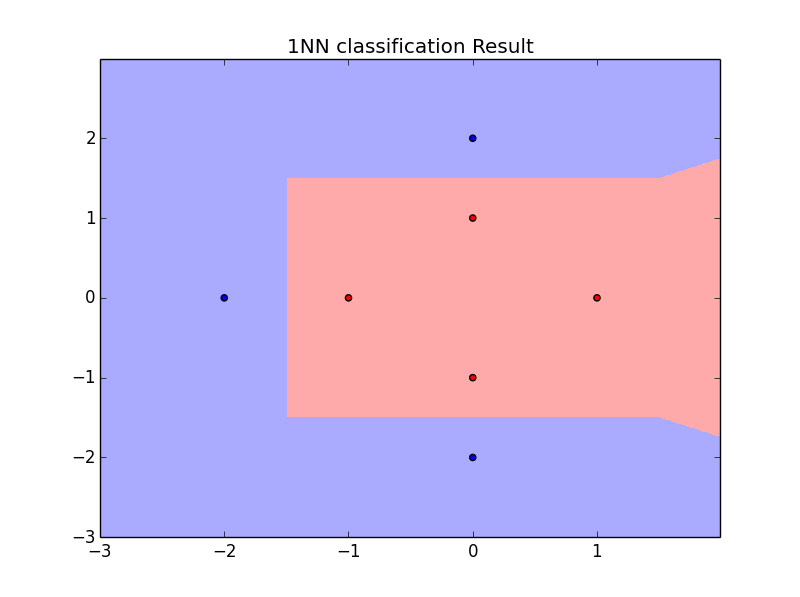
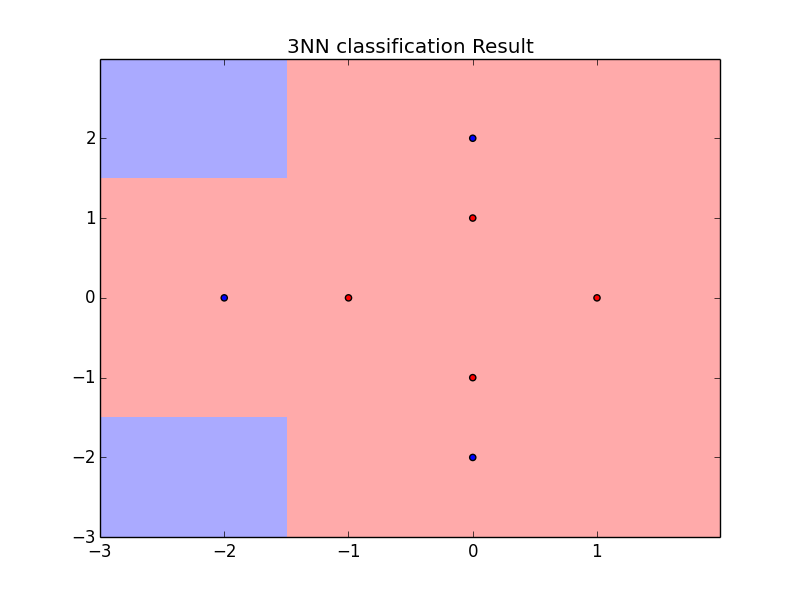
**Problem 6.1**

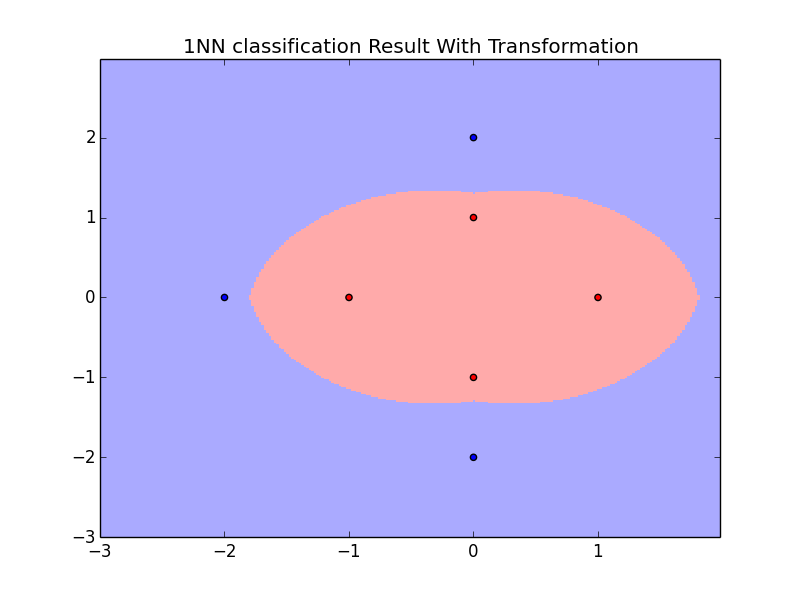
1. The 1-NN and 3-NN classification results for original training data without transformation is as below:

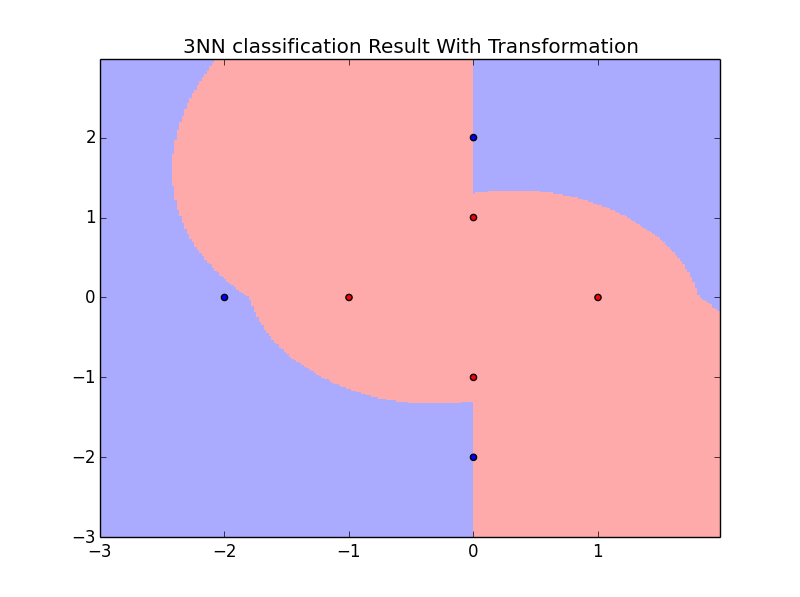




As there are only 7 points where 4 of them are red, when we use 3-NN for classification, it is very likely to misclassify.

1. The 1-NN and 3-NN classification results with transformed input are as below:

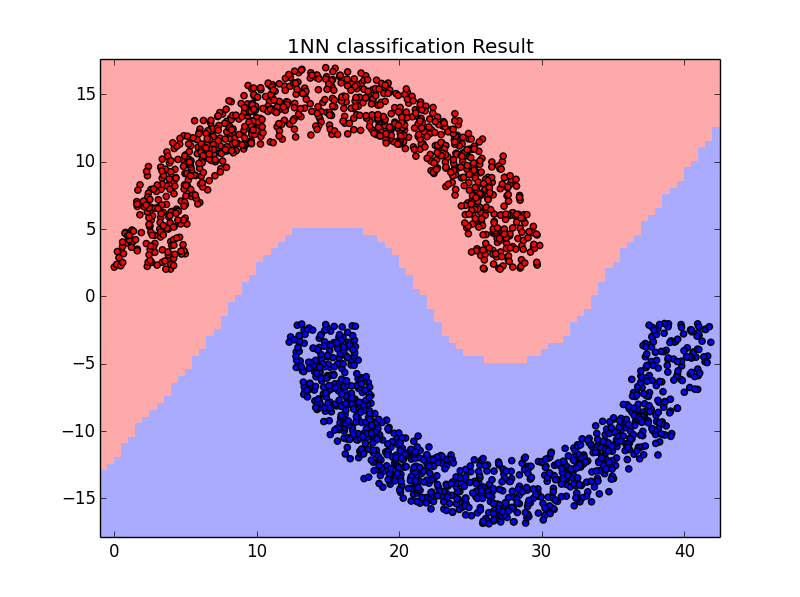


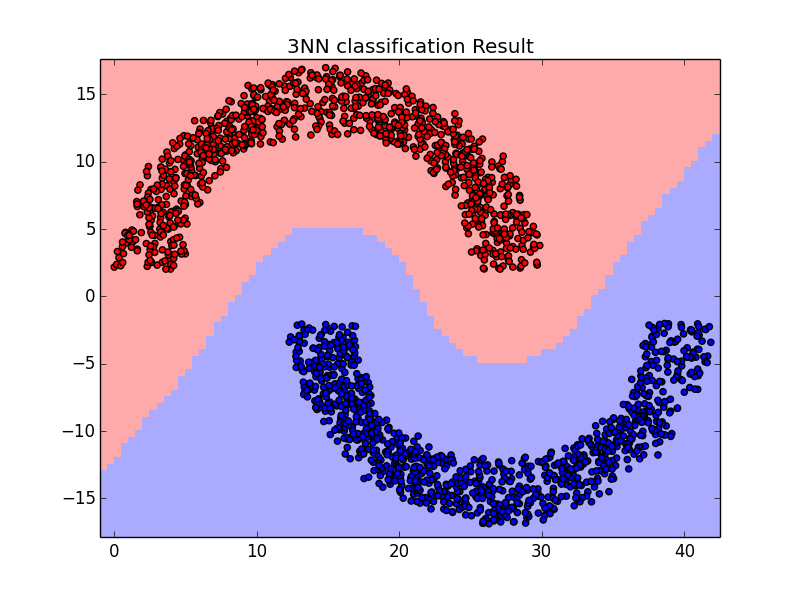


With transformed features, the same models have significant different explanatory power.

**Problem 6.4**

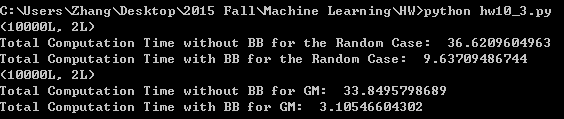
The 1-NN and 3-NN classification results are as below:

****

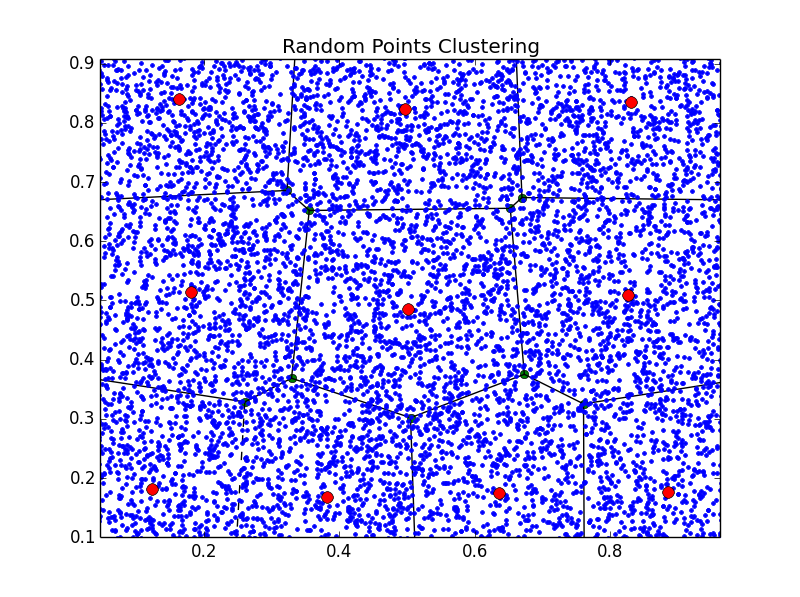
****

The two classification results looks quite similar as there are enough well separated training data points so when K is not large the KNN results are similar to each other.

**Problem 6.16**

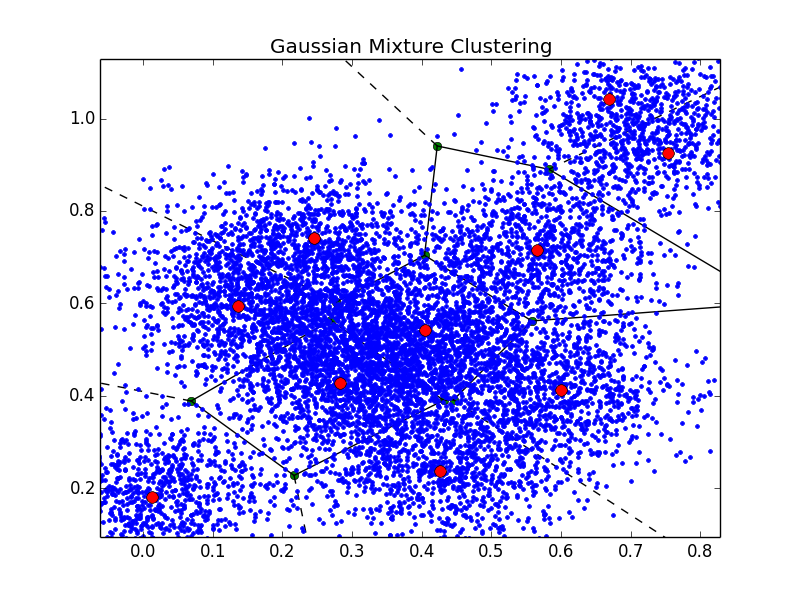


1. The clustering result is as below



Without branch and bound, the algorithm takes 36.6 seconds to converge. With branch and bound, it only takes 9.6 seconds, much faster than the brute force solution.

1. The clustering result is as below



Without branch and bound, the algorithm takes 33.8 seconds to converge. With branch and bound, it only takes 3.1 seconds. Much faster using BB, and also significant faster than partitioning with BB in part A.

1. In both cases, the branch and bound approach reduces the computational effort for convergence. Compare the two examples, the first example does not have obvious clustering effects (all the points are randomly drew from the same process). The second example has obvious different modes (10 Gaussian mixtures). With the original brute force approach, the computational time for the two examples are close, as the sample sizes are the same so that they would require same computational budget. But with branch and bound approach, the points in example B are more condensed in clusters thus require less computational effort for convergence.
2. Yes. The more test points I want to evaluate, the more I would like to use branch and bound to save the computational effort of calculating the distance. The computation complexity of the original brute force approach is a higher order of the sample size N while branch bound approach reduce the complexity to a lower order. Thus, larger value of N encourages the use of branch and bound.