Coursework commentary 2017–2018

CO3311 Neural networks

Coursework assignment 1

General remarks

On the whole coursework assignment 1 was well answered, with a majority of students gaining very good or excellent marks. However, it was disappointing to see a number of students choose not to answer both questions. Leaving out a question is a guarantee of losing a significant number of marks. The average mark for each section was around 60%.

Comments on specific questions

Question 1

This first question was designed to give students some exposure to a wider family of neural networks than can be covered in a course of the length of CO3311. It is based on the Asimov institute's website's poster called 'The Neural Network Zoo' available at http://www.asimovinstitute.org/neural-network-zoo/ . Students were asked to summarise the contents of that poster.

Good answers used knowledge gained from the CO3311 subject guide, and its associated readings, in formulating their answers.

Students were aided with a set of eight headings (a-h) under which to structure their answers. It was disappointing to see many answers ignore this aid. Clearly, sections which are omitted cannot be awarded marks, so advice such as this is best heeded. Examiners did, however, look for appropriate material wherever in the answer it might be placed.

Good answers included an introduction (part a) which gave some motivation for the 'neural network zoo'. Students were asked to comment on the graphical notations used in the poster in section (b). Clearly it would be appropriate to comment on its features and to contrast it with the other notations used in CO3311. The next section (c) was to be an account of the main features distinguishing networks. It is inappropriate to list common features of units here as this was to be the contents of the fourth section (d). Good answers talked about gross architecture, recurrence, layers, and so on. The focus should be on the networks and not on the units themselves.

Section (d) on unit types and their common features, inputs, weights, net, activation, and so on. should have been an easy section to write for those students who read both the poster and the questions carefully.

Similarly, the main differences between units, how they are connected, the formulae for their activations and the learning schemes was to be the subject of the fifth section (e).

This section was then to be followed by a summary (f) of the major types of network – it was open to students just how this was to be organised but could usefully have been given in a hierarchical manner – perhaps with a diagram.

The final two sections, conclusions (g) and references (h) enabled students to summarise their thoughts and to give due credit to any other sources that they used.

Amongst the most common weaknesses of answers were:

- not including material for ALL sections – even if headings are missing the material was looked for;
- lack of overall coherence of the answers;
- lack of citations and references where appropriate.

Question 2

Question 2 required students to implement a network. Real neural network applications tend to be very demanding in both computer runtime and memory. This coursework assignment presented students with the opportunity to see how a change in the desired output set of a network might impact on the weights. The idea is to train a network on one set of desired outputs before training the resulting network on another – a sort of evolution from one output set to another.

Two types of behaviour might be expected:

Firstly the 'retraining' time required might depend on how close the two output sets are, in some measure of 'close'. However, it is also possible that two output sets that seem to be close can have very different sets of weights.

This coursework assignment required students to try a few simple examples and to report on their findings. These examples are of one-unit Perceptron networks which learn simple three input logic functions.

Questions such as this are designed to give students simple examples and some experience implementing simple networks. Another major benefit is in giving students the opportunity to write up the results of their experiments. It is important to practice answering questions in the coursework assignment and therefore be better prepared for the examinations.

Good students were able to choose weights for the initial function (S) by inspection as hinted in the question. Less experienced students used the learning algorithm to learn both S and T from S. Being able to design a simple network is a useful skill – especially when answering some examination questions. Changing the weights to integers was often omitted and this might indicate a student's lack of understanding of how this might easily be achieved. Multiplying all weights by a positive real number does not change the function being modelled by a unit of this type.

The question also hints at linearly separable versus non-linearly separable training sets. Good answers discussed this issue in depth. This important distinction seems to have been lost on some students.

In writing up results for the CO3311 coursework assignments, students should heed the advice and requirements given in the questions – we are particularly interested in the detail required for others to duplicate the work being described.

Testing the behaviour of a trained network is an important step in validating what one has done. It was disappointing to see that a number of students failed to include such testing in their reports.

It was interesting to see the ingenuity of students when it came to visually presenting the way that weights change from learnt function S to function T. Despite this being a requirement of the question, some students seemed to feel it was not necessary to show the changes graphically and just presented

the 'before' and 'after' weights. Many did choose to show how the weights morphed from those appropriate to S to those appropriate to T.

Few students took up the suggestion of looking systematically at how training time depends upon the pairs (S, T).

Although there were many very good write-ups, there were a disappointing number that just presented the barest results with little or no analysis. Thinking about what experiment results mean, such as those obtained for this question, helps to develop a deeper understanding of the ways that units are able to mimic logic gates. It is a pity not to take this opportunity to learn.

Reports were expected to have six sections:

- An introduction, which set the scene. This was neglected in many answers.
- A description of the program or spreadsheet and how it was developed and tested. Again, testing seems to be missing from many students' development methodologies.
- Results with the given set of S and T showing how the weights morph from one set of values to another during training. This seemed to be the focus of most answers at the expense of the other sections.
- An analysis of the results obtained including any surprising or interesting observations that were made. Good answers gave much here but many answers lacked any analysis.
- A conclusion, beyond just the bare results. Good answers summed up the work done in this section.

References should also be included, as it is unlikely that students did not seek help from sources other than the guides and it would not be very good practice to omit such research.