GLASGOW COLLEGE UESTC

Main Paper - Fall 2022-2023

ARTIFICIAL INTELLIGENCE & MACHINE LEARNING (UESTC3036)

Date: 23rd Feb 2023
Time: 09:30-11:30

Attempt all PARTS. Total 100 marks

Use one answer sheet for each of the questions in this exam.
Show all work on the answer sheet.

Make sure that your University of Glasgow and UESTC Student Identification Numbers are on all answer sheets.

An electronic calculator may be used provided that it does not allow text storage or display, or graphical display.

All graphs should be clearly labelled and sufficiently large so that all elements are easy to read.

The numbers in square brackets in the right-hand margin indicate the marks allotted to the part of the question against which the mark is shown. These marks are for guidance only.

Continued overleaf

Question 1

Multiple-choice questions [25 in total]

(a) Which of the following description is correct? [4]

A. Machine learning is a subset of AI.
B. AI and machine learning are two separate research domains without interaction.
C. AI and machine learning refer to the same thing and are interchangeable.
D. AI is a subset of machine learning.
E. AI is a subset of evolutionary computation.

(b) Which of the following description is correct? [3]

A. The hill climbing method has the mechanism to jump out of local optima.
B. The hill climbing method is a global optimization method.
C. The hill climbing method is a local search method.
D. The hill climbing method cannot obtain the global optimum in any case.
E. In the hill climbing method, using large neighborhood are cheap to explore but you might need to explore many of them compared to using small neighborhood.

(c) Which of the following description is correct? [3]

A. Quadratic assignment problem is a combinatorial optimization problem.
B. Quadratic assignment problem is a continuous optimization problem.
C. Quadratic assignment problem is neither a combinatorial nor a continuous optimization problem.
D. In quadratic assignment problem, the candidate decision variables can always be represented by binary strings.
E. In quadratic assignment problem, the candidate decision variables can always be represented by real numbers.

(d) Which of the following description is correct? [3]

A. An optimization problem can only have one objective but can have more than one constraint.
B. An optimization problem can have more than one constraint and objective at the same time.
C. When an optimization problem has more than one objective, it cannot have constraints. Otherwise, it is unsolvable.
D. In engineering optimization, constraint satisfaction is often not important when the objective function has an excellent value.
E. Feasible solutions refer to solutions with good objective function values.

Continued overleaf

(e) Which of the following description is correct? [3]

A. Pareto optimal solutions are proposed for single objective optimization problems.

B. For solutions A and B, if solution A satisfies all the constraints, while B does not, then A dominates B.

C. For solutions A and B, if solution A satisfies all the constraints, while B does not, then B dominates A.

D. Pareto set is the same as the Pareto front.

E. Pareto optimal solutions are non-dominated solutions.

(f) Which of the following description is correct? [3]

A. In the selection of GAs, better solutions tend to have fewer copies in the mating pool.

B. In the selection of GAs, better solutions tend to have more copies in the mating pool.

C. GA selection has nothing to do with fitness values.

D. Proportional selection is stronger than tournament selection.

E. Fitness scaling can improve the performance of tournament selection.

(g) Which of the following description is correct? [3]

A. In GA, the main goal of crossover is to decrease the average fitness.

B. In GA, the main goal of crossover is to increase the average fitness.

C. In GA, crossover has nothing to do with the average fitness.

D. In GA, the main goal of crossover is to carry out exploitation.

E. In GA, the crossover and mutation probability should be large.

(h) Which of the following description is correct? [3]

A. Multimodal optimization problem means that the problem has more than one local optimum.

B. Multimodal optimization problem means that the problem has more than one decision variable.

C. Multimodal optimization problem means that the problem has more than one simulation model.

D. Multimodal optimization problem means that the problem has more than one machine learning model.

E. Multimodal optimization problem means that the problem has more than one objective function.

Continued overleaf

Question 2

An electronic engineer is designing a satellite antenna, which should meet the following specifications.

Bandwidth: 1.1 GHz to 1.7 GHz

For any candidate design x, the reflection coefficient S11(x) and axial ratio AR(x) can be obtained by simulation. The |S11(x)| (i.e., magnitude) and AR(x) are vectors: For each frequency point within the bandwidth (e.g., 1.11 GHz, 1.12 GHz, …), there is a corresponding value of |S11|, as well as AR, which is a real number.

The maximum of |S11(x)| over the bandwidth: -14 dB

The maximum of AR(x) over the bandwidth: 3 dB

While satisfying the |S11| and AR specifications, the engineer hopes to make the size of the antenna as small as possible. You are the consultant helping him to design the antenna using AI techniques.

(a) Please formulate the above design problem into an optimization problem. [5]

(b) You are using the penalty function method to handle the constraints. Please formulate the penalized objective function. [5]

(c) You are using a genetic algorithm to optimize the penalized objective function. Please draw a flow diagram of genetic algorithm, showing the fundamental algorithm operators and the control flow. [5]

Please indicate: (1) What crossover and mutation rate will you use, [2]

and (2) the possible stopping criteria. [3]

(d) You obtained optimal design parameters x' satisfying the specifications and with a compact antenna size. However, the engineer hopes to see if better results are possible. Please list the possible actions that you will carry out aiming to improve the current design. [5]

Continued overleaf

Question 3

(a) Multivariant methods: In regression, fitting a quadratic is equivalent to fitting a linear model with an extra input corresponding to the square of the input. Can we also apply the same concept in classification? Justify your choice and give an example. [6]

(b) What is an outlier? How can we make *k*-means robust to outliers? [4]

(c) An engineer is given a dataset of cellular images from patients with and without cancer. Answer the following questions and explain your reason(s) clearly.

(i) If he is required to train a classifier that predicts the probability that the patient has cancer, would he use **Decision trees** or **Logistic regression**? Explain your reason(s) clearly for the chosen classifier. [5]

(ii) Suppose the dataset in question 3c(i) had 900 cancer-free images and 100 images from cancer patients. If he trains a classifier which achieves 85% accuracy on this dataset, can he say it is a good classifier? Explain your reason(s) clearly. [5]

(iii) Is it true that the error of a hypothesis measured over its training set provides a pessimistically biased estimate of the true error of the hypothesis (i.e., the real-world performance of the hypothesis would be better than the training set performance)? Justify your answer. [5]

Continued overleaf

Question 4

(a) An engineer wants to build a Neural Network that classifies two-dimensional data (i.e., X = [x1, x2]) into two classes: circles and triangles. He has a set of training data that is plotted in Figure 4a.

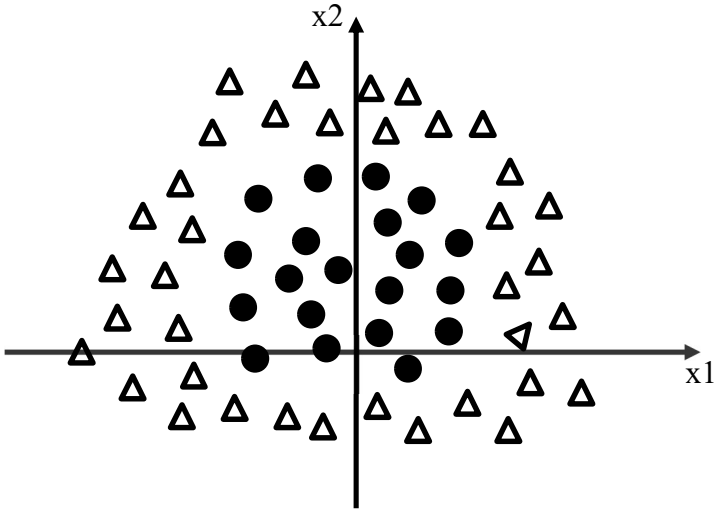

Figure 4a

(i) Draw a network that can solve this classification problem. [8]

(ii) Justify your choice of the number of nodes and the architecture. [4]

(iii) Draw the decision boundary (on the given Figure) that your network can find on the diagram. [4]

(b) Neural Network: Does increasing the number of hidden nodes in a multilayer perceptron improve generalisation? Justify your answer in both cases. [5]

(c) Naive Bayes: $X = (X_1, X_2)$ is drawn from a two-dimensional Gaussian distribution with a diagonal covariance matrix.

$$X = X_1, X_2 \sim \aleph(\mu, \sum)$$

$$\sum = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$$

where a and b are some real numbers.

Are X_1 and X_2 independent? Explain as briefly as possible. [4]

End of question paper