CMSC726- Homework#1 Solution

1. a. Bad inputs: Noisy data, Weak features.
   b. Overfitting
   c. Underfitting

2. Assume that I want to predict an image is a face image or not. My algorithm is $K$-NN:
simply compare a test image with all the training images pixel-wise and look at $K$ nearest
neighbors of it and if the majority of them were face then label the test image as a face.
   a. If images had low quality in training you may find some non-face images similar to
      face images
   b. If $K$ value chosen very small then you may only be able predict the faces that are
      actually identical to the training faces and not other than that.
       b. If $K$ value chosen very large your prediction will be independent to the test data
          because for any test image you may look at all the training set and pick the majority of the
          labels.

3. It can be a good idea in the sense of saving storage under the assumption that we have a
   good non-linear model. But it will not be a good idea generally because it may code some
   categories close to each other in hamming space and some others far in which they may not
   be really close or far.

4. $\| a - b \| = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \ldots + (a_D - b_D)^2}$ each terms $(a_i - b_i)^2$ is a random
   variable with this expectation: $\int_0^1 \int_0^1 (a - b)^2 dadb = \frac{1}{6}$ So the overall expectation of $\| a - b \|$ will be $\sqrt{\frac{D}{6}}$

5. The convergence theorem assume that data points are linearly separable and it tells us after
   a finite number of iterations we achieve the best best classifiers that separate all the positive
   and negative data which means it has zero number of errors.