Problem Statement 1

The images are 28 x 28 pixels in gray scale. The categories are 0, 1, ... 9. We are going to concatenate the image rows into a 784-dimensional (784 = 28 x 28) vector and treat this as your feature, and assume the feature vectors in each category in the training data have Gaussian distribution.

Problem Statement 2

We will be classifying the images in the testing data set using 0-1 loss function and Bayesian Decision Rule and report the performance. Lets assume that the class-conditional 784-dimensional image vector has Gaussian distribution. Additionally, we will explore why it doesn't perform as well as many other methods?

Problem Statement 3

We will train a multi-class logistic regression model.