S1443483

INF2B-CW2 REPORT 3

Task 3

4.1

The Correct Classification Rate for k-nearest neighbours in the 2D-space is 0.3130 as opposed to 0.8520 when we use the full feature space. The Correct Classification Rate for the Gaussian Models is 0.4190 for Gaussian_full in 2D as opposed to 0.8680 when using the full feature space and 0.4010 for Gaussian_lda in 2D as opposed to 0.8560 when using the full feature space. Thus, we conclude that the correct classification rate is reduced when use a smaller part of the full feature space.

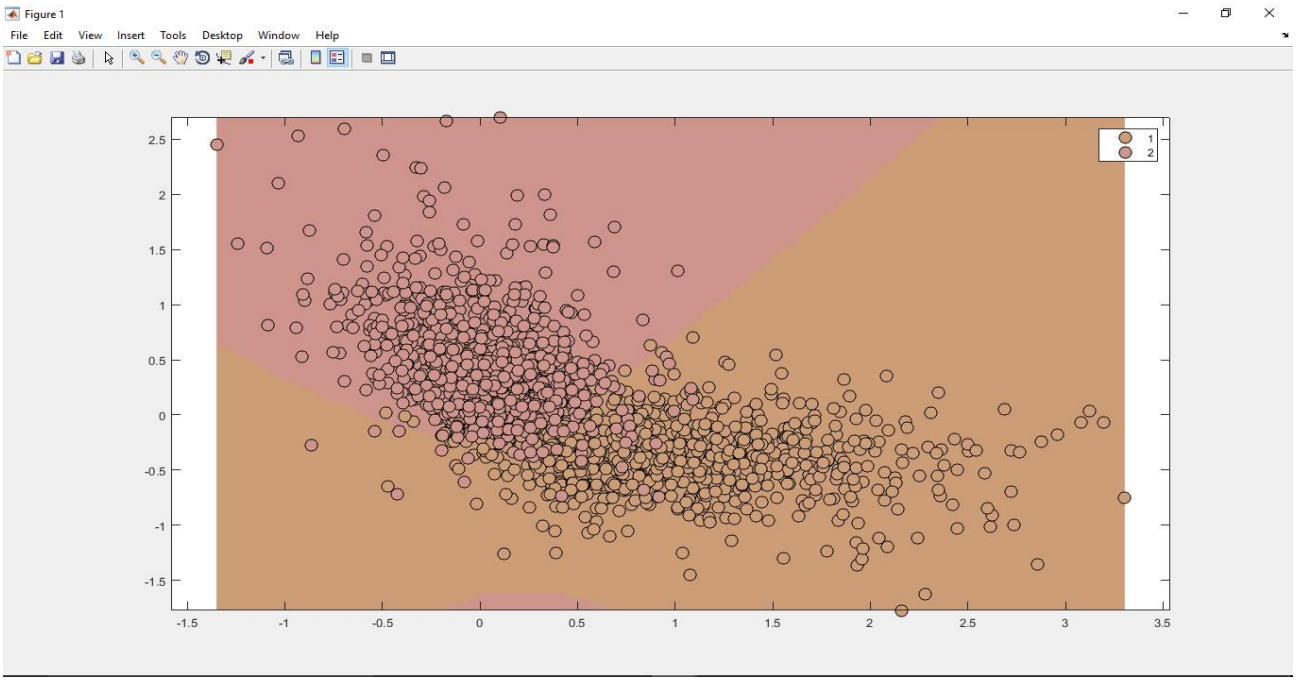
4.2 & 4.3

I wrote the code to visualise decision boundaries and decision regions of the features according to the class given by the K-nearest neighbours classifies and of the two types of Gaussian Models classifiers.

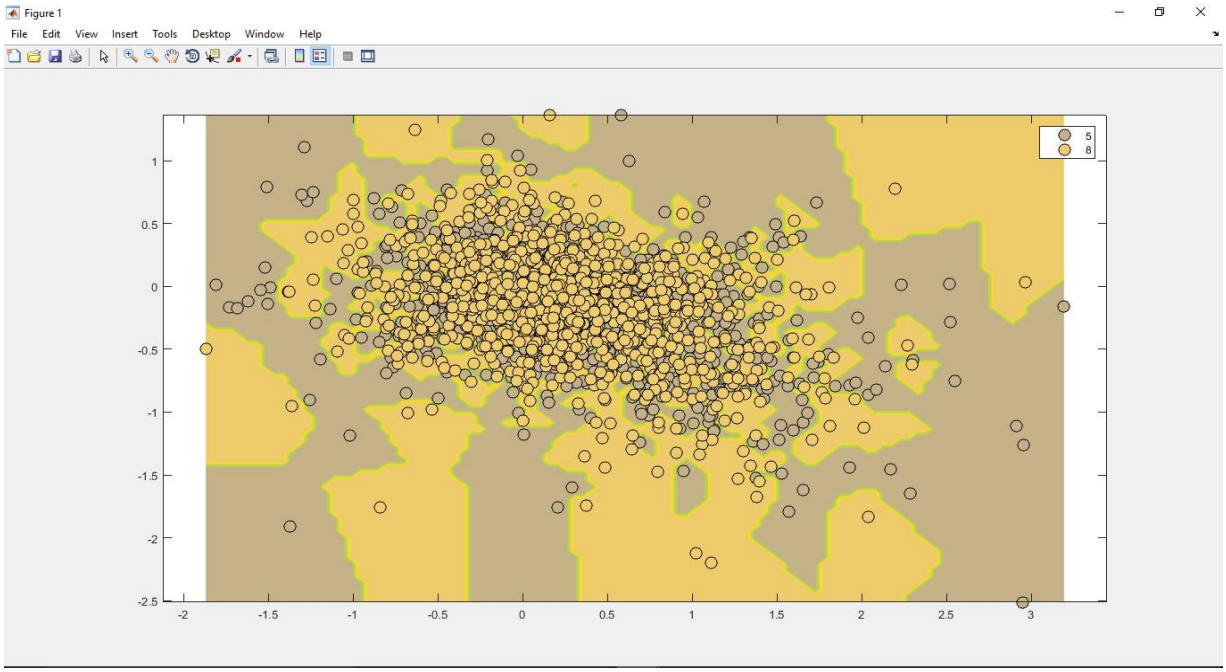
4.4

The shapes of the KNN method decision boundaries are non-specific and kind of messy. They look like pieces of land, resembling countries. There are even separate parts of the same classes' decision boundaries, as if one decision boundary is a country and it has colonies all over the map (plot), resulting in several same class decision boundaries existing in different places. On the other hand, the Gaussian Models methods decision boundaries have straight lines which separate each class from one another. We could say they look like trapezoids, with no specific similarity between their sides. Contrary to the KNN method, these boundaries have not separate parts. In other places, so it would be like a country with straight boundaries and without any colonies.

As we can see below, the decision boundaries and data points corresponding to class 1 and class 2 are clearly separable.



As we can see below, the decision boundaries and data points corresponding to class 5 and class 8 are NOT clearly separable.



Some classes are harder to separate because they contain points which are almost identical. In terms of images, we could say we have two different images which are very similar and this translates to a harder classification, making the separation of the two classes more difficult.