A23406

No calculator permitted in this examination

UNIVERSITY OF BIRMINGHAM

School of Computer Science

Final Year – Degree of BSc with Honours

Artificial Intelligence and Computer Science
Computer Science

06 20416

Neural Computation

Summer Examinations 2010

Time allowed: 1 ½ hours

[Answer ALL Questions]

[Answer ALL questions]

1. (a) Draw a diagram of a *McCulloch-Pitts Neuron* and explain how it relates to the components of a biological neuron. [6%]

   (b) Write down the equation for the output of a *McCulloch-Pitts Neuron* with n inputs. For the case of two inputs, derive the equation for the decision boundary that it computes in the input space, and explain what it means. [6%]

   (c) In terms of such decision boundaries, describe with the aid of diagrams which logical functions can be computed by a single *McCulloch-Pitts Neuron* and which cannot. [5%]

   (d) Explain what the *Perceptron Learning Rule* is, and discuss its advantages and disadvantages over attempting to compute decision boundaries analytically. [8%]

2. (a) Suppose you have been given a set of training data and asked to use it to train a *Multi-Layer Perceptron (MLP)* neural network using the following cost function

$$E = -\sum_{p} \left[targ^{p} . \log\left(out^{p}\right) + \left(1 - targ^{p}\right) \log\left(1 - out^{p}\right) \right]$$

   What is this cost function called and what kind of problem would it be appropriate for? Describe the network architecture you would use, including your choice of activation functions. Outline (without going into mathematical details) how you could derive a simple gradient descent based learning algorithm for your network. [9%]

   (b) Why might you consider using a *line search* based learning algorithm instead? Describe how such an algorithm works in practice, and outline the main advantages and disadvantages this might have over simple gradient descent. [8%]

   (c) Suggest a change you might consider making to the given cost function to improve your network's generalization performance? Give an intuitive explanation of why you think it might work. Explain the effect of any associated parameter values and how you would optimize them. [8%]

3. (a) Describe the architecture of a standard *Radial Basis Function (RBF)* network and explain what is computed at each layer. [8%]

   (b) What is the most computationally efficient way to use a set of training data to determine appropriate weights/parameters for such a network that can be expected to produce good generalization performance? [9%]

   (c) Explain what is meant by *ensemble averaging* and how it might improve the generalization that is achieved by the RBF network. [8%]

4. (a) Explain what *dimensional reduction* means and how it is relevant to *data compression*. [5%]

   (b) Describe the architecture of a *Kohonen Network* and outline how it can be trained to carry out dimensional reduction of a set of data. [10%]

   (c) In this context, explain what *Voronoi Tessellation* means. [3%]

   (d) Explain what is meant by *Learning Vector Quantization (LVQ)*, and why it might result in more useful outputs for a Kohonen Network. [7%]

End of Paper