Discussion L4.1 (Practical aspects of problem solving with MLPs)

1) How would you represent text data (e.g., letters, n-grams, words) in ANN-based text prediction approach?

2) How do you deal with normalization or any data transformation with respect to training and (unseen) test data?

3) You develop a triage system that supports early prioritisation of patients at an emergency unit. There are five classes of the severity (1 – hardly any problem…5 – top priority) and the ratio of available samples per class is highly skewed, i.e. class 1: 35%, class 2: 25%, class 3: 15%, class 4: 20%, class 5: 5%. What are major concerns? How would you approach them (focus on key aspects relevant in the application context)?

4) You are supposed to set up a neural network based system for predicting political preferences helping make predictions about the results of elections based on web based questionnaires (without explicitly asking who people are going to vote for, just their views on burning societal etc. issues). What could be a problem with a data collection process?

Discussion L4.2 (Loss function and optimisation)

1) Give some examples where loss function does not directly capture all meaningful aspects of the evaluation of the network's functionality in a specific application. How could you minimise this discrepancy?

2) You are supposed to train an MLP for predicting lifetime of an expensive product. The predictor is supposed to suggest when the customer should stop using the product to avoid failure resulting in critical implications. What cost function would you choose for training the MLP predictor?

3) What implications for the network's predictions would have the choice of the error function, L2-norm, log(L2-norm) or exp(L2-norm), for training?

4) What is the major challenge with deploying Newton's method in MLP training? How can this be mitigated?

5) In what way are popular backprop based algorithms adaptive, e.g. Adam?

6) When and why would you recommend using ROC curves for evaluation?

7) Why is error/loss function different sometimes from problem dependent performance measure?

Discussion L4.3 (Performance evaluation)

1. You receive brain scans (neuroimaging) from patients suspected to suffer from Alzheimer's disease and collected by different hospitals. You want to investigate the feasibility of a general ANN-based diagnosis support system. Please describe your approach, list what you plan to do, and how/what you evaluate.

2. How would you compare two different neural network models in a weather prediction problem based on data from 10 different regions in Europe?

   - case A: choose the better architecture that could be locally tuned to guarantee best region specific performance
   - case B: choose the architecture that generalises best across different regions
   - case C: choose the right architecture and train the model for weather prediction on yet unseen data
   - case D: compare a single best generalising model with a committee of networks