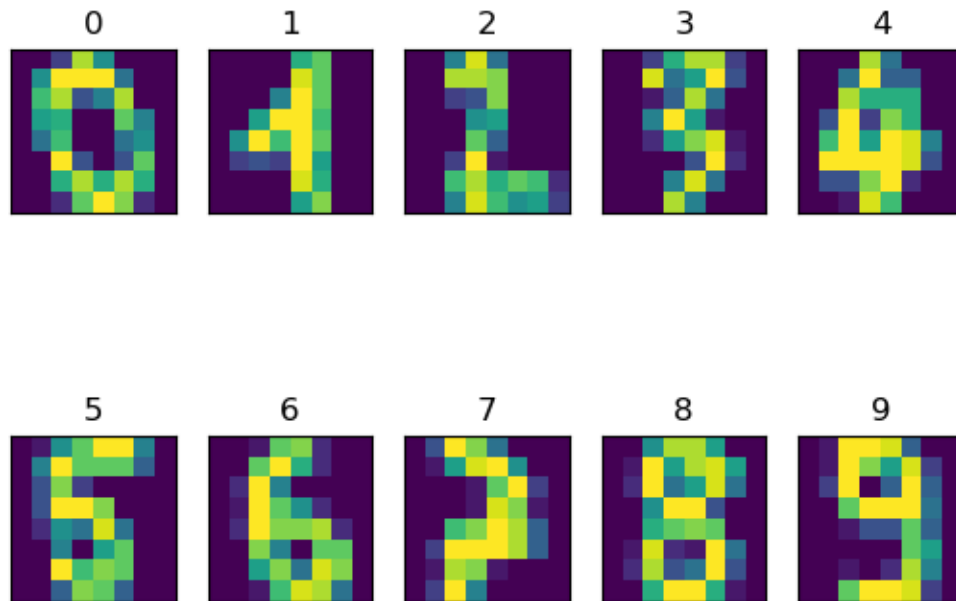


Report on code and paper by hamed khashehchi

In this paper (Mini Batch Normalization Mutual Information) Author and his colleagues (Thejas et. al.) proposed a new hybrid method for feature selection. Their hybrid model consists of two main parts. First, they rank the features by the Normalized mutual information of each feature on classification with batch-k-means (because of the speed of this approach) the problem at hand. Second, they choose a subset of features in which they get best results for the training set and then they used it on the test set.

They used two major approach to create the mentioned subsets. First, in this approach they used Inclusion technique which they added a feature until it gets a better result from higher ranking features, if on any case it made a result worse they will delete the feature from the subset. Second, Using Exclusion in which they started from the lowest ranking feature and delete each one of them as long as it didn't impair the results.

After I wrote the code I execute it for two results desired and mentioned at the question. My data was handwritten recognition at UCI:



*Figure 1- Sample of data at hand*

The first inquiry in the question were about each feature quality of data which were each pixel with a max normalization preprocess.

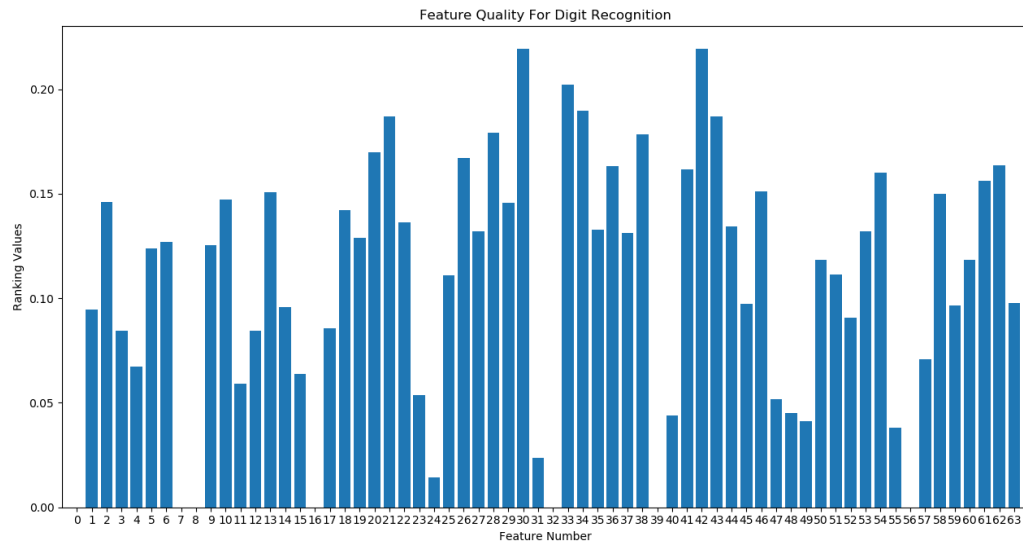


Figure 2- feature quality of each feature

The second inquiry at hand were about number of features in classification and its effect on accuracy.

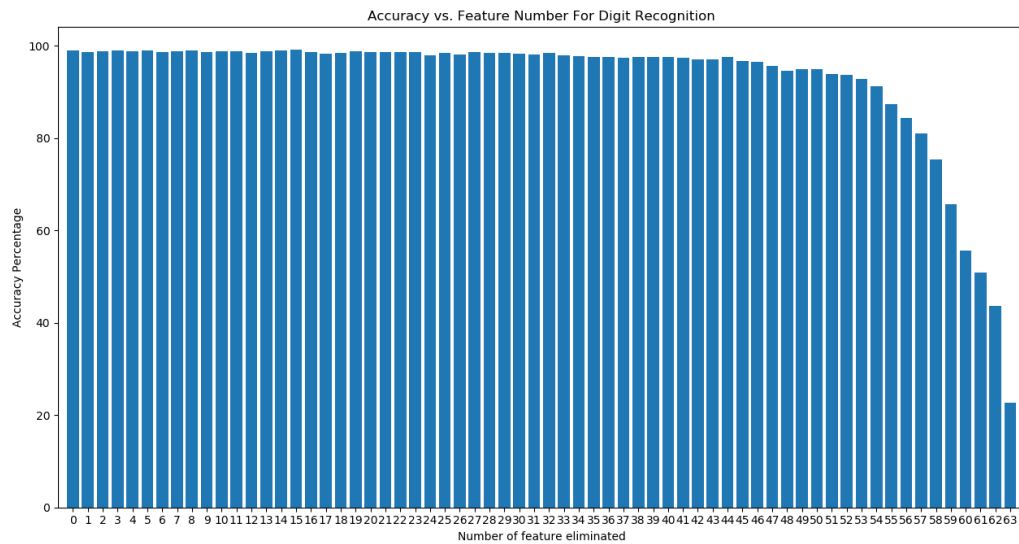


Figure 3-Comparing accuracy against number of features used in classification