LNMIIT/B. Tech./PE/2019-20/ODD/CSE/GAA/ET

LNMIIT

The LNM Institute of Information Technology

Department Name: CSE

Ujjawal Kumar
17UCS171.

Genetic Algorithms & Applications(GAA 3031)

Exam Type(End Term)

Time: 180 Minutes **Date: 11/12/2019** **Max. Marks: 40**

Instruction: *Answer must be in brief and to the point. No query will be handled during exam. Though careful proof reading has been done, even then if you have any doubt/confusion regarding the question you can make your assumptions. You must write your assumptions clearly before you start attempting that question. If instructor thinks that your doubt/confusion/assumption is genuine, then it will be entertained.*

Question 1 **Total Marks 10 [3(1.5+1.5)+1+1+3(2+1)+1+1]**

Design a genetic algorithm based rule set (Containing two rules) for the classification of the following data. You will have to provide complete details regarding representation scheme, fitness function, modified/usual genetic operators (Crossover and mutation) and finally two rules (one for yes class and second for no class) for the classification problem after termination of genetic algorithm.

TABLE 1. Object Classification Training Set

Index	Size	Color	Shape	Decision
1	medium	Red	brick	Yes
2	small	Blue	wedge	No
3	small	Red	sphere	Yes
4	large	Red	wedge	No
5	large	Green	pillar	Yes
6	small	Blue	wedge	No
7	large	Green	sphere	Yes

Note: Choose the following parameters for genetic algorithms

Population size: 4; Crossover probability: 1; Mutation probability: 1; Generation model with elitism; Termination condition: Two generations including initial generation;

Parent selection: All distinct chromosomes will be in mating pool; Crossover points: 1 and 15 (If required); Mutation point: 3 (If required); Rules for these two classes will be formed based on the best two chromosomes (for different class) in the last generation.

Question 2 **Total Marks 25 [{3+3+2(1+1) +1+4+1+2+4+1}+{2+2}]**

Consider the following data [described in Table 2(Rating Matrix) and Table 3(Item Features)] for designing a prototype of hybrid recommender system based on collaborative & content based filtering for active user U1 using Genetic algorithm. In this prototype, genetic algorithm will be used to learn appropriate weights for different recommenders. You will have to provide complete details regarding representation scheme, fitness function, genetic operators (Crossover and mutation) and finally appropriate weight vector for the process of collaborative & content based filtering after the termination of genetic algorithm.

Use this weight vector in the developing a prototype of hybrid recommender system and compute mean absolute error for user U1. You will have to discuss each step with details including necessary equations.

Tables 2 [Rating Matrix] and 3 [Item Features]

LNMIIT

User/Item	M1	M2	M3	M4	M5	M6
U1	4	3	4	3	4	5
U2	3	4	1	3	4	4
U3	4	3	3	2	3	4
U4	5	4	4	5	4	3

Items/Features	G1	G2	G3	G4	G5
M1	1	0	1	1	0
M2	1	1	0	1	1
M3	1	0	1	0	1
M4	0	1	0	1	1
M5	0	1	1	1	1
M6	1	0	0	1	1

Please consider the following parameters for the above problem:
Active user: U1; Training movies for U1:M1, M2, M3, M4; Test movies for U1: M5 & M6; Population size: 4; Crossover and mutation probabilities: 1; Parent selection: All distinct chromosomes with elitism; Termination condition (number of generations): 2 (including initial generation); Neighbourhood Size: 2; Best chromosome will be selected from the terminating generation.

Question 3 **Marks 1+1+1+1+1= 05**

a) Write at least 4 chromosome set, which are identified by schema S = (01 *1 * *)
b) Find the defining length and order of the following schema.
$S_1 = (0 * * 11 * 0 * * *)$
c) Discuss the bloat problem in genetic programming with some possible solutions
d) Compute the disruptive single point crossover probability for the following schema
H =* * 1 0 1 1 * * Note: Crossover probability is 0.9
e) Construct a parse tree for the following logical formula

$$(x \vee true) \rightarrow \big((x \wedge y) \wedge (z \leftrightarrow (x \vee y))\big)$$

Some Important Formulae

$$Sim(u_a, u_x) = \frac{\sum_{i \in S_{u_a,u_x}} (r_{u_a,i} - \overline{r_{u_a}})(r_{u_x,i} - \overline{r_{u_x}})}{\sqrt{\sum_{i \in S_{u_a,u_x}} (r_{u_a,i} - \overline{r_{u_a}})^2} \sqrt{\sum_{i \in S_{u_a,u_x}} (r_{u_x,i} - \overline{r_{u_x}})^2}}$$

$$P(u_a, i) = \overline{r_{u_a}} + \frac{\sum_{u \in N} (r_{u,i} - \overline{r_u})}{\sum_{\in N} |Sim(u_a, u)|}$$

$$Sim(M_x, M_y) = \frac{\overrightarrow{M_x} \cdot \overrightarrow{M_y}}{\|\overrightarrow{M_x}\| \, \|\overrightarrow{M_y}\|}$$

Note: You can interpret the symbols based on the class.