

NANYANG TECHNOLOGICAL UNIVERSITY**SEMESTER 2 EXAMINATION 2022-2023****EE6227 – GENETIC ALGORITHMS AND MACHINE LEARNING**

April / May 2023

Time Allowed: 3 hours

INSTRUCTIONS

1. This paper contains 4 questions and comprises 5 pages.
2. Answer all 4 questions.
3. All questions carry equal marks.
4. This is a closed book examination.
5. Unless specifically stated, all symbols have their usual meanings.

1. Question 1 consists of two independent parts.

- (a) Suppose at birth, every human being is given a fixed number of heartbeats, and once these heartbeats are used up, life ends. How should one optimally use these heartbeats to have a life as long as possible? It is tempting to suggest that one is better off staying in bed and resting, to maintain a low heartbeat. However, the science of good health does point to the fact that a trained heart beats more slowly when the person is at rest. On this note, consistent exercising is not a bad idea.

Suppose that the untrained heart beats 80 times per minute when a person is at rest, and during exercise it beats 120 times per minute. If a person exercises the fraction x of its time, the average heartbeats per minute is as follows:

$$f(x) = 120x + (1-x)g(x)$$

To determine an appropriate function $g(x)$, we can impose a requirement that $g(x)$ should be close to 80 for small x , which means that hardly any exercise is done. When x approaches 1, we assume $g(x)$ to be close to 50, meaning the person is well-conditioned. Let's assume a simple model for $g(x)$ to be:

$$g(x) = 50 + 30e^{-100x}$$

Note: Question No. 1 continues on page 2.

Produce a table according to the format shown in Table 1 in your answer script.

- (i) Fill in values for x , $g(x)$ and $f(x)$ in the table. You are free to select 20 suitable values of x to work out an approximate optimum value for exercising. State the optimum time for exercising in minutes per day.
 - (ii) Write a suitable fitness function $f^*(x)$ to scale the fitness values so that they are normalized from 0 to 1. Fill in the values for the column labelled $f^*(x)$ accordingly.
 - (iii) From an optimization perspective, justify in your own words the level of difficulty in solving for the optimum value of x .

(13 Marks)

Table 1

- (b) Based on the categorization of the type of problems according to the generic input-output model block architecture, which type of problems would you categorize the following. Provide explanations for the choice of your categorization.

 - (i) Prediction of IBM share price based on historical valuation for trading purpose.
 - (ii) Packing of odd size boxes into a truck for delivery.
 - (iii) Policy making for country's economic planning unit.
 - (iv) Rostering of security guards in a gated apartment complex

(12 Marks)

2. You are given a multi-dimensional function which is to be optimized.

$$f(X) = 0.5 + \frac{\sin^2(\sum_{i=1}^3 \sqrt{(x_i - \delta)^2}) - 0.5}{\left[1 + 0.001(\sum_{i=1}^3 \sqrt{(x_i - \delta)^2})\right]^2}$$

where:

$$\delta = 0$$

$$-10 \leq x_i \leq 10 \text{ for } i = 1, 2 \text{ and } 3$$

$$f_{min}(X^*) = 0$$

$$x_i^* = 0$$

Note: Question No. 2 continues on page 3.

- (a) Reproduce Table 2 in your answer script and complete the missing values in the table. Present the values with up to 3 decimals accuracy. (7 Marks)
- (b) Apply the arithmetic recombination $z_i = \alpha x_i + (1 - \alpha)y_i$ with $\alpha = 0.4$ to produce two offspring from each crossover operation. Fill up the columns labelled Child1 and Child2. (8 Marks)
- (c) The representation of the chromosome can be a binary string. Write the 18-bit binary strings for the 2 fittest chromosomes. (5 Marks)
- (d) Let $\delta = 0.32$. Write the 18-bit binary string for X^* . (5 Marks)

Table 2

Chromosome ID	$X=(x_1, x_2, x_3)$	$f(X)$	Crossover Mate	Child1	Child2
1	(-2.3, 1.5, 0.5)		3		
2	(1.1, -0.3, -5)		4		
3	(1.2, 0.8, -1.8)		5		
4	(7.8, 9.2, 3.3)		1		
5	(-0.5, 0.5, 0.8)		2		

3.

- (a) For a given dataset of a pattern classification problem, it is known that most of the features are insignificant or redundant, which may have detrimental effects on the classification performance. Suggest a feature selection method to address the issue and describe the procedure of the suggested feature selection method.

(9 Marks)

- (b) A training dataset contains 400 samples with equal proportions from two classes, 1 and -1 . Assume that the samples in the two classes follow two independent normal distributions, and the mean vectors and scatter matrices of classes 1 and -1 are given as follows:

$$\mathbf{m}_1 = \begin{bmatrix} -0.98 \\ -1.93 \\ -0.96 \end{bmatrix}, \quad \mathbf{m}_2 = \begin{bmatrix} 0.96 \\ 2.10 \\ 1.04 \end{bmatrix}$$

$$\mathbf{S}_1 = \begin{bmatrix} 215.67 & -5.16 & -29.73 \\ -5.16 & 182 & -2.36 \\ -29.73 & -2.36 & 22.52 \end{bmatrix}, \quad \mathbf{S}_2 = \begin{bmatrix} 205.13 & -1.39 & 7.87 \\ -1.39 & 182.41 & 15.11 \\ 7.87 & 15.11 & 159.30 \end{bmatrix}.$$

Design a Fisher linear discriminant classifier for the 2-class classification problem.

(8 Marks)

- (c) In the evaluation of a pattern classifier, the repeated k -fold cross validation method is often used. Briefly describe the method and discuss the rationale of the method.

(8 Marks)

- 4 (a) There are five samples, s_1, s_2, s_3, s_4 and s_5 . The Euclidean distance matrix of the samples is given as follows:

	s_1	s_2	s_3	s_4	s_5
s_1	0	0.1732	1.6673	1.1489	1.1045
s_2	0.1732	0	1.6401	1.0724	1.0247
s_3	1.6673	1.6401	0	1.0677	0.9695
s_4	1.1489	1.0724	1.0677	0	0.1414
s_5	1.1045	1.0247	0.9695	0.1414	0

Note: Question No. 4 continues on page 5.

Cluster the samples using the agglomerative hierarchical clustering algorithm and draw the dendrogram. Assuming that the single linkage method is used.

(15 Marks)

- (b) In a clustering task, it is known that the data is noisy and form clusters with irregular shape in the feature space. However, the number of clusters is unknown. Select a method to solve the clustering problem, explain the rationale for your selection, and describe the procedure of the selected method.

(10 Marks)

END OF PAPER

EE6227 GENETIC ALGORITHMS & MACHINE LEARNING

Please read the following instructions carefully:

- 1. Please do not turn over the question paper until you are told to do so. Disciplinary action may be taken against you if you do so.**
2. You are not allowed to leave the examination hall unless accompanied by an invigilator. You may raise your hand if you need to communicate with the invigilator.
3. Please write your Matriculation Number on the front of the answer book.
4. Please indicate clearly in the answer book (at the appropriate place) if you are continuing the answer to a question elsewhere in the book.