# Machine Learning Lab4

Naïve Bayesian Classifier

1. Consider each dimension of the digits is conditionally independent given the digit label, and assume a uniform prior, please implement a naïve Bayesian classifier.

    For estimating the likelihood term, we can assume that the values at each dimension follow a Gaussian distribution
    http://en.wikipedia.org/wiki/Normal_distribution
    and the mean and standard deviation of the Gaussian function can be easily estimated in Matlab by using mean and std function.

2. If the naïve Bayesian classifier can't perform well, please think by yourself the reasons that cause its failure. If you think its failure is caused by the digit's high dimension, then Principle Component Analysis (PCA) might help. (You will learn PCA in Topic 6.) Try to use PCA to project the digits to low dimensions, and perform Naïve Bayesian Classifier in that low dimensional space. You can also try to visualize the PCA projections as 'ghost digits' shown below.
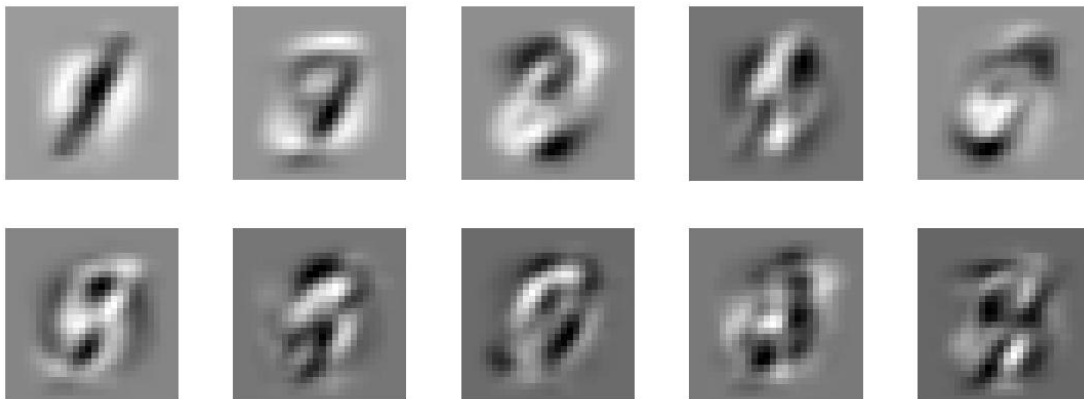
Instance based learning: K-NN

1. Try to implement the K-NN algorithm. Simply set K to arbitrary number, and see what accuracy you could get on the testing set.
2. Use cross validation technique to automatically select an optimal K.
3. You can also try to perform K-NN in the PCA reduced low dimensions.