Two hours

QUESTION PAPER MUST NOT BE REMOVED FROM THE EXAM ROOM AND MUST BE RETURNED

UNIVERSITY OF MANCHESTER
SCHOOL OF COMPUTER SCIENCE

Machine Learning and Optimisation

Date: Monday 12th January 2015
Time: 09:45 - 11:45

Answer ALL 20 multiple choice questions in Section A
Answer ALL Questions in Section B
Answer ALL Questions in Section C

Use a SEPARATE answerbook for each SECTION.

This is a CLOSED book examination

The use of electronic calculators is permitted provided they are not programmable and do not store text

[PTO]
Section A

This Section contains Multiple Choice Questions and is therefore restricted
Section B

Answer ALL questions from this section.

1. a) Imagine we apply a Perceptron classifier to the data below. Do you expect good performance? You must provide a concise reason for your answer. (1 mark)

![Data Points](image)

b) On the same data, we now apply a $k$-nearest neighbour classifier. Give pseudo-code for the testing phase of the KNN, which should work for any value of $k$. (2 marks)

c) State the learning rule for the perceptron learning algorithm, giving definitions for all the mathematical terms used, and stating which parameter(s) can be altered manually by a user to control the learning process. (5 marks)

d) Write out full pseudo-code for the ID3 algorithm, being sure to state base cases, and state precisely (i.e. with a mathematical equation) how you would determine the most important feature at each splitpoint. (7 marks)
2. Clustering analysis is an unsupervised learning process and often required by real world applications.

a) Briefly describe the main steps in the Agglomerative algorithm for clustering analysis and give one factor that may significantly affect the performance of this algorithm.  

(5 marks)

b) YouTube website contains a large set of video clips. Design an efficient method that can group such video clips into a set of clusters effectively.  

(10 marks)