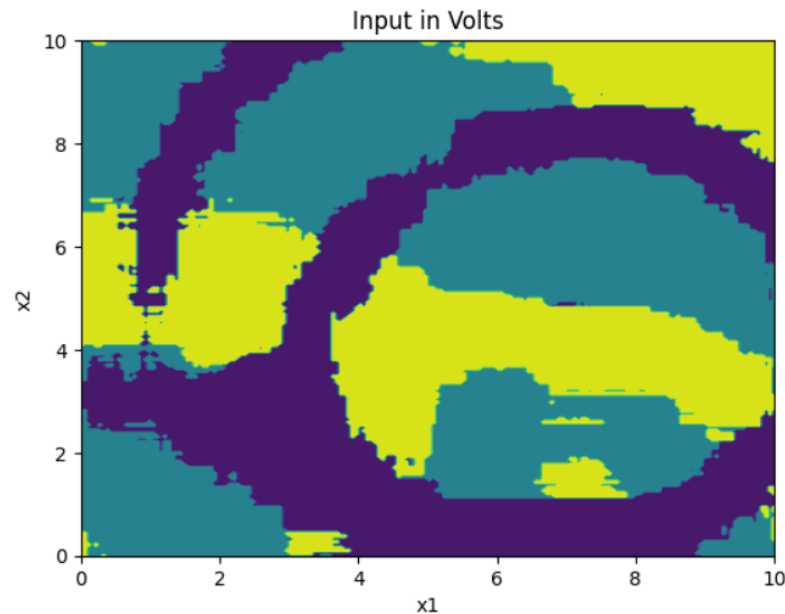


# Machine Learning Coursework 3

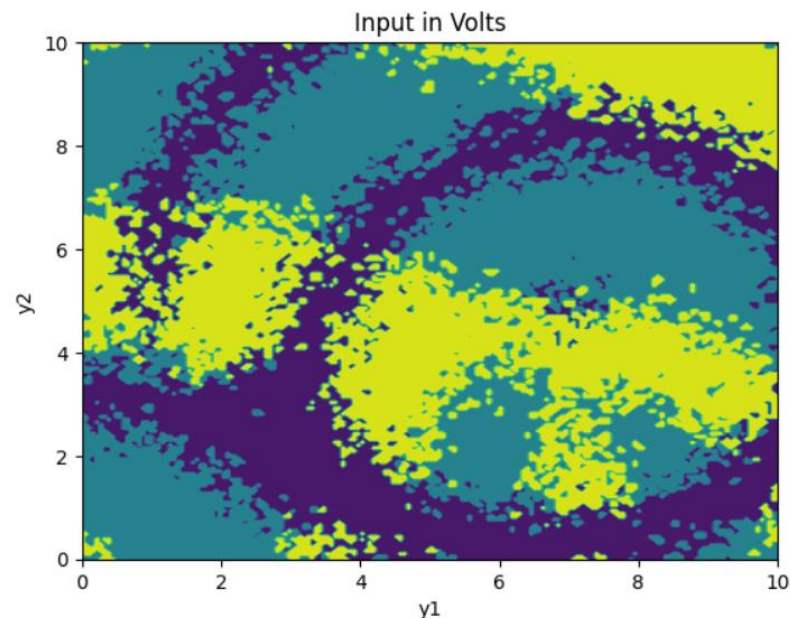
A. The parameters are given as  $y_1, y_2$  as follows:

1. 3.19 1.91
2. 7.20 7.71
3. 2.89 6.40

B. The graph using the random forest.



C. The graph using the nearest neighbour approach.



D.

1. The random forest performs best and is more robust to outliers. Best performance is when the model represents the underlying physics of the problem most accurately, hence smoother regions are preferable.
2. 3D higher order Regression could be used to attain a function with how  $y_1$  varies with  $x$  inputs and another function with how  $y_2$  varies with  $x$  inputs.
3. Yes, a function could be obtained such that the output could be described by the two inputs, however this would be a very high order function to fit this discrete data properly and would be at risk from overfitting.