Machine Learning for Applications in Computer Vision: Week 3

Liu, Nan
Wiedemann, Andreas

GitHub: `https://github.com/DJLoco/TUMMachineLearning.git`

Exercise 1: Boosting

The objective is to use the Boosting framework in MATLAB to train a AdaBoost classifier with the MNIST hand-written digits dataset. Here is our work:

- Load the MNIST data into MATLAB program, unpack the dataset and convert it into arrays:
  - Training images: 60000*784 uint8
  - Training labels: 60000*1 uint8
  - Testing images: 10000*784 uint8
  - Testing labels: 10000*1 uint8
- Train the default AdaBoost on the first 1000 samples using fitensemble:
  ClassTreeEns = fitensemble(TrainImages(1:1000,:),TrainLabels(1:1000),'AdaBoostM2',1000,'Tr
- Test against the test set using misclassification error:
  L = loss(ClassTreeEns,TestImages,TestLabels);

  The classification error is 0.3765.
- Plot the training error for 100, 200,… until 1000 the find the optimal number of learning rounds. According the figure, when learning rounds is 800, the resubstitution loss is the least, which is 0.3.
- Run the training on the entire dataset. The classification error becomes 0.3300.
- Train the LPBoost on the first 1000 samples:
  ClassTreeEns = fitensemble(TrainImages(1:1000,:),TrainLabels(1:1000),'LPBoost',1000,'Tree');

  The classification error is high, which is 0.6478.

Exercise 2: Gaussian Process Classification

The objective is to get familiar with the Gaussian Process Classification tool and compare with the Boosting classification method.

Download, install and run the demoClassification.m. Play around with the parameters. (Change the kernel which is to change mean and covariance function, the inference method, the likelihood function and evaluate on a new test set) I only change the inference method, for the others, which will be quite time-consuming because you have to take care of parameters of all the functions.

The default parameters are: meanfunc = @meanConst; covfunc = @covSEard; likfunc = @likErf; covfuncF = @covFITC,covfunc,u; inffunc = @infFITC$_EP.Theresul$ $Functionevaluation38; Value4.413667e{+}011.inferencemethod| > inffunc =$ @$infFITC_LaplaceTheresultis : Functionevaluation9; Value4.304855e{+}01$.

3) Evaluate the classifier with respect to uncertainty estimation. I couldn't continue from here because I don't know how to tell if the classifier is correct or not. I have asked several people I know from our class, they haven't gone that far. About the histogram, MATLAB have the function to compute it. 4) Use Boosting Classifier and compare the histogram plots. This is not difficult because the Boosting classifier is easy to use, which you can find from my code! Thank you again, and good luck!

Reference