Examiners' commentaries 2009

2910311 Neural networks – Zone B

General remarks

At the end of this academic year 2008–2009 the candidates from zone B have achieved good results in the CSI311 exam on Neural networks. I am inclined to conclude that the candidates from the external programme have worked hard on the material from this course, and they succeeded in understanding the principles of neural network modeling.

The exam questions were subdivided, as is traditional, into two 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 performing classification and regression using neural network training algorithms.

This year the external candidates were more successful on the theoretical questions about single-layer and multilayer feedforward networks, such as Questions 3 and 4. Showing good results on Question 4 is important since multilayer feedforward networks are the most widely used kind of neural networks. Having a good theoretical knowledge of these networks means that candidates should also be able to deliver a practical application. The candidates were also successful in answering correctly the practical Question 6 devoted to unsupervised learning in counterpropagation networks. That is why, I think, the performance of candidates in this course is satisfactory and that the candidates have acquired the potential to work in this area.

Question 1

Traditionally Question 1 is devoted to the general theory of neural networks, including single layer and multilayer architectures. This year, it begins by asking how many layers are necessary in a network in order to classify a point into a concave polytope; and what exactly is meant by the notion of a concave polytope. The second question 1b) asked the candidates to explain briefly if there is some delay between the firing of the neurons in the first and second layers in multilayer networks. Subquestion 1c) asked what the implication is of using the value-descending strategy for Boltzmann training of multilayer neural networks, while subquestion 1d) asked candidates to give the formula for computing the error necessary to evaluate the performance of a multilayer network trained by the Boltzmann algorithm. The final subquestion 1e) asked candidates to train a single discrete neuron with a given alternative training algorithm. Surprisingly most candidates avoided this question, which is intended to be an easy question devoted to the most general ideas behind classical neural networks. As many candidates did not attempt this question, I am rather inclined to think that maybe they have concentrated their attention on specific network training algorithms, rather than studying the general principles in this area.

Question 2

This question was devoted to networks of single neurons. The first subquestion 2a) required candidates to explain how a single neuron with two inputs and given connection weights could be trained. More precisely, it asked candidates to give the equation of the line modelled by this neuron, to draw the line modelled by this neuron in the two-dimensional plane, and to show in which half of this plane the points corresponding to positive outputs are. Although this seems to be a purely theoretical question on the learning capacity of single neurons, many candidates attempted it and achieved good results. This means that they have studied hard the section in the subject guide, in which the geometry of single neurons is discussed.

Part 2b) was a practical question and required candidates to train a single discrete neuron with two inputs and a bias using given examples. Most candidates performed relatively well on this subquestion. However some made mistakes in the numerical calculations, which were supposed to be easy. In general I am satisfied with the results for this question as a whole, because most of the candidates who attempted it achieved relatively high marks.

Question 3

This question was devoted to single layer networks with sigmoidal activation functions. This is different from question 2 since it requires candidates to understand a more complex training algorithm, using a more complex training rule which is a milestone for understanding multilayer networks. Subquestion 3a) asked candidates to reason whether there will be a change in the performance of a single neuron with two inputs and a discrete activation function, if the weights and the threshold are multiplied by a negative constant. The answer was straightforward because if the weights and the threshold of such a neuron are multiplied by a negative constant the line that the neuron produces will be the same, so the output will remain the same.

The second subquestion 3b) asked candidates to explain briefly which of the Boolean functions, AND, OR, XOR and NOT, can be modelled by a single neuron with two inputs: the answer is that only the XOR function cannot be learned. The third subquestion 3c) required problem solving. Candidates were asked to apply the Widrow-Hoff rule for sigmoidal neurons to train a sigmoidal neuron with two inputs. Most of the candidates performed well on this question, and especially on 3c) which indicates that they can address practical tasks that require modelling by sigmoidal neurons. Many candidates demonstrated that they are able to apply their knowledge to practical problem solving using single layer nets.

Question 4

This question was devoted to the theory of multilayer feedforward neural networks. It asked candidates to provide the equations of the backpropagation training algorithm necessary for training a given multilayer neural network with one hidden neuron and one output neuron, both of which use sigmoidal activation functions, and two inputs.

This was a purely theoretical question that did not require any numerical calculations. It was designed to test the analytical abilities of candidates, and to see if they are able to reason theoretically about training multilayer neural networks - the most popular and widely used neural networks in practice. It was quite satisfactory to find that most candidates succeeded in deriving the 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 network node. This means that the candidates have acquired enough knowledge and understanding of multilayer feedforward neural networks, and they seem to be really confident in their knowledge. Many candidates achieved more than twenty marks on this question which is a very good outcome.

Question 5

This question was devoted to Hopfield neural networks. This topic is well explained in the subjecct guide; however many candidates avoided this question. Subquestion 5a) asked which is the most distinguishing characteristic of Hopfield neural networks with respect to the connectivity pattern, and most candidates answered correctly that the Hopfield network is a recurrent neural network with feedback connections.

In part 5b) candidates were asked to give the formula for updating the state in Hopfield neural networks, but many of them answered wrongly, giving the formula for training of the weights, rather than the state. In part 5c) candidates were asked to demonstrate training of a Hopfield network with three neurons, having been given the initial weight matrix. Although many candidates attempted this subquestion, they did not achieve satisfactory results in general, because they failed to properly apply the training algorithm. The reason for this could be that this training algorithm is quite different from the others studied in the subject, making it seem harder to understand.

Question 6

This question was devoted to training of the Kohonen layer of counterpropagation networks. The first subquestion 6a) required candidates to explain how the number of neurons in the Kohonen layer of counterpropagation networks impacts on their initialisation. The second subquestion 6b) required candidates to say how we train Kohonen networks: in supervised or unsupervised mode. Surprisingly candidates did not generally give a satisfactory answer to this question. A good answer is that Kohonen networks are trained in unsupervised mode because typically only training inputs are provided, without corresponding target outputs.

The final subquestion 5d) required candidates to train a Kohonen network with two neurons having four inputs each, using given initial weight vectors and one training input vector. After each input training vector there had to be computed the summation block and after that, there had to be calculated the corresponding weight updates. Candidates traditionally like questions on such self-organizing networks and most of them attempted this question. Unfortunately, this year many candidates did not realise that the given initial weights and input vectors are pre-normalized, that is they do not need to be normalized in advance. Overall however, candidates performed very well on this question and achieved good results.