Examiners' commentaries 2009

2910311 Neural networks – Zone A

General remarks

This academic year, 2008–2009, the results of the candidates from zone A in the CIS311 exam on Neural networks were relatively good. I am inclined to think however, that the candidates did not work equally hard on all topics. For example, most candidates avoided the first question. This seems surprising since Question 1 was intended to be the easiest question. The exam questions were subdivided, as is traditional, into two different parts: theoretical and practical. The theoretical parts asked questions on the properties of feedforward and recurrent neural networks, and also questions on the corresponding training algorithms for these networks. The practical parts asked questions on the applicability of the network algorithms to regression and classification tasks.

Overall I found out that the external candidates who sat this paper were slightly better at answering the more theoretical questions, such as Questions 2, 3 and 4. Many of them showed also good knowledge on Question 6 which has larger practical subsections. Surprisingly most candidates avoided Question 1 as mentioned above, and Question 5 although this is as practical as Question 6. The performance of candidates in this subject is, in general, satisfactory. Most have demonstrated their ability to understand the subject material and apply this to work on tasks in the area of neural networks.

Question 1

This question is traditionally devoted to the general theoretical underpinnings of single neuron and multilayer neural networks. Subquestion 1a) asks candidates to explain why many layers in a network are sufficient to classify an n-dimensional point into a polygon, while the next subquestion 1b) asks whether we can design networks in which the neurons use different activation functions. Subquestion 1c) asked candidates to describe whether the value-descending strategy for Boltzmann training of multilayer networks can reach a solution with zero error using the notions of local and global minimum. Subquestion 1d) asked candidates to give the formula for accepting weight changes in the Boltzmann training algorithm for multilayer networks

The last subquestion 1e) asked candidates to train a single discrete neuron with a given alternative training algorithm. It should be noted that only a few candidates attempted Question 1, and they achieved only an average success on it, which for me was a little disappointing as this was supposed to be a relatively easy question.

Question 2

This question was devoted to networks of single discrete neurons. The first part subquestion 2a) requires candidates to describe the learning potential of a single neuron that uses a discrete activation function. More precisely, the question required candidates to first write the equation of the line modelled by this neuron, and then plot in two-dimensional space the line modelled by this neuron, and finally to identify in which half of the plane are the points corresponding to negative outputs. I was pleased to find that many candidates attempted this question, which is purely theoretical. This indicates that they have studied hard using the subject guide where the relevant material is well explained.

The second part, subquestion 2b), asked candidates to perform training of a single discrete neuron with two inputs using the given initial weights and one training example. It should be noted that this is not a difficult task, and most of the candidates who answered this question achieved very good marks.

Question 3

This question was also devoted to single layer networks, but it included single discrete neurons as well as sigmoidal neurons. As expected many candidates worked on this theoretical question and achieved quite good results. Subquestion 3a) asked when is the neural network training process unsupervised. Subquestion 3b) asked which of the following Boolean functions, AND, OR, XOR and NOT, cannot be modelled by a single neuron with two inputs. Surprisingly many candidates failed to answer correctly that only the XOR function cannot be modelled by a single neuron with two inputs.

Subquestion 3c) asked candidates to explain how the performance of a single neuron with two inputs and a discrete activation function will be affected when the weights and the threshold are multiplied by a positive constant. The answer is relatively easy, but many candidates failed to realise that if the weights and the threshold of such a neuron are multiplied by a positive constant the line (or the hyperplane) that the neuron produces will be the same, and its output remains the same for each pair of inputs. The last subquestion 3d) asked candidates to train a sigmoidal neuron with two inputs, given initial weights and two training examples. It should be noted that many candidates worked properly and produced correct results on this subquestion, which is quite satisfactory as a good knowledge of sigmoidal neurons is essential for understanding the more complex multilayer neural networks.

Question 4

This question required candidates to train a multilayer feedforward neural network with one hidden neuron, one output neuron, and two inputs. The task was to demonstrate knowledge and understanding of the basic backpropagation training algorithm for multilayer neural networks, which is a milestone question in the field of neural networks. The candidates were given only symbolic variables, and were asked to derive corresponding backpropagation rules for training the connections between the inputs and each hidden node, as well as training the connections between each hidden node and the output node. They had to develop the

equations for computing the hidden and output node outputs, the equations for computing the changes to weights on the connections toward the output node, and equations for computing the weight changes on the connections toward the hidden node. This was essentially a theoretical question that did not involve any numerical calculations, and was designed to test the analytical abilities of candidates to reason theoretically about training the multilayer neural networks that are most popular and widely used in practice. Most candidates attempted this question and achieved quite good results. I am therefore satisfied, given the answers of the candidates, that they have enough knowledge and understanding of the most essential part of the material.

Question 5

This question was devoted to Hopfield neural networks. The first subquestion 5a) asked candidates to discuss whether the training process in Hopfield networks can reach an energy well starting from any initial state. The answer is that the final state to which the network converges depends not only on the initial state but also on the order of firing of the neurons. The energy of the initial state will be greater than or equal to the energy of the well it reaches. The second subquestion 5b) asked candidates to write the formula for computing the energy level in Hopfield neural networks. In subquestion 5c) candidates were asked to train a Hopfield network with three neurons, and a given initial weight matrix with a given input pattern. Similar examples are given in the subject guide, but surprisingly only a few candidates attempted this question. It could be that the candidates found Hopfield networks, such as the one in this question, harder to understand than other networks. This explains why they were unable to properly update the weights using the Widrow-Hoff rule, or to retrain the network after that. Although their results were not excellent, the performance of candidates who attempted this question was acceptable overall.

Question 6

The final question was devoted to self-organizing neural networks of Kohonen type. The first part subquestions 6a) and 6b) asked candidates to give the role of Kohonen networks as unsupervised learning tools, and the alternative ways for initialisation of the weights in the Kohonen layer of counterpropagation networks. Many candidates answered these two subquestions correctly. Most candidates also wrote down the correct normalisation formula for Kohonen networks required in subquestion 6c). The next subquestion 6d) required candidates to demonstrate training of a simple self-organizing neural network with two neurons in the Kohonen layer having four inputs each. The two initial weight vectors and two input training vectors were given. Unfortunately, many candidates did not realize that the given initial weights and input vectors are directly given as pre-normalized, that is they do not need to be pre-normalized before using them for training. However, with this exception, candidates performed well on this question and I am satisfied with their results on this final question in the exam paper.