Reg. No															

B.Tech DEGREE EXAMINATION, NOVEMBER 2023

Seventh Semester

18ECE471T - DATA SCIENCE FOR COMMUNICATION NETWORKS

(For the candidates admitted during the academic year 2020 - 2021 & 2021 - 2022)

Note:

i. **Part - A** should be answered in OMR sheet within first 40 minutes and OMR sheet should be handed over to hall invigilator at the end of 40th minute.
ii. **Part - B** and **Part - C** should be answered in answer booklet.

Time: 3 Hours **Max. Marks: 100**

PART - A (20 × 1 = 20 Marks) Marks BL CO

Answer **all** Questions

1. Formula for correlation between two variables x,y is_____ 1 1 1
   (A) cov(x,y)/sd(x)sd(y) (B) cov(x,y)/var(x)var(y)
   (C) cov(x,y)/cov(x)cov(y) (D) avg(x,y)/sd(x)sd(y)
2. Propotion of 0.95 means 1 2 1
   (A) 95% of data is within ±2sds (B) 95% of data is within ±1sds
   (C) 95% of data is within ±3sds (D) 95% of data is within ±4sds
3. Consider the differences between the first and third quartiles to the median If one of these differences is larger than the other then it is called _____ 1 2 1
   (A) skew (B) Mean
   (C) IQR (D) Outlier
4. y(x) = e(a0 + a1x1 + a2x2 + … + aixi) / (1 + e(a0 + a1x1 + a2x2 + … + aixi)) is the equation for _____ 1 2 1
   (A) Logistic Regression (B) Linear Regression
   (C) Polynomial Regression (D) Supervised Learning
5. The three transition strategies from IPv4 to IPv6 are 1 1 2
   (A) Dual stack, Tunneling strategy, Header translation strategy (B) Dual queue, Tunneling strategy, Header translation strategy
   (C) Tunnel stack, Dual strategy, Footer translation strategy (D) Header stack, Tunneling strategy, Dual translation strategy
6. Call Switching Infrastructure is an attribute of 1 1 2
   (A) IP traffic flow records (B) Telephony call record details
   (C) Protocol transitions (D) Mobile Communication Networks
7. BSC in Mobile Communication Networks stands for 1 2 2
   (A) Base Service Control (B) Building Station Communication
   (C) Base Station Controller (D) Base Station Communication
8. VANETs consist of three components namely OBU,TA and _____ 1 2 2
   (A) TRU (B) RTU
   (C) BTU (D) RSU
9. This component belongs to the Seven visual variables 1 1 3
   (A) Saturation (B) Texture
   (C) Hue (D) Luminance

10.	Examples of Ineffective visualizations are (A) short tables and simple charts (B) lengthy tables and simple charts (C) lengthy tables and complex charts (D) right chart type and appropriate tables	1	2	3
11.	The four types of data visualizations are (A) Descriptive ,datadriven, exploratory, conceptual (B) Declarative, cognitive ,exploratory, conceptual (C) Declarative, datadriven, supplementary , conceptual (D) Declarative,datadriven,exploratory, conceptual	1	1	3
12.	In supervised learning, the algorithm is trained on a _____where the input data is paired with the corresponding output or target variable. (A) labeled dataset (B) unstructured dataset (C) ordinal dataset (D) raw dataset	1	2	3
13.	The fundamental unit of network is (A) brain (B) nucleus (C) neuron (D) axon	1	1	4
14.	The independent variable is used to explain the dependent variable in _____. (A) Linear regression analysis (B) Multiple regression analysis (C) Non-linear regression analysis (D) Polynomial regression analysis	1	2	4
15.	Among the following option identify the one which is not a type of learning (A) Reinforcement learning (B) Supervised learning (C) Semi unsupervised learning (D) Unsupervised learning	1	2	4
16.	Which of the following machine learning techniques helps in detecting the outliers in data? (A) Clustering (B) Classification (C) Anomaly detection (D) Regression	1	2	4
17.	NAT stands for _____ (A) Network activation translation (B) Network activation transmission (C) Network address translation (D) Network address transmission	1	1	5
18.	QoS Metrics for Measuring the ISP network are_____ (A) bandwidth (B) spectrum efficiency (C) Throughput (D) jitter and packet reordering	1	2	5
19.	Challenges of ISP data analysis include scale, complexity, and _____nature of these networks (A) static (B) dynamic (C) expanding (D) contracting	1	2	5
20.	CLV stands for_____ (A) Customer Lifetime Value (B) Communication Lifetime Value (C) Customer Long term value (D) Communication Long term value	1	1	5

PART - B (5 × 4 = 20 Marks)

Answer **any 5** Questions

		Marks	BL	CO
21.	For the following data given, apply different binning methods to clean the data 4,8,9,15,21,21,24,25,26,28,29,34.	4	3	1
22.	Write short notes on data transformation techniques	4	1	1
23.	Draw the block diagram of Mobile Communication Network components.	4	3	2
24.	Briefly discuss about the attributes of IP traffic flow records generated by routers.	4	2	2

		Marks	BL	CO
25.	Differentiate data mining from machine learning.	4	2	3
26.	Briefly explain about regression.	4	1	4
27.	Write short note on the importance of Customer Behavior analysis in telecom industry.	4	2	5

PART - C (5 × 12 = 60 Marks)

Answer **all** Questions

		Marks	BL	CO
28.	(a) With neat sketches explain the various Visualization techniques of categorical and numerical variable dataset.	12	3	1
	(OR)			
	(b) What is the significance of Summary statistics of exploratory data analysis? Elaborate on each one.			
29.	(a) Explain in detail about the collection of Public safety/security data in Vehicular networks.	12	2	2
	(OR)			
	(b) Discuss in detail about the various issues related to Security in data acquisition?			
30.	(a) Discuss the features of any four Modern visualization tools and techniques available for data visualization in detail.	12	2	3
	(OR)			
	(b) Explain about the types of Learning Paradigms for Data Science in detail.			
31.	(a) Explain in detail about data science problem in wireless networks	12	2	4
	(OR)			
	(b) Discuss the importance of linear regression and its applications in detail.			
32.	(a) Discuss the various challenges associated with ISP data analysis.	12	2	5
	(OR)			
	(b) Elaborate on the techniques for Customer Churn Prevention using data analytics.			

* * * * *