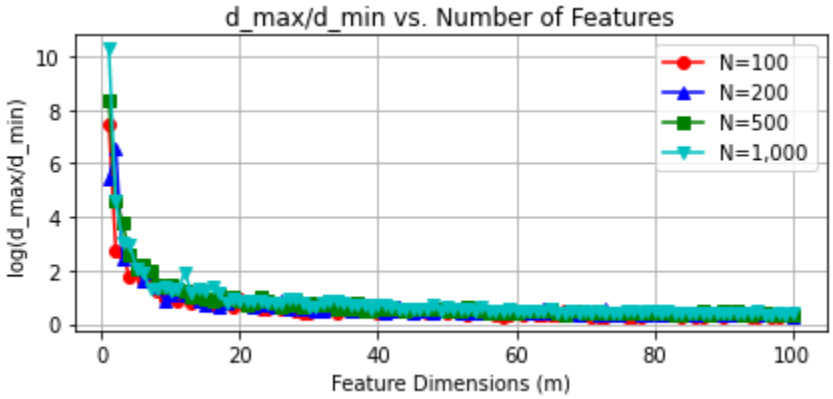
The random forest classifier was the second best with a best score of about 94%, test score of 95% and training score of about 99%. While there are a few clear spots of inaccuracy it is far more reliable to generate a high classification accuracy score.



Classifier: AdaBoost
Best Score: 0.6266666666666667
Training Score: 0.68
Test Score: 0.565

AdaBoost classifier comes in at the second worst score with a best score of about 63%, test score of only 56% and training score of 68%. While it is able to deliver decent results, it is not on par with random forest, decision tree or KNN classifiers.

Part 2



The wikipedia link explores the complex nature of working with high-dimensional data: talking about both the curse and the blessing of dimensionality. This concept of the blessing of dimensionality arose in the late 1900s and is now a key concept in machine learning. The blessing refers to specific benefits, such as random data points being **linearly separable** and the **concentration of measure** (as dimension increases, the distances between these points become nearly the same as one another). These properties can simplify data analysis and make simple methods, like linear classifiers, unexpectedly effective. The blessing is especially apparent when the added dimensions provide **signal** (useful information) rather than **noise** (irrelevant features). The point is that even though it can be difficult to implement, high-dimensional problems can be easily tackled as long as the dimensions provide higher signals and lower noise.