UNIVERSITY OF SOUTHAMPTON CM311W1

SEMESTER 1 EXAMINATION 2002/2003

NEURAL COMPUTATION

Duration: 120 mins

Answer ALL *questions from section A (20 marks)*
and ONE *question from section B (25 marks)*
and ONE *question from section C (25 marks)*

This examination is worth 70%. The coursework was worth 30%.

Calculators without text storage MAY be used

 Number of pages 6

Section A

Question 1

a) Describe what is meant by the terms *training set*, *validation set* and *testing set*.

(3 marks)

b) Describe what is meant by the terms *classification*, *regression* and *density estimation*.

(3 marks)

c) Describe a useful real-world application of machine learning.

(4 marks)

d)

e)

f)

Section B

Question 2

a) Explain what is meant by the term over-parameterisation. State how it is removed from the hyperplane, $\boldsymbol{w}^\mathsf{T}\boldsymbol{x} + b = 0$, in the linear Support Vector Machine formulation to produce a canonical hyperplane.

(3 marks)

b) State the condition for separability of the two-class data-set $\mathcal{D} = \{\boldsymbol{x}_i, y_i\}_{i=1}^n, \boldsymbol{x}_i \in \mathbb{R}^d, y_i \in \{-1, 1\}$ with this canonical hyperplane.

(3 marks)

c) Describe the maximum margin principle and show that the resulting optimisation problem is given by the Lagrangian,

$$\Phi(\boldsymbol{w}, b, \boldsymbol{\alpha}) = \tfrac{1}{2}\|\boldsymbol{w}\|^2 - \sum_{i=1}^{n} \alpha_i \left(y_i \left[\boldsymbol{w}^\mathsf{T}\boldsymbol{x}_i + b \right] - 1 \right), \quad \alpha_i \geq 0.$$

(8 marks)

d) Solve the Lagrangian problem, $\max_{\boldsymbol{\alpha}} \left(\min_{\boldsymbol{w},b} \Phi(\boldsymbol{w}, b, \boldsymbol{\alpha}) \right)$, to show that the solution for the Lagrange multipliers can be written as a quadratic program,

$$\begin{gathered} \min_{\boldsymbol{\alpha}} \tfrac{1}{2}\boldsymbol{\alpha}^\mathsf{T} H \boldsymbol{\alpha} + \boldsymbol{c}^\mathsf{T}\boldsymbol{\alpha}, \\ \text{subject to the constraints,} \\ \alpha_i \geq 0, \quad \textstyle\sum_{j=1}^{n} \alpha_j y_j = 0. \end{gathered}$$

(8 marks)

e) What are the Support Vectors and how do these relate to the Lagrange multipliers?

(3 marks)

Question 3

a) Describe the K Nearest Neighbour method.

(8 marks)

b) Describe the methods of holdout validation, cross-validation and bootstrap validation.

(9 marks)

c) Given the following 2-class dataset, use stratified 4-fold cross

x_1	x_2	y
2.5	1.2	A
5.3	2.3	A
4.2	1.1	A
5.3	2.3	A
1.2	2.7	B
4.2	3.4	B
2.3	1.3	B
1.2	2.7	B

validation to estimate the optimal value for K, in the K nearest neighbour algorithm, and discuss the result.

(16 marks)

Section C

Question 4

Question 5

END OF PAPER