BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI

Work Integrated Learning Programmes Division

Cluster Programme - M. Tech in AI & ML II Semester, 2022 – 23(July,2023) Mid semester Examination (MAKEUP)

Course No : AIMLC ZC418

Course Title : Introduction to Statistical Methods

Nature of Exam. : Open Book(Online)

Weightage : 30 Marks
Duration : 120 minutes

Date : 29th July,2023_FN

Number	of questions:4	
Number	of Pages: 2	

la a sallaga 1									
In a college 45% students registered for Machine Learning (ML) course, 60% registered for Introduction to Statistical Methods(ISM) course and 50% registered for Deep Learning (DL) course. 15% registered for ML & ISM,25% registered for ISM & DL, 20% registered for ISM & DL. 5% registered for all the three courses. <i>Validate this data</i> . If valid, then find the following i).Percentage of students registered for ML course only. ii).Percentage of students registered for ML but not DL. iii). Percentage of students registered for ML and ISM but not DL. iv).Percentage of students registered for ML and DL but not ISM. Following is the statistical summary of a data set. Write any 3 useful observations from this summary which helps you in Proceeding further with ML									
modelling.	count	mean	std	min	25%	50%	75%	max	
symboling	205.0	0.834146	1.245307	-2.00	0.00	1.00	2.00	3.00	
wheel_base	205.0	98.756585	6.021776	86.60	94.50	97.00	102.40	120.90	
length	205.0	174.049268	12.337289	141.10	166.30	173.20	183.10	208.10	
width	205.0	65.907805	2.145204	60.30	64.10	65.50	66.90	72.30	
height	205.0	53.724878	2.443522	47.80	52.00	54.10	55.50	59.80	
curb_weight	205.0	2555.565854	520.680204	1488.00	2145.00	2414.00	2935.00	4066.00	3M
engine_size	205.0	126.907317	41.642693	61.00	97.00	120.00	141.00	326.00	SIVI
bore	205.0	3.329366	0.270858	2.54	3.15	3.31	3.58	3.94	
stroke	205.0	3.256098	0.313634	2.07	3.11	3.29	3.41	4.17	
compression_ratio	205.0	10.142537	3.972040	7.00	8.60	9.00	9.40	23.00	
horsepower	205.0	104.165854	39.529733	48.00	70.00	95.00	116.00	288.00	
peak_rpm	205.0	5126.097561	477.035772	4150.00	4800.00	5200.00	5500.00	6600.00	
city_mpg	205.0	25.219512	6.542142	13.00	19.00	24.00	30.00	49.00	
highway_mpg	205.0	30.751220	6.886443	16.00	25.00	30.00	34.00	54.00	
	registered for registered for three courses i).Percentage ii).Percentage iii).Percentage iv).Percentage Following is the Write any 3 use Proceeding fur modelling. symboling wheel_base length width height curb_weight engine_size bore stroke compression_ratio horsepower peak_rpm city_mpg highway_mpg	registered for Deep registered for ISM 8 three courses. Valid i). Percentage of studii). Percentage of studiii). Percentage of studiiii). Percentage o	registered for Deep Learning registered for ISM & DL, 20% r three courses. Validate this day i). Percentage of students regisii). Percentage of students regisii). Percentage of students regisii). Percentage of students regisiv). Percentage of students regisivolents regisivol	registered for Deep Learning (DL) course registered for ISM & DL, 20% registered three courses. Validate this data. If valid, i). Percentage of students registered for ii). Percentage of students registered for iii). Percentage of students registered for iv). Percentage of students	registered for Deep Learning (DL) course. 159 registered for ISM & DL, 20% registered for ISM three courses. Validate this data. If valid, then fi). Percentage of students registered for ML cou ii). Percentage of students registered for ML an iii). Percentage of students registered for ML an iv). Percentage of students registered for ML an Following is the statistical summary of a data so Write any 3 useful observations from this summ Proceeding further with ML modelling. count mean std min symboling 205.0 0.834146 1.245307 -2.00 wheel_base 205.0 174.049268 12.337289 141.10 width 205.0 65.907805 2.145204 60.30 height 205.0 65.907805 2.145204 60.30 height 205.0 53.724878 2.443522 47.80 curb_weight 205.0 2555.565854 520.680204 1488.00 engine_size 205.0 126.907317 41.642693 61.00 bore 205.0 3.256098 0.313634 2.07 compression_ratio 205.0 10.142537 3.972040 7.00 horsepower 205.0 104.165854 39.529733 48.00 peak_rpm 205.0 5126.097561 477.035772 4150.00 city_mpg 205.0 25.219512 6.542142 13.00 highway_mpg 205.0 30.751220 6.886443 16.00	registered for Deep Learning (DL) course. 15% regist registered for ISM & DL, 20% registered for ISM & DL. three courses. Validate this data. If valid, then find the i). Percentage of students registered for ML course onlii). Percentage of students registered for ML and ISM iv). Percentage of students registered for ML and ISM iv). Percentage of students registered for ML and DL but not Dl but	registered for Deep Learning (DL) course. 15% registered for registered for ISM & DL, 20% registered for ISM & DL. 5% registered for ISM but not DL. III). Percentage of students registered for ISM but not ISM but not ISM percentage of students registered for ISM and ISM but not ISM percentage of students registered for ISM and ISM but not ISM percentage of students registered for ISM and ISM but not ISM percentage of students registered for ISM and ISM but not ISM percentage of students registered for ISM and ISM but not ISM but	registered for Deep Learning (DL) course. 15% registered for ML & registered for ISM & DL, 20% registered for ISM & DL. 5% registered three courses. Validate this data. If valid, then find the following i).Percentage of students registered for ML course only. iii). Percentage of students registered for ML but not DL. iii). Percentage of students registered for ML and ISM but not DL. iv).Percentage of students registered for ML and ISM but not DL. iv).Percentage of students registered for ML and DL but not ISM. Following is the statistical summary of a data set. Write any 3 useful observations from this summary which helps you is Proceeding further with ML modelling. Count mean std min 25% 56% 75%	registered for Deep Learning (DL) course. 15% registered for ML & ISM,25% registered for ISM & DL, 20% registered for ISM & DL. 5% registered for all the three courses. Validate this data. If valid, then find the following i).Percentage of students registered for ML course only. iii).Percentage of students registered for ML and ISM but not DL. iii). Percentage of students registered for ML and ISM but not DL. iii). Percentage of students registered for ML and ISM but not ISM. Following is the statistical summary of a data set. Write any 3 useful observations from this summary which helps you in Proceeding further with ML modelling. count

	$I(X)D(A^*IA \cap B)$	1
	iv) $P(A^c A \cap B)$	
Q.2.b)	Probabilities of Mr.Steves and Mr. John graduating from a University are 0.6 and 0.4 respectively. Probabilities of they getting placement offer after graduation are 0.15 and 0.25 respectively. Find the following i) Probability that Mr Steves was not graduated given that he got the placement offer. ii) Probability that Mr John was not graduated given that he got the	3M
	placement offer.	
	iii) Probability that both got placement offer but not graduated.	
Q.3.a)	Consider the following function:	
	$f(x,y) = \frac{x+ky}{2}, 0 < x < 1 \text{ and } 0 < y < 2$	4 M
	i).Is there any possibility of considering f(x) as probability density function? Discuss.	
	ii).If possible find mean and variance of the distribution.	
Q.3.b)	Consider the probabilities of two random variables X and Y, which are independent. $P(x = 0) = 0.20$, $P(x = 1) = 0.50$, $P(x = 2) = 0.30$, $P(y = -1) = 0.25$, $P(y = 0) = 0.35$, $P(y = 1) = 0.40$. i). Find the joint the probability distribution of (X, Y) .	4 M
	ii). $P(x < 2 / y = -1)$.	
	iii). $P(x < 1/y < 1)$.	
Q.4.a)	It is assumed that average marks of students in a course is 65 with standard deviation 5. Sampling is done by considering samples of size 45. Then find	4 M
	i).the probability that mean of the sampling distribution lies between	
	55 and 70.	
0.4 b)	ii).the probability that mean of the sampling distribution is 70.	
Q.4.b)	Consider that the time taken for processing a product(X)follows normal distribution with mean 20 minutes with variance 4. Then find the following i). $P(X > 25)$ ii) $P(X > 28)$ iii) $P(X = 22)$	4 M

XXXXXXX