NTUST, CSIE
Machine Learning (CS5087701), Fall 2017

Homework 3 (14pts)

Due date: Dec. 19

Question 3.1. [4pts] In the ANN lecture, we discuss learning an identity function by training eight data (slides no. 50 to 56). Discuss what will happen if we choose the number of hidden nodes to be fewer than 3 or more than 3. Writing a code or using some ANN package to find the answer is optional. No matter what, you need to explain your answer.

Question 3.2. [10pts] Analyze the following two datasets:

The MNIST Database of Handwritten Digits
http://yann.lecun.com/exdb/mnist/, and
ECML/PKDD 15: Taxi Trip Time Prediction (II)
https://www.kaggle.com/c/pkdd-15-taxi-trip-time-prediction-ii

You are recommended to use logistic regression, LDA (Linear Discriminant Analysis) or ANN with a diverse set of learning algorithms to analyze the datasets. That is, you should discuss at least two learning algorithms to find out the performance difference if there is any. After your analysis, you should write a short report and the report should be around three pages with a discussion section. The discussion part should be at least one full page long.

In your report, you should include the following items:

(a) List all the parameters for the models that you used.
(b) The prediction accuracy with cross-validation and possible different data partitions.
(c) What kind of feature set you should use for the datasets?
(d) Explain how many hidden layers are good to apply to the datasets and how many nodes in each of the layers should be appropriate based on your study. It is also possible if you choose only a single perceptron to learn the patterns hidden in the data. In this case, you should emphasize more about the different result from different learning algorithms.
(e) Give the reasons why the result is good (or bad) for different experimental settings (the number of learning iterations, etc.).