Robert Crimi

1) What is the role of the number of training points to accuracy?
   As the number of training points increases, so does the accuracy of classification. This is due to the fact that given a test example, with more training points, KNN is more likely to find the correct classification. More training data will give a better picture of the "real world" and thus given examples will be categorized into their "real world" classifications.

2) What is the role of k to accuracy?
   K represents the number of closest neighbors to find. On one extreme, if k is the size of the training set, any example will be classified as the majority of the training set. This is because the K closest neighbors would be the entire set. Thus, when taking the majority of the closest neighbors, it would result in the majority of the set. On the other extreme, very small values of k would result in inaccurate results. For example, if k=1 the single closest neighbor would be the classification of the example. This neighbor may not be the best representation of the example. This may be caused by a training set that is not representative of the "real world." In practice, it is best to experiment with different values of k to see what results in the most accurate classifications.

3) What numbers get confused with each other most easily?
   4 and 9 get confused the most amount of times. This is due to the shape of the numbers as they are visually similar. 7 and 9 were also numbers that got confused with one another easily. If the training set could be plotted, points corresponding to 4, 7, and 9 would be relatively close to one another. Depending on the example point location and value of k, these numbers will be confused with each other easily.