Examiners' commentary 2018–2019

CO3311 Neural networks – Zone B

General remarks

As usual, the examination consisted of six questions. This aims to allow candidates to display their knowledge of a broad set of topics from the course. A balance between text and calculations was sought so that candidates had varying opportunities to score highly.

Candidates' familiarity with the following was expected and tested:

1. material from the subject guides
2. directed reading from the guides
3. coursework assignments.

Some scripts were very difficult to read. Candidates whose handwriting is not clear are well advised to consider printing (that is, using small capital letters rather than cursive handwriting) and to leave a blank line between each line of their answer scripts.

Questions 1, 3 and 4 were the most popular and gained good marks on average. Question 5 was also answered by most candidates, who on average achieved slightly less than those answering Questions 1, 3 and 4. Few attempted Question 6 but those who did scored reasonably well. On average, marks gained answering Question 2 were disappointing.

Comments on specific questions

Question 1

Question 1 focused mainly on the Perceptron.

Perceptrons and their learning algorithm were developed as a model of biological (natural) neurons. Parts (a) and (b) of Question 1 required candidates to list their main features and to explain the main differences in their operations. Poor answers tended to focus on the similarities while ignoring the considerable differences. For example, the spiking and seemingly haphazard nature of natural neurons contrasted with the more easily understood net and activation functions of artificial ones. Good answers often included diagrams of both types, and although not asked for, these can nevertheless save many words of description.

Part (c) was bookwork, requiring an expression for the activation of a threshold unit with given inputs and bias. Most candidates had little difficulty obtaining the full 2 marks for this.

The subject guides give three main reasons or motivations for studying artificial neural networks. Part (d) required each of these to be described. Poor answers simply listed these three reasons. Better answers showed insight into the mindset of investigators and the value of such studies to humanity.

Question 2

This question aimed to test candidates' ability to design perceptron networks that distinguish between the inside and outside of a simple geometric shape, an approximation to a circle. It did this by posing a sequence of sub-problems leading up to this.

Part (a) required a simple sketch of a given Perceptron. Candidates who used the notation introduced in the subject guides tended to fare better on this than those using more standard notation, because the guide's notation, when properly used, reminded candidates of the necessary parameters needed to fully describe a unit.

Next, in part (b), the expressions for net and activation were asked for. This was bookwork, which most candidates should have been able to answer without difficulty.

Part (c) required the design of a network that outputs a 1 when presented with points outside of a given triangle. Except for some issues with signs of weights, this proved straightforward in the majority of cases. Candidates are wise to show their working for such problems, because credit can be given even if the final result has errors. If no working is shown, arithmetic errors and small slips can have greater effects on the marks.

The final part, (d), reminded candidates that it is possible to approximate a circle by using a number of straight lines. Using this insight, candidates were to explain (using a diagram) how to produce a Perceptron network that gives an output of 1 for points inside the circle and 0 for those outside it. Some good answers were produced but on the whole answers tended to be muddled.

Question 3

Attention was turned back to Backpropagation. For part (a), a diagram of a specified network was required. While this should have been straightforward to answer, a surprising number of candidates failed to obtain full marks. Common errors were the omission of details. Had candidates used the notation introduced in the subject guides, they may have reduced these omissions.

The learning algorithm for the network described previously was required for part (b). Good answers gave the algorithm and included both the formulae and the meaning of the symbols in the formulae. Both forward and backward passes needed to be mentioned for full marks.

Given the parameters of a single Backpropagation unit, part (c) required candidates to calculate the weights after training with some simple examples. When selecting examples, examiners often use easy weights with few decimal places in order to simplify the working, and this was the case here. A common error was to forget to give the weight updates for the last example.

For part (d) the effectiveness, or otherwise, at learning the examples was required. It was disappointing to see how many candidates failed to notice that, as the target output for some examples was out of range for sigmoid activation, the network could not succeed no matter how many iterations through the training set were made.

Question 4

Question 4 tested candidates' knowledge of the details of Kohonen-Grossberg networks. Understanding that a main application is to find sets of classes in data sets is all that is required for part (a). Good answers stated this with some typical examples. Poor answers were too vague.

Normalisation, an important aspect of Kohonen training, has a few technical issues, which were the topic of part (b).It was disappointing to read that some answers confused accounts of when normalisation is or is not appropriate.

Part (c) was bookwork requiring a detailed account of the algorithm for training the Kohonen layer. Many answers were almost word for word copies of the subject guide – a good sign in this case, as sometimes vital steps can be lost in paraphrasing.

Parts (d) and (e) tested knowledge of how classes and their number are initially chosen. Candidates should have some understanding of these from the coursework exercises. Good answers stressed the need for experimentation and replication as well as strategies of growing or pruning networks.

Question 5

The main types of recurrent neural network from the course are Hopfield networks, and these were the main topic of this question.

It started by asking for examples of how these differ from non-recurrent networks. A few candidates confused the concepts of backpropagation and recurrence, though most had no such difficulties.

The detailed algorithm used to determine the state transition table of a Hopfield network was required in answer to part (b). Although answers to later parts of the question showed that most understood the algorithm, many found expressing their understanding in simple English quite hard.

Part (c) sought examples of terms introduced in the work on Hopfield networks. Terms such as state, predecessor, successor, source sink, etc. were all acceptable.

The final part of the question required the production of state transition tables and diagrams to be produced for a set of weights. On the whole, this part was well answered. Errors in calculation were the most common problems encountered, though some candidates did not remember the order of columns needed for an efficient calculation.

Question 6

The final question on the paper tested candidates' knowledge of developments in Artificial Neural Networks over the decade since the subject guide was written. Three sets of changes were called for. **Developments in architecture** were reasonably well answered, although some forgot that architecture includes algorithms and activation functions, as well as more obviously the ways that units are connected.

Second, **technological advances** have made previously impractical sizes of networks, dataset sizes and numbers of iterations possible. Good answers mentioned all of these as well as giving, for example, the use of GPUs as one of these drivers.

The **existence of large datasets** has also, coupled with the technological advances, made applications that seemed impossible in 2009 a possibility. **Deep learning** has made possible many new applications of ANNs.