COMP24111

Two hours - online hybrid

EXAM PAPER MUST NOT BE REMOVED FROM THE EXAM ROOM

UNIVERSITY OF MANCHESTER
DEPARTMENT OF COMPUTER SCIENCE

Machine Learning and Optimisation

Date: Monday 20th January 2020

Time: 09:45 - 11:45

This is a hybrid examination with sections to be answered online and questions to be answered on paper.

Please answer All Questions in Section A and Section B online and All Questions in Section C in a separate answerbook

The examination is worth a total of 70 marks

This is a CLOSED book examination

Electronic calculators may be used in accordance with the University regulations

Sections A and B contain restricted multiple choice questions (MCQs) and are NOT published

Section C

Answer this section in an answer booklet.

1. You are asked to build a linear classifier based on four training samples. Each sample is characterised by two features: $[f_1, f_2]$. The four training samples are $[1,1]$, $[0,1]$ and $[1,0]$ from class A, and $[0,0]$ from class B.

   a) Write down the formulation of your linear classifier and state what model parameters need to be optimised. (3 marks)

   b) Use the sum-of-squares error function to train your classifier. Explain how you are going to set your target output and write down your optimisation objective function. (3 marks)

   c) Derive the partial derivatives of the objective function. (3 marks)

   d) One way to do the training is by letting the partial derivatives equal to zero and solving the resulting linear equations. Following this approach, derive the trained classifier. (3 marks)

   e) Instead of setting the partial derivatives to zero, stochastic gradient descent can be used for training. At the t-th iteration, only the sample $[1,1]$ from class A is used to estimate the gradient. Write down the model updating rule for this iteration, given learning rate fixed as $\eta = 1$. (3 marks)

   f) What happens when using $[0,0]$ from class B to update the model parameters? (3 marks)

2. a) Give pseudo-code for the testing phase of the k-nearest neighbour regression algorithm.
(4 marks)

b) List two advantages and two disadvantages of the k-nearest neighbour algorithm.
(4 marks)

3. You have learned that the gradient descent approach can be used to minimise a given function. Now you are asked to maximise a function. Design an approach for this and explain your design. (4 marks)