PES UNIVERSITY

PES University, Bengaluru

(Established under Karnataka Act No. 16 of 2013)

UE18CS303

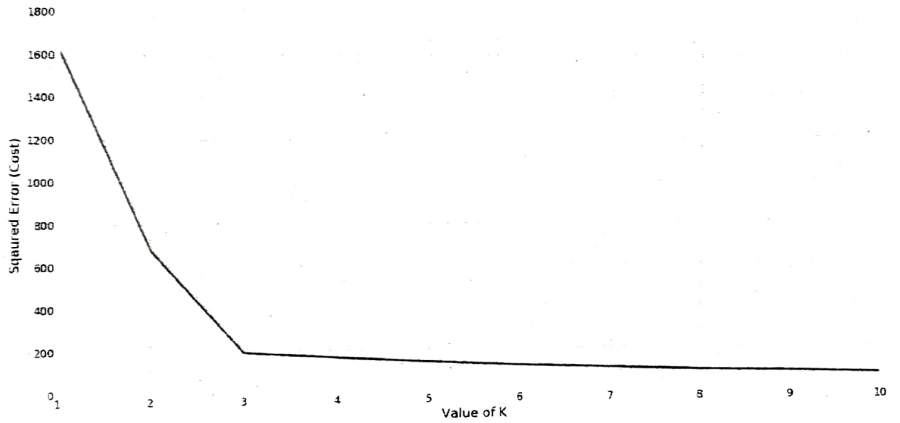
NOVEMBER 2020: IN SEMESTER ASSESSMENT B Tech CSE 5th SEMESTER TEST – 2

UE18CS303 – MACHINE INTELLIGENCE

Time: 80 Minutes	Answer All Questions	Max Marks: 40

1. a) (i) Explain how to use the elbow method to find the most optimal value of k in k-means clustering. Explain what is an elbow point. (2+1)
(ii) Is convergence always guaranteed in k-means? (1)
(iii) In the given figure below what is the most optimal value of k? (1)

5



b) Determine the frequent itemsets for the following transactions with a minimum support value of 50%. Build the associate rule graph and determine any two associate rules having a confidence value greater than 75%.

5

Transactions	Potato	Onion	Milk	Bread	Canned Drink
t1	1	1	0	1	0
t2	1	0	1	1	0
t3	0	0	1	0	1
t4	1	1	1	0	0
t5	1	1	0	1	1
t6	1	1	1	1	0

2.	a)	Find the frequent item sets ending with 'm' using FP Growth Algorithm. Build the fp-tree and condition based pattern tree (cbp-tree) and assume the threshold to be 50%.	5

Transactions	Items
t1	f,a,c,d,g,i,m,p
t2	a,b,c,f,l,m,o
t3	b,f,h,j,o
t4	b,c,k,s,p
t5	a,f,c,e,l,p,m,n

	b)	Find the singular values of the matrix $B = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$.	3
	c)	Explain the Steps of PCA	2
3.	a)	(i) List the 5 main features of genetic algorithms (1) (ii)What two requirements should a problem satisfy in order to be suitable for solving it by a GA? (2) (iii)What kind of a population is a genetic algorithm (GA) for optimisation most likely to succeed in? What should ideally be the crossover and mutation probability? (2) (iv) For a single point crossover with cross over point in the middle what will be the child chromosomes for the following two parent chromosomes 110111 and 100011 (2)	7
	b)	A budget airline company operates 3 plains and employs 5 cabin crews. Only one crew can operate on any plane on a single day, and each crew cannot work for more than two days in a row. The company uses all planes every day. A Genetic Algorithm is used to work out the best combination of crews on any particular day. a) Suggest what chromosome could represent an individual in this algorithm? b) Suggest a fitness function for this problem.	3
4.	a)	Suppose a genetic algorithm uses chromosomes of the form x = abcdefgh with a fixed length of eight genes. Each gene can be any digit between 0	7

		and 9. Let the fitness of individual x be calculated as: $f(x) = (a + b) - (c + d) + (e + f) - (g + h)$, and let the initial population consist of four individuals with the following chromosomes: x1 = 6 5 4 1 3 5 3 2 x2 = 8 7 1 2 6 6 0 1 x3 = 2 3 9 2 1 2 8 5 x4 = 4 1 8 5 2 0 9 4 a) Evaluate the fitness of each individual, showing all your workings, and arrange them in order with the fittest first and the least fit last. (2) b)Perform the following crossover operations: (i) Cross the fittest two individuals using one–point crossover at the middle point. (1) (ii) Cross the first and third fittest individuals (ranked 1st and 3rd) using a uniform crossover. (1) (iii) Evaluate the fitness of the new population consisting of the 4 offspring resulting from the crossover previous questions. Has the overall fitness improved? (3)	
	b)	i) If the population size is N and the number of iterations is T, how many times the objective function is to be evaluated (including the initial objective function evaluations) in PSO? ii) How many user defined parameters are required in PSO apart from the population size and number of iterations?	2
	c)	The VC dimension of hypothesis space H1 is larger than the VC dimension of hypothesis space H2. Which of the following can be inferred from this? A) The number of examples required for learning a hypothesis in H1 is larger than the number of examples required for H2. B) The number of examples required for learning a hypothesis in H1 is smaller than the number of examples required for H2. C) No relation to number of samples required for PAC learning.	1