GLASGOW COLLEGE UESTC

Main Paper - Fall 2023-2024

ARTIFICIAL INTELLIGENCE & MACHINE LEARNING (UESTC 3036)

Date: 25th December 2023
Time: 19:00 – 21:00 (2 hours)

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

An electronic engineer is designing a microwave filter.

For any candidate design x, passband reflection coefficient S11(x) can be obtained by simulation. The |S11(x)| (i.e., magnitude) considering different frequencies is a vector: For each frequency point within the considered bands, there is a corresponding value of |S11|, which is a real number.

The design specifications are as follows: The maximum of passband reflection coefficient |S11(x)| between 296 GHz to 304 GHz should be smaller than -20 dB. The minimum of passband reflection coefficient |S11(x)| between 280 GHz to 292 GHz should be larger than -1 dB. The minimum of passband reflection coefficient |S11(x)| between 308 GHz to 320 GHz should be larger than -1 dB.

You are the consultant helping him to design the filter using AI-based optimization techniques.

(a) Please formulate the above design problem into an optimization problem by listing the objectives (if there are any) and constraints (if there are any). [5]

(b) You need to integrate the objectives (if there are any) and constraints (if there are any) into a single function for optimization. Please write that function. [5]

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

(d) If the engineer has an initial design having reasonable performance, how will you use the initial design to benefit your GA optimization and how can your GA be benefitted? [5]

(e) After a few hours, your current best design has not yet satisfied the above design specifications and has not been improved for many generations. The engineer is. wondering whether the current optimal design can be further improved much or not. What will you check to answer him and why? [5]

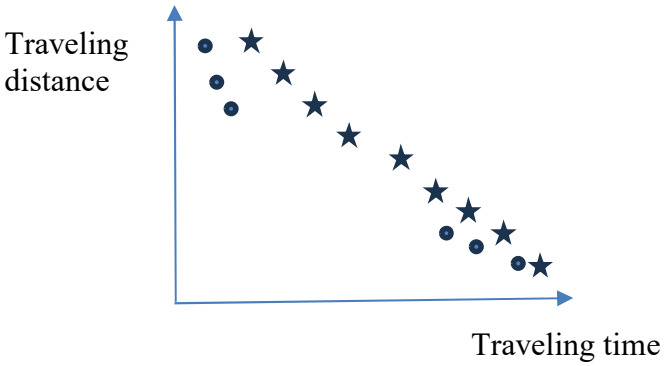
Continued overleaf

Question 2

A delivery company needs to find its optimal delivery plan. Besides the standard traveling salesman problem aiming to find the shortest traveling distance, traveling time due to traffic issues also needs to be considered. It is found that the two goals conflict with each other. You are the information engineer helping them.

(a) Do you suggest using single objective constrained optimization or multiobjective optimization? What is the reason? [5]

(b) Considering you are using multiobjective optimization and tried two multiobjective evolutionary algorithms. The generated Pareto fronts are as follows.



Circles refer to the result of algorithm 1, and stars refer to the result of algorithm 2. Please compare the results of the two algorithms and make comments. [8]

(c) You are asked to design an algorithm using GA operators to solve this problem. In the following, there are two candidate solutions A and B, assuming there are 6 places in the delivery plan. Please design a crossover operator, showing the crossover process and an example result of the crossover. [8]

A:

1	3	2	6	5	4

B:

5	2	4	3	1	6

(d) Please suggest two multiobjective evolutionary algorithm alternatives to solve this problem. [4]

Continued overleaf

Question 3

K-Nearest Neighbor (KNN) Classifier

(a) Consider a set of five training examples given as $((x_i, y_i), ci)$ values, where x_i and y_i are the two attribute values (positive integers) and c_i is the binary class label: $\{((1, 1), -1), ((1, 7), +1), ((3, 3), +1), ((5, 4), -1), ((2, 5), -1)\}$.

(i) Classify a test example at coordinates $x_t = (3, 6)$ using a KNN classifier with $k = 3$ and Manhattan distance defined by $d((x_1, y_1), (x_2, y_2)) = | x_1 - x_{12}| + | y_1 - y_2|$. Your answer should be either +1 or -1, and you must show all your calculations to receive full marks. [9]

(ii) Will the results of a general KNN classifier that uses Euclidean distance change if we multiply each example's attribute value by 0.5? Justify your answer in both cases. [4]

(iii) Why implementing a KNN model on a very large training dataset is challenging in practice? [4]

Support Vector Machines

(b) For each of the following cases, state whether it would be best to use the primal or dual SVM formulation.

(i) We apply a feature transformation that maps the input data into a feature space with infinite dimension. [4]

(ii) We apply a feature transformation that doubles the dimension of the input data. The input data has billions of training examples and is linearly separable. [4]

Continued overleaf

Question 4

Convolutional Neural Networks (CNNs)

(a) In a given object classification task, can CNNs learn to recognize an object in an image no matter how the object is translated (i.e., shifted horizontally and/or vertically), even if the training set only includes that object in one position? Justify your answer in both cases. [5]

(b) Consider a convolutional neural network that has an input layer containing a 13 x 13 image that is connected to a convolution layer using a 4 x 4 filter and a stride of 1 (i.e., the filter is shifted horizontally and vertically by 1 pixel, and only filters that are entirely inside the input array are connected to a unit in the convolution layer). There is no activation function associated with the units in the convolution layer. The convolution layer is connected to a max pooling layer using a 2 x 2 filter and a stride of 2. (Only filters that are entirely inside the array in the convolution layer are connected to a unit in the pooling layer.) The Output layer contains 4 units that each use a ReLU activation function, and these units are fully connected to the units in the pooling layer.

(i) How many units are in the convolution layer? Justify your calculations/answer with an explanation. [4]

(ii) How many distinct weights must be learned for the connections to the convolution layer? Justify your calculations/answer with an explanation. [4]

(iii) How many units are in the pooling layer? Justify your calculations/answer with an explanation. [4]

(iv) How many distinct weights must be learned to connect to the output layer? Justify your calculations/answer with an explanation. [4]

(c) Can a neural network be used to model the following machine learning algorithms? If so, state the neural network structure (how many hidden layers are required) and the activation function (s) used at the internal and output nodes. If not, briefly describe why not.

(i) K-Nearest Neighbors (KNN) [2]

(ii) Linear regression [2]

End of question paper