Problem 3

as the dimensionality of the problem grows, the higher-dimensional space is less densely occupied by the training data, and we need to search a large volume of space to find neighbors of the test point. The pair-wise distance between points grows as we add additional dimensions.

And in that case, the neighbors may be so far away that they don't actually have much in common with the test point.

In general, the length of the smallest hyper-cube that contains all k nearest neighbors of a test point is:

$$\frac{(k/n)^1}{d}$$

for N samples with dimensionality d.

From the expression above, we can see that as the number of dimensions increases linearly, the number of training samples must increase exponentially to counter the "curse".