

## Assignment 2

Use the ReducedMNIST which is a reduced version of the MNIST data set.

- **ReducedMNIST training:** 1000 examples for each digit.
- **ReducedMNIST test:** 200 examples for each digit.

1. Use the ReducedMNIST data to generate these features for each of the images of the training and testing sets:
  - a. DCT features (180 dimensions)
  - b. PCA (use a number of dimensions so that the total variance is at least 95% of the total variance when using all the 784 dimensions).
  - c. A feature of your creation
2. Then train these classifiers using the training set of the MNIST data for each of the above features:
  - a. K-means for each class. Try 1, 4, 16 clusters for each class.
  - b. GMM for each class. Try 1, 4, 16 GMM for each class.
  - c. SVM. You may try the linear and the nonlinear kernels (like RBF). In this case state clearly, what kernel have you used.

Then use the resulting models to classify the test set. Compare among the different features and the different classifiers:

1. You must add a final table to summarize all your results (accuracy and processing time) in a comparative way, as the table shown below.
2. Only for the best result of each classifier put it in a confusion matrix among the 10 digits.
3. Add your final conclusions.

		Features					
		DCT		PCA		created features	
Classifier		Accuracy	Processing Time	Accuracy	Processing Time	Accuracy	Processing Time
K-means Clustering	1						
	4						
	16						
GMM	1						
	4						
	16						
SVM	Linear						
	nonlinear*						
*mention which kernel							