			Statistics	G		
Day Time	& Da : 03	ate: Friday, 26-04-2019 :30 PM to 06:00 PM		Max. Marks: 70		
Instr	ucti	ons: 1) All questions are 2) Figures to the rig	e compulsory. ght indicate full mai	ks.		
Q.1	 Choose Correct Alternative from the fol 1) The problem of finding hidden structure a) supervised learning c) mixed learning 			l owing. in unlabeled data is called b) unsupervised learning d) all of these		
	2)	The full form of KDD is _ a) Knowledge database b) Knowledge discovery c) Knowledge data hou d) Knowledge data defi	 e in domain y in database se nition			
	3)	Which of the following isa) Clusteringc) Summarization	a predictive mode b) d)	l? Regression Association rule		
	4)	Extreme values that occ a) outliers c) dimensionality reduc	ur infrequently are b) tion d)	called as rare values all of the above		
	5)	Market-basket problem a) Agrawal et al. c) Toda et al.	was formulated by b) d)	 Steve et al. Simon et al.		
	6)	The proportion of transa called a) confidence c) support count	ction supporting X b) d)	in Transactional database is support all of the above		
	7)	The Apriori algorithm is a) top-down search c) depth first search	a b) d)	breadth first search bottom-up search		
	8)	The goal of is to data set a) Classification c) Association rule	discover both the b) d)	dense and sparse regions of a Clustering Genetic Algorithm		
	9)	clustering technic tries to split that cluster a) Divisive c) Agglomerative	ques starts with all into small pieces. b) d)	records in one cluster and then Numeric Partition		
	10) Each neuron is made u a) Molecules c) Dendrites	p of a number of no b) d)	erve fibres called Atoms Electrons		

Page 1 of 2

SLR-ES-388

Set

Ρ

Seat No.

M.Sc. (Semester - IV) (New) (CBCS) Examination March/April-2019

Q.

- a) Moleculesc) Dendrites

SLR-ES-388

	11) In a feed- forward network, the connect from input to outputa) bidirectionalc) unidirectional	tions b)	between layers are multidirectional	
	 12) Data by itself is not useful unless a) It is massive b) It is properly stated c) It is collected from diverse sources d) It is processed to obtain information 	u)	Directional	
	13) Clustering isa) supervised learningc) mixed learning	b) d)	unsupervised learning None of these	
	14) A Data that are not of interest to the daa) changing datac) noisy data	ta m b) d)	nining task is called as missing data irrelevant data	
Q.2	 Answer the following (Any four) 1) Define association rule. 2) Define confusion matrix. 3) Define Information Gain. 4) Define metadata. 5) Define Gini index. 			08
	 B) Write Notes on (Any two) 1) Define activation function. Also desc 2) What are the major tasks in data min 3) Write a short note on Agglomerative 	ribe ning Hie	its types. ? rarchical clustering method.	06
Q.3	 A) Answer the following (Any two) 1) Write a short note on classification b 2) Describe single linkage method of cl 3) Explain SEMMA data mining proces 	ase luste s.	d on regression model. ering.	08
	 B) Answer the following (Any one) 1) Explain McCulloch-Pitts AN model. 2) Explain the following: i) Training data ii) Testing data iii) Validation of model 			06
Q.4	 A) Answer the following (Any two) 1) Explain k- means clustering. 2) Explain support vector machine in bit 3) Define: i) Accuracy ii) Sensitivity iii) Specificity iv) Precision, in the context of evaluation 	rief. atinę	g classifier performance	10
	 B) Answer the following (Any one) 1) Describe Apriori algorithm with an ex 2) Explain: Mini-max normalization 	xam	ple.	04
Q.5	 Answer the following (Any two) a) Explain Naive Baye's classifier. b) Write a note on grid based clustering matching c) Explain briefly different classification technology 	ethc chnic	d. ques.	14

			DATA N	IININ	G	
Day & Time	& Date : 03:00	: Thu) PM	ursday, 14-11-2019 To 05:30 PM			Max. Marks: 70
Instr	uction	1) 2)	All questions are compulsory. Figures to the right indicate ful	ll mark	S.	
Q.1	Fill ir	14				
	1)	Rem a) c)	noving duplicate records is a pro recovery data washing	ocess b) d)	called data cleaning data pruning	
	2)	Whi a) c)	ch of the following is the other r Exploratory data analysis. Deductive learning.	name d b) d)	of Data mining? Data driven discovery All of the above	<i>י</i> .
	3)	The a) b) c) d)	full form of KDD is Knowledge database. Knowledge discovery in datab Knowledge data house. Knowledge data definition.	ase.		
	4)	Tasl a) c)	k of inferring a model from labe supervised learning both (a) and (b)	led trai b) d)	ining data is called unsupervised learning none of these	 J
	5)	a) c)	maps data into predefined g Regression Prediction	roups. b) d)	Time series analysis Classification	
	6)	a) c)	is a the input to KDD. Data Query	b) d)	Information Process	
	7)	Trea a)	ating incorrect or missing data is selection	s calleo b)	d as preprocessing	
	8)	c) a)	transformation data are noisy and have ma Discretized Real-world	d) ny mis b) d)	interpretation sing attribute values. Cleaned Transformed	
	9)	Marl a)	ket-basket problem was formula Agrawal et al.	ated by b)	/ Steve et al.	
	10)	c) The data a)	absolute number of transaction base is called	a) ns supp b)	Simon et. al. porting X in Transaction	nal
	11)	c) The a)	support count second phase of Apriori algorit Candidate generation	d) hm is ₋ b)	none of the above Itemset generation	

Statistics

Pruning

C)

M.Sc. (Semester - IV) (New) (CBCS) Examination Oct/Nov-2019

SLR-JS-389

Set P

- b) Itemset generation
- d) Partitioning

Seat No.

12)	clustering technique starts with as many clusters as there are
	records, with each cluster having only one record.

c) Partition d) Numeric

	13)	In algorithm each cluster is represented by the centre of gravity of the cluster. a) Factor analysis b) k-means c) STIRR		
	14)	c) STRRd) ROCKThe sigmoid function also knows as functions.a) regressionb) logisticc) probabilityd) neural		
Q.2	A)	Answer the following questions. (Any Four)01)What is meant by data mining?2)Define association rule.3)State anti-monotone property.4)Give an example of an activation function.5)Define metadata.	8	
	B)	Write Notes. (Any Two)01)What is unidirectional association? Explain with suitable example.2)What are the major tasks in data mining?3)Write a short note on Divisive Hierarchical clustering method.	6	
Q.3	A)	Answer the following questions. (Any Two)01)Explain CRISP data mining process.2)Describe complete linkage method of clustering.3)Write a short note on Outlier Analysis.	8	
	B)	Answer the following questions. (Any One)01)Explain McCulloch-Pitts ANN model.2)Describe single layer feed forward network in the context of ANN.	6	
Q.4	A)	 Answer the following questions. (Any Two) 1) Write a note on grid based clustering method. 2) Explain support vector machine in brief. 3) Define: i) Accuracy ii) Sensitivity iii) Specificity and iv) Precision, in the context of evaluating classifier performance. 		
	B)	Answer the following questions. (Any One)01)Describe supervised and unsupervised learning.2)Write a note on Market Basket Analysis.	4	

SLR-JS-389

Q.5 Answer the following questions. (Any Two)

- 1) Write a note on Density Based Spatial Clustering of Application with Noise (DBSCAN) algorithm.
- 2) Consider the following transactional database D. Assuming minimum support 60% and minimum confidence of 80%, find all frequent items using apriori algorithm. Also give strong association rule.

TID	Items
T100	K,A,D,B
T200	D,,A,C,E,B
T300	C,A,B,E
T400	B,A,D

3) Explain Naive Baye's classifier.

14