Drew Currie

EELE-491 Homework 2 Writeup

1. Training of the initial model resulted in a model that performed well. Additionally, this model was quick to train and run, compared to later models in the homework.

Accuracy	98.90%
Precision	99.09%
Recall	99.80%
F1 Score	0.99

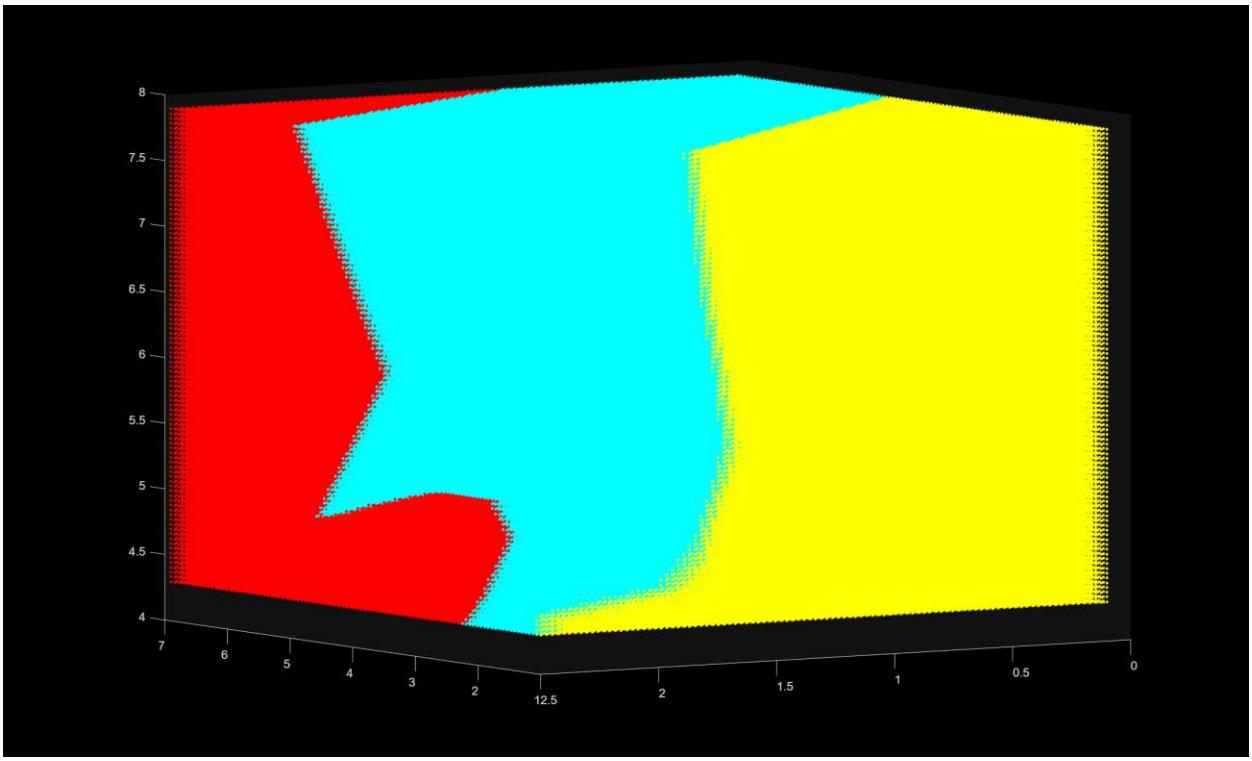
   The boosted and under-sampled LinearSVM seemed to work best at a doubling of the minority class and a 50% down-sampling of the majority class.

Accuracy	97.66%
Precision	99.52%
Recall	98.08%
F1 Score	0.99

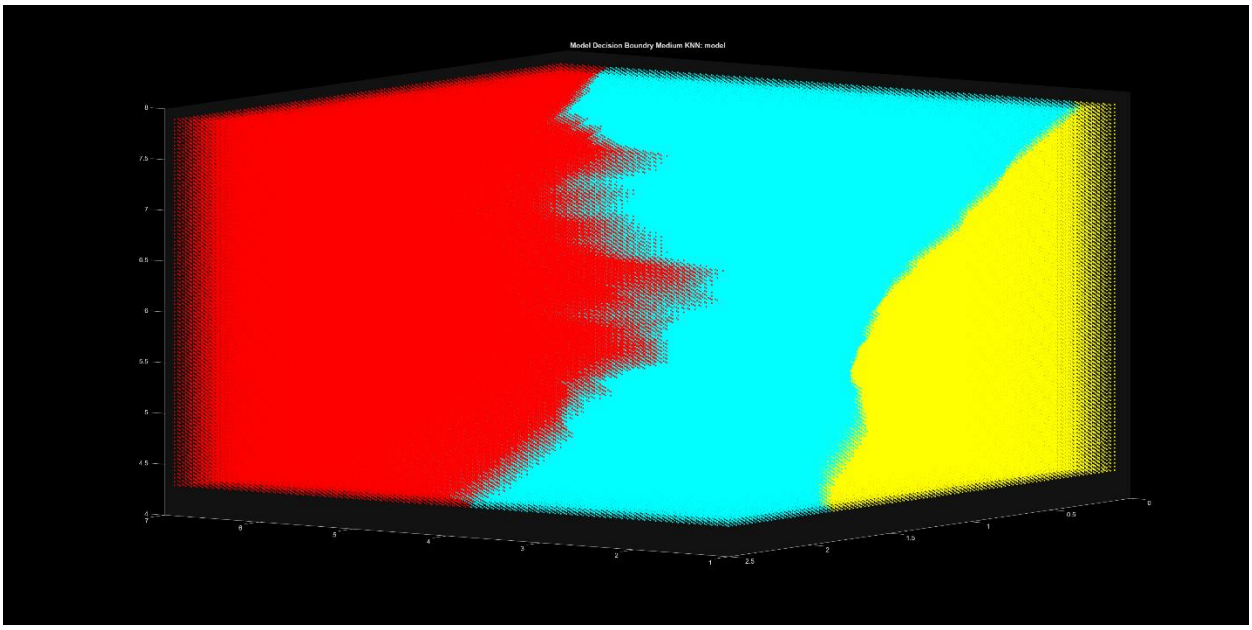
2. See provided PDF of code and results.

3. Plots for this are provided here, code is attached to the upload.

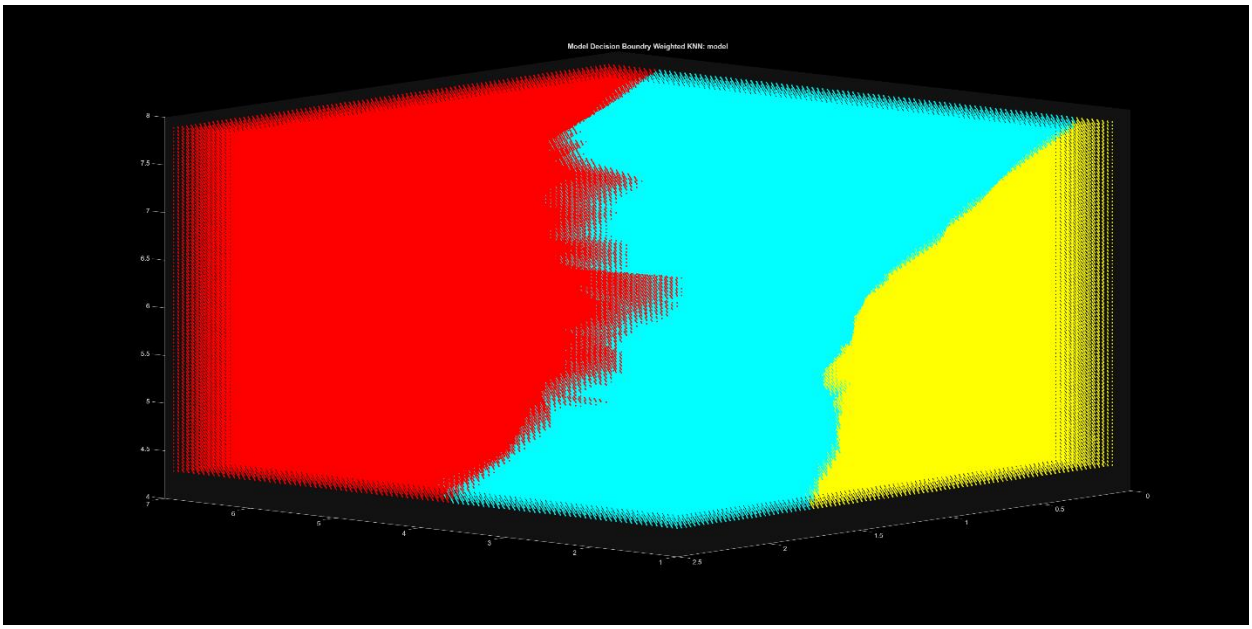
Trilayered Neural Network:



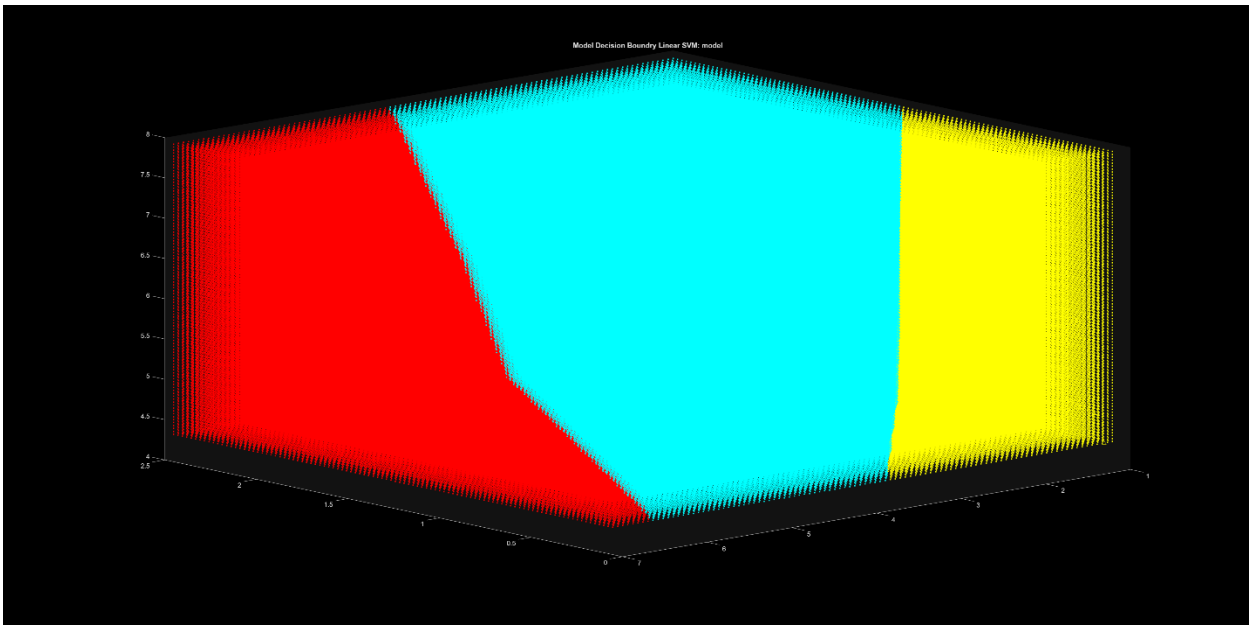
Medium KNN:



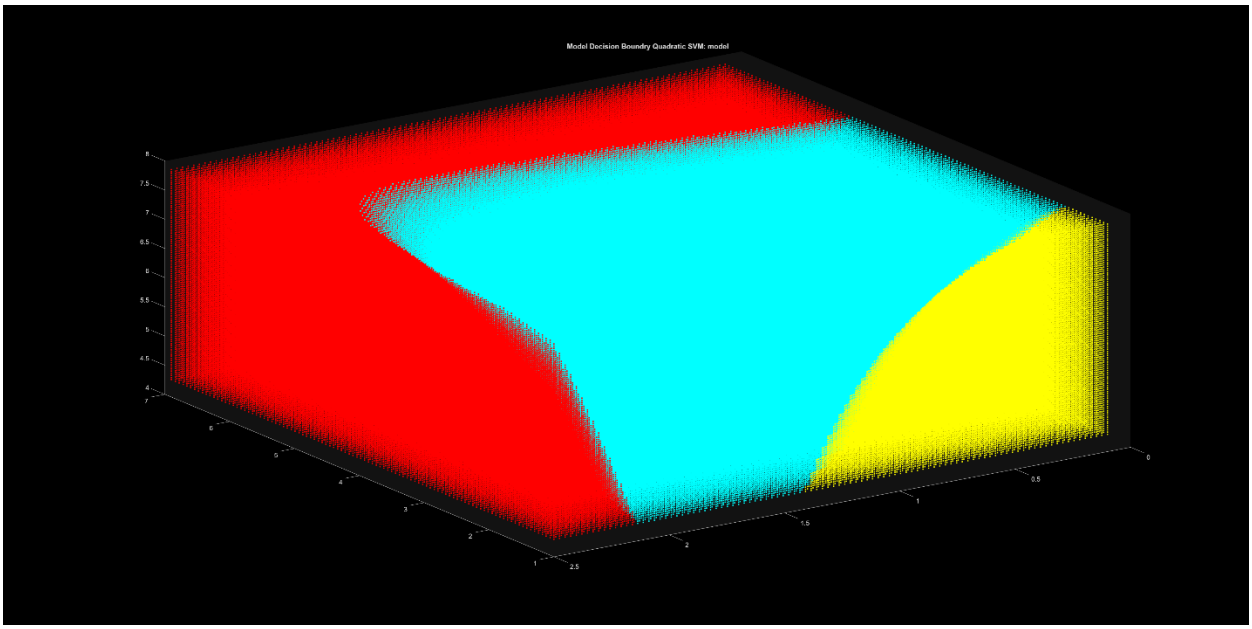
Weighted KNN:



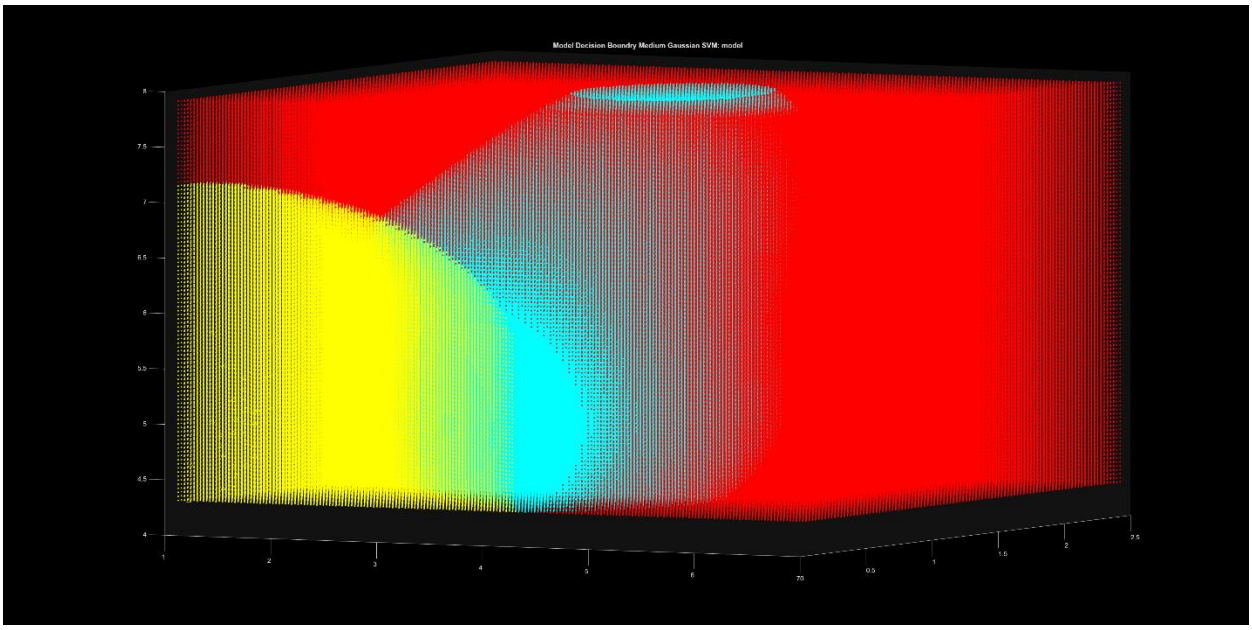
Linear SVM:



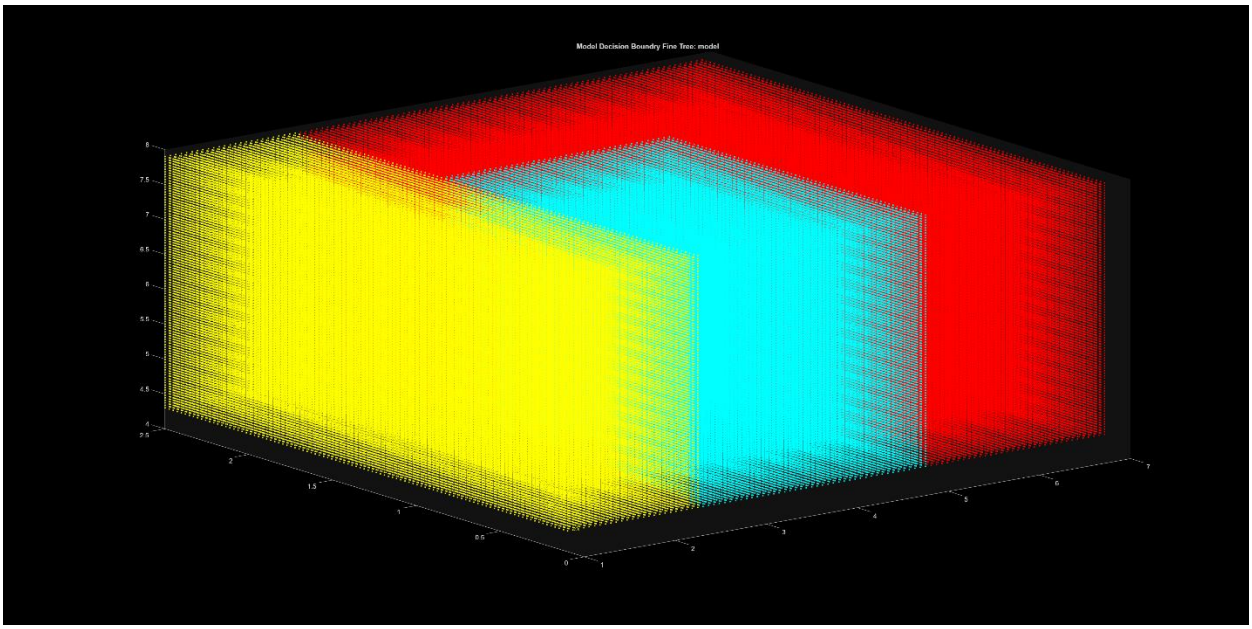
Quadratic SVM:



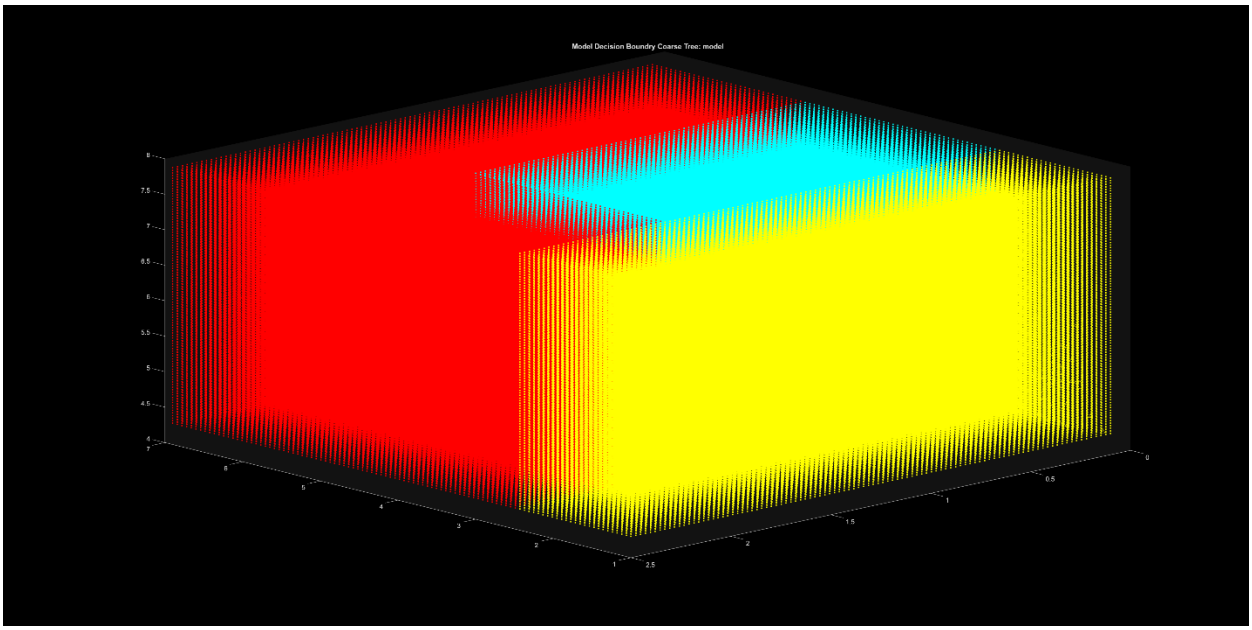
Medium Gaussian SVM:



Fine Tree:


Model Decision Boundry Fine Tree: model

Coarse Tree:


Model Decision Boundry Coarse Tree: model

4. Data is provided as an attached PDF to this assignment. What I found most interesting was how consistent some of these models were. Despite changing the amount of training data, the models were still able to perform very nicely. Perhaps this is more down to the quality of the training and test data compared to the model structure or the method of storing validation set of data. I generally think this shows that in applications that machine learning is applicable, there are a lot of ways to achieve the same results. This means that in a situation where a 100% accuracy, precision, recall, and F1 score is needed but running time isn't important that can be achieved. While a trade off, of only ~2% on average, could decrease training time and run time of the model.