Examiners' commentary 2018–2019

CO3311 Neural networks – Zone A

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 2 and 6 were the least popular whereas Question 1 was the second most popular. These three questions achieved on average fewer marks than the other three questions. The most answered question, Question 4 was the second highest scoring one, although Questions 3 and 5 were not far behind in popularity and had similar average marks.

Comments on specific questions

Question 1

This question concerned Perceptrons and their relationship with biological neurons.

Parts (a) and (b) required a description of the action of both types of neurons, together with an account of how they differ in action. A sketch of both was required but some poorer answers lacked this. In answering questions like this it is important to highlight any contrasts and not to leave it to the reader to find differences between two descriptions.

Perceptrons have a number of limitations; describing these and the strategies used to overcome some of them was the subject of parts (c) and (d). Many answers explained linear separability; and good answers gave the XOR or a similar function and a network that models it as examples. Fewer candidates remembered that binary outputs are a limitation that is harder to overcome.

The existence of a learning algorithm used for training a single unit Perceptron is perhaps its greatest strength and the last part of the question simply required an account of this algorithm. On the whole, this part was well-answered with the good responses giving the algorithm, its equations and the meaning of all of the symbols and terms appearing in it.

Question 2

Question 2 was also about Perceptrons but this time focused on their relationship to lines (or higher dimensional subspaces) and how they can be combined to classify points within a polygon drawn on the x, y plane.

Part (a) required the relationship between an arbitrary line and the weights of a corresponding Perceptron to be spelt out. Poor answers resulted from not checking carefully the points actually on the line. Good answers expressed the correspondence between bias and other weights with the equation $y = mx - c$; and that multiplying the weights by any non-zero number does not change the line. Very good candidates also noted that for the origin to output a 1 requires the bias to be non-negative.

For part (b) this was applied to form a Perceptron network giving an output of 1 for the inside of a triangle. Poor answers failed to make a connection with part (a), whereas good answers showed how the result could be simply achieved using three lines and a Perceptron having an OR truth table. Poorer answers also tended to give explanations that were far from complete.

Part (c) took this idea further and only required a description, in strategic terms, of how the ideas of the previous parts could be generalised to arbitrary polygons. Good candidates had little difficulty expressing this in English whereas poorer ones fumbled with their explanations. Choosing a good form of notation is the key to answering this type of question.

Question 3

In contrast with previous questions, Question 3 turned its attention to Backpropagation networks.

Part (a) asked for a comparison (and contrast) of these type of network with Perceptron networks. Besides the obvious issues of analogue versus binary outputs, good answers talked about the universality of backpropagation networks with one hidden layer as well; as the uncertainty of their convergence.

The two important learning rues were the topic of parts (b) and (c), which sought a comparison of Hebb's rule with that of Widrow-Hoff. Though confusing the two names is no great issue, it is important that candidates understand that increasing correlations between units that are required to fire together is just one possible way of learning, while gradient descent of the error surface is another valid option. This was clearly stated in good answers. Poor answers confused these with the contrast between the Perceptron learning algorithm and that of Backpropagation.

Part (d) tested candidates' ability to work through a simple example of Backpropagation. Candidates were given a set of initial weights along with the desired outputs, assuming four examples of two binary inputs for a single Backpropagation unit.

The weights after training were required. This was a straightforward exercise for many. Common errors included:

i. not using the updated weights for the next input

ii. forgetting to update the weights after the fourth example

iii. sloppy rounding errors.

The final part, (e), seems to have puzzled many candidates. The target values for some members of the training set were out of the range of Backpropagation networks and so the set as a whole could not be learnt without modification of some sort. Too many candidates did not notice this.

Question 4

Moving on from Perceptrons and Backpropagation, parts (a) and (b) started by asking for a comparison of supervised and unsupervised learning. Candidates were mostly able to do this successfully as well as give suitable examples.

Part (c) asked for the main features of a Kohonen-Grossberg neural network. Good answers mentioned that the Grossberg conditioning layer is often omitted.

The algorithm for training the Kohonen layer is needed to answer part (d). Each step and each formula required a line or two of explanation. A line of mathematics on its own is not enough for good marks. Fortunately, the majority of candidates were able to produce good answers for this. However, a common error was in the normalisations needed by the algorithm.

Given the weights of a four-unit Kohonen-Grossberg network, candidates were required to calculate the results of training the network using a given training set in part (e).

Not all candidates were able to put the algorithms that they gave previously into practice. Common errors, besides slips in calculations, were the omission of normalisation and the lapse into using sigmoid rather than winner-takes-all activations.

Question 5

Part (a) asked for a labelled diagram of a Hopfield network 'Using the notation for units introduced in the guides'. Some candidates failed to follow this instruction, making their task more difficult and often losing marks by omitting vital details.

For part (b) the function and use of each component was required. Here, failure to heed the requirement for the notation used in the subject guides often led to missing components and so to losing marks.

Producing a state transition table and associated diagram, given a set of weights is laborious but should present little technical difficulty, if the layout suggested in the guide is strictly adhered to. Most candidates were able to do this in answer to parts (c) to (e) although some suffered from arithmetic errors and others confused the order of columns in the table. A common error in the production of the diagram was the omission of circles on states which are their own successors.

Question 6

Question 6 gave candidates the opportunity to write about their knowledge of applications of ANNs. The question encouraged candidates to give an account of a single application, its aims, objectives; and how attempts to develop such applications have changed over recent times.

For high marks, answers to such questions need to include technical details and not just the sort of information often produced for marketing or advertising purposes.

Excellent answers included details of architectures, success criteria, and where further progress is needed.