



MSC Examination by course unit

Friday, 15 May 2015 10:00 am

ECS708P/U/D Machine Learning Duration: 2 hours 30 minutes

**YOU ARE NOT PERMITTED TO READ THE CONTENTS OF THIS QUESTION PAPER UNTIL
INSTRUCTED TO DO SO BY AN INVIGILATOR**

Answer ALL FOUR questions

Calculators are permitted in this examination. Please state on your answer book the name and type of machine used.

Complete all rough workings in the answer book and cross through any work that is not to be assessed.

Possession of unauthorised material at any time when under examination conditions is an assessment offence and can lead to expulsion from QMUL. Check now to ensure you do not have any notes, mobile phones, smartwatches or unauthorised electronic devices on your person. If you do, raise your hand and give them to an invigilator immediately. It is also an offence to have any writing of any kind on your person, including on your body. If you are found to have hidden unauthorised material elsewhere, including toilets and cloakrooms it will be treated as being found in your possession. Unauthorised material found on your mobile phone or other electronic device will be considered the same as being in possession of paper notes. A mobile phone that causes a disruption in the exam is also an assessment offence.

EXAM PAPERS MUST NOT BE REMOVED FROM THE EXAM ROOM

Examiners:

Yiannis Patras

Roya Haratian

Question 1: *Linear Regression and Logistic Regression*

- (a) Describe the difference between “Linear Regression” and “Logistic Regression” in terms of their output.

[4 marks]

- (b) Explain the difference between Maximum a Posterior (MAP) estimates and Maximum Likelihood (ML) estimates.

[4 marks]

- (c) Explain the “gradient descent” algorithm.

[4 marks]

- (d) Explain what is “Regularization” and how it is applied in “Linear Regression”. Explain the problem and the “Regularization” solution.

[5 marks]

- (e) Explain “Linear Basis Function Models” for “Linear Regression”. Give three examples of the basis functions.

[8 marks, 2 for the main explanation, 2 for each example]

[Total 25]

Question 2: Neural Networks

- (a) Explain the process of training a neural network by optimising a cost function using the method of steepest gradient descent. In your answer, discuss the role of the network weights and the error measure, and explain why the use of a validation set is important.

[6 marks]

- (b) Sketch a diagram of a standard 3-layer Multilayer Perceptron (MLP) for classifying inputs x_1, \dots, x_d to a single target value t . What key limitation of the original Perceptron is overcome by the MLP?

[6 marks]

- (c) Briefly explain “feed forward neural network” and “backpropagation” technique.

[4 marks]

- (d) Implement the logical XOR function using a neural network

[9 marks]**[Total 25]**

(2015)

Question 3: K-Means

- (a) Describe the difference between supervised and unsupervised learning. Give an example of a supervised learning problem and an example of an unsupervised learning problem.

[4 marks]

- (b) Describe in detail the steps of the K-means algorithm. Make sure that you define the input to the algorithm, the output, and the dimensionality of all the variables that you use.

[8 marks]

- (c) Give the formula that defines the quantity that is optimised by the K-means algorithm.

[4 marks]

- (d) What is the optimisation method used by K-means? Using a sketch show that this general method is warranted to converge. Show that it might converge to a local minimum.

[4 marks]

- (e) The K-Means algorithm terminates at local minimum. Describe a practical way to deal with this problem.

[5 marks]

[Total 25 Marks]

Question 4: Markov Models

- (a) With a help of a diagram explain the main principles of the first-order Markov Model. Explain what is meant by the term “first-order”. What are the differences with a hidden Markov model (HMM)? In your answer, define the states ω_i , the symbols v_k , and the matrices $\mathbf{A} = [a_{ij}]$ and $\mathbf{B} = [b_{jk}]$.

[6 marks]

- (b) The decoding problem can be stated as follows: Given an HMM and a sequence of observation symbols $\mathbf{V}^{1:T}$ determine the most likely sequence of hidden states $\omega^{1:T}$. What are the other two types of problems considered in the context of HMMs?

[6 marks]

- (c) What is the name of the algorithm that is used to solve the decoding problem described in 4.b). Give a brief description of the algorithm using pseudo code.

[13 marks]

End of Paper