UNIVERSITY OF EDINBURGH

COLLEGE OF SCIENCE AND ENGINEERING

SCHOOL OF ENGINEERING

Convenor of the Board of Examiners: Professor Mike Davies

External Examiner: Professor John Soraghan

Wednesday 11th May 2011

2.30pm – 4.00pm
(1.5 hrs)

PGEE11020 ADVANCED CONCEPTS IN SIGNAL PROCESSING

This paper consists of TWO sections.

Candidates should attempt THREE questions, chosen as follows:

SECTION A: (ONE question). Answer the whole section.
SECTION B: Answer TWO out of THREE questions.

Students should assume reasonable values for any data not given in a question nor available on a datasheet, and should make any such assumptions clear on their script.

Students in any doubt as to the interpretation of the wording of a question, should make their own decision, and should state it clearly on their script.

Only a calculator from the list approved by the College of Science and Engineering may be used in this examination

The examination paper that you are now sitting is to be marked ANONYMOUSLY.

Please write your name in the space indicated at the top right hand corner on the front cover of the answer book. Also enter your examination number in the appropriate space on the front cover.

Write **ONLY** your examination number on any extra sheets, worksheets or graph paper used and firmly attach these to the answer book(s).

SECTION A

a) Draw a diagram of a simple perceptron. Compare this with a biological neuron, stating the biological equivalents of the important parts. **(5)**

b) Explain what is meant by the bias variance trade-off and its role in model learning. **(3)**

c) Describe how to minimize a cost function, $J(\mathbf{w})$, with respect to $\mathbf{w}$, using a fixed step size gradient descent. Show that the speed of convergence is controlled by the eigenvalues of the Hessian matrix, $\nabla^2 J(\mathbf{w})$. **(6)**

d) Consider a 3-state left-to-right Hidden Markov Model (HMM) with state transition matrix $\mathbf{A}$ and 3 visible symbols, v_1, v_2, and v_3 with an emission probability matrix $\mathbf{B}$, where $\mathbf{A}$ and $\mathbf{B}$ are given by:

$$\mathbf{A} = [a_{ij}] = \begin{pmatrix} 0.0 & 0.5 & 0.5 \\ 0.0 & 0.5 & 0.5 \\ 0.0 & 0.0 & 1.0 \end{pmatrix} \text{ and } \mathbf{B} = [b_{jk}] = \begin{pmatrix} 1.0 & 0.0 & 0.0 \\ 0.5 & 0.5 & 0.0 \\ 0.0 & 0.0 & 1.0 \end{pmatrix}.$$

Given an observed sequence $\{v(1), v(2), v(3)\} = v_1 v_2 v_3$, and assuming that the HMM is in the left-most state at time $t = 0$, calculate:

(i) $\mathbb{P}(\{v(1), v(2), v(3)\} = v_1 v_2 v_3)$,

(ii) the most likely state sequence $\hat{\omega}^3 = \{\omega(1), \omega(2), \omega(3)\}$ to generate the observation sequence,

(iii) $\mathbb{P}(\hat{\omega}^3 | \{v(1), v(2), v(3)\} = v_1 v_2 v_3)$. **(6)**

SECTION B

Question B1

Consider a scalar 2-category classification problem with Gaussian class conditional densities $p(x|\omega_i)$.

a) Show that the Bayes decision boundary (BDB) can be written in terms of the prior probabilities P_1, P_2, the means μ_1, μ_2 and the variances σ_1^2, σ_2^2, as follows:

$$\frac{P_1}{\sigma_1} \exp\left\{\frac{(x-\mu_1)^2}{2\sigma_1^2}\right\} = \frac{P_2}{\sigma_2} \exp\left\{\frac{(x-\mu_2)^2}{2\sigma_2^2}\right\}.$$

(3)

b) If $\sigma_1 = \sigma_2 = \sigma$ and $P_1 = P_2$ show that the BDB is given by $x = (\mu_1 + \mu_2)/2$. **(5)**

c) Show that when $\mu_1 = \mu_2$ and $\sigma_1 > \sigma_2$ then no BDB exists if:

$$\frac{P_1}{P_2} > \frac{\sigma_1}{\sigma_2}.$$

Explain what this means. **(7)**

d) Given class labelled training data $\{x_i^{(1)}\}, i = 1, \dots n$ in ω_1 derive expressions for the Maximum Likelihood estimates for μ_1 and σ_1^2. **(5)**

Question B2

a) Explain what is meant by the concept of unsupervised learning. **(2)**

b) Suppose that a sequence of scalar values, $x(1), \ldots, x(n)$, is drawn from a Mixture of Gaussians (MoG) model with k components, $c_1, \ldots, c_k$, where each component has a mean, μ_i, a variance, σ_i^2, and a prior probability, $w_i = P(c_i)$.

(i) Write down an expression for the log likelihood function for this model. **(3)**

(ii) Describe how to estimate the means, μ_i, and variances, σ_i^2, using the Expectation Maximization (EM) algorithm (assume that the prior probabilities are known). **(5)**

(iii) Indicate the similarities between the EM algorithm and the maximum likelihood estimates for a single Gaussian model. **(1)**

c) (i) What is Independent Component Analysis? Your answer should describe the model assumptions and the inherent ambiguities in the solution. **(3)**

(ii) Show that the log likelihood function for the unmixing matrix $\mathbf{W}$, given the data vectors $\mathbf{x}(1), \ldots, \mathbf{x}(n)$, can be written in the form:

$$\mathcal{L}(\mathbf{W}) = n \log \det \mathbf{W} + \sum_{i=1}^{n} \sum_{k=1}^{d} \log f([\mathbf{W}\mathbf{x}(i)]_k)$$

where $[\mathbf{W}\mathbf{x}(n)]_k$ denotes the kth element of $\mathbf{W}\mathbf{x}(n)$. State all your assumptions. **(6)**

Question B3

a) Explain how a 2-category classification problem can be converted into a regression problem using a mean squared error criterion. **(2)**

b) For a 2-category linear classifier, write down the closed form solution for the minimum squared error solution for the weights, identifying all your terms. **(2)**

c) Identify whether each of the following two input logic operations are linearly separable and can therefore be represented by a simple perceptron. Where possible provide an appropriate weight vector that outputs the correct symbol 0 or 1.

(i) x_1 AND x_2

(ii) NOT $(x_1$ OR $x_2)$

(iii) $(x_1$ AND NOT $x_2)$ OR $(x_2$ AND NOT $x_1)$

(7)

d) Explain how a Radial Basis Function (RBF) network can be used to transform a linearly nonseparable problem into a linearly separable one. **(3)**

e) Using part (d) above or otherwise, show that the RBF network given below can solve the XOR problem.

output unit: $z = w_1 y_1 + w_2 y_2 + w_0$

input units: $y_i = \exp(-||\mathbf{x} - \mathbf{c}_i||^2)$

where $\mathbf{x} = [x_1, x_2]^T$ are the network inputs and $\mathbf{c}_1 = [0, 0]^T$ and $\mathbf{c}_2 = [1, 1]^T$ are the RBF centres. **(6)**

END OF PAPER