

Time: 3 Hours

Max Marks: 80

Instructions:

- Figures to the right indicate max marks.
- Draw appropriate diagram whenever applicable.
- Assume suitable data wherever applicable. State your assumptions clearly.
- **Question number 1 is compulsory.**
- Attempt **any Three** questions from remaining questions.

**Q1 Attempt Any Four from the following. (5 marks each) 20**

- Differentiate different Machine Learning approaches. What is Cross Validation? Discuss bias variance trade-off with suitable diagram.
- Explain SVD and its applications.
- Discuss Support Vector Machines.
- Explain Eigen values and vectors.
- Implement XOR function using McCulloch Pitts Model

**Q2 a) In the classification model, the values for the observations are as follows. True Negatives(TN) =300, True Positive(TP)=500, False Positive (FP) = 50, False Negatives (FN)=150. Evaluate the performance of the model by finding values of Accuracy, Precision, Recall and F1-Score. 10**

b) What is the trace of a Matrix. What are its properties? 10

**Q3 a) Diagonalize the Matrix 10**

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}$$

- b) Find Singular Value Decomposition of given matrix and indicate insights about linear transformations conveyed by this method. 10

$$A = \begin{bmatrix} 3 & -5 \\ 4 & 0 \end{bmatrix}$$

**Q4 a) Discuss different activation functions used in Neural Networks. (Formula, Graph and Range). 10**

- b) Implement the **ANDNOT** logic functions using McCulloch Pitts Model. 10

**Q5 a) Implement OR function (logic gate) using single layer perceptron. Assume initial values of weights and learning rate as follows  $w_1=0.6$ ,  $w_2=1.1$  learning rate = 0.5, Threshold = 1 10**

- b) Explain Multilayer perceptron with a neat diagram and its working with flowchart or algorithm. 10

- Q6** a) Why Dimensionality Reduction is very Important step in Machine Learning? Apply PCA on the following data and find the principle component. **10**

<b>X</b>	2	1	0	-1
<b>Y</b>	4	3	1	0.5

- b) Explain Back Propagation Neural Network with flowchart. **10**

---

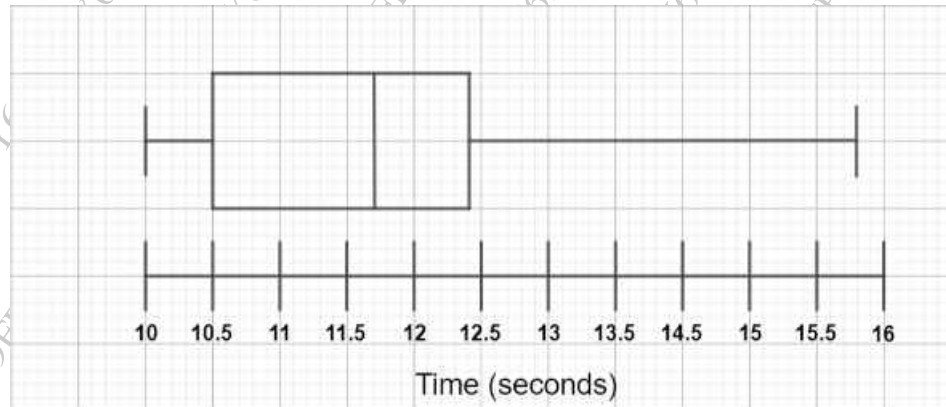
**(3 Hours)**

**Total Marks: 80**

- N.B. 1. Question No. 1 is compulsory.  
 2. Attempt any three questions out of remaining five.  
 3. All questions carry equal marks  
 4. Assume Suitable data, if required and state it clearly.

1 Attempt any four:

- (a) The box plot below was constructed from a collection of times taken to run a 100 m sprint. Using the box plot, determine the range and interquartile range.



- (b) Define Type I and Type II Errors **5**  
 (c) In 800 families with 4 children each. Classify according to given criteria, how many families would you expect to have?  
 2 boys and 2 girls **5**  
 (d) A coin is tossed three times. Calculate the probability of obtaining more heads than tails. **5**  
 (e) Explain the various decomposition models used in time series data. Also, state which decomposition model will be appropriate for the following condition: **5**  
 a) When the seasonal variation is relatively constant over time.

2 (a) You have just taken ownership of a pizza shop. The previous owner told you that you would save money if you bought the mozzarella cheese in a 4.5 pound slab. Each time you purchase a slab of cheese, you weigh it to ensure that you are receiving 72 ounces of cheese. The results of 7 random measurements are 70, 69, 73, 68, 71, 69 and 71 ounces. Are these differences due to chance or is the distributor giving you less cheese than you deserve?

- a. State the hypotheses.  
 b. Calculate the test statistic.  
 c. Would the null hypothesis be rejected at the 10% level? The 5% level? The 1% level?

(b) Elaborate moving average and exponential smoothing techniques?

3 (a) Define sampling and central limit theorem ? Elaborate stratified sampling, judgment sampling, systematic sampling and cluster sampling

- (b) Use multiple regression derive equation for y given x1 and x2. 10

y	x1	x2
-3.7	3	8
3.5	4	5
2.5	5	7
11.5	6	3
5.7	2	1

- 4 a) In a manufacturing unit, four teams of operators were randomly selected and sent to four different facilities for machining techniques training. After the training, the supervisor conducted the exam and recorded the test scores. At 95% confidence level does the scores are same in all four facilities? (Kruskal–Wallis and chi-square table) 10

Facility 1	Facility 2	Facility 3	Facility 4
88	77	71	52
82	76	56	65
86	84	64	68
87	59	51	81

- b) In the context of Multiple linear regression explain what is Over fitting & multicollinearity? 10

- 5 a) Some vehicles pass through a junction on a busy road at an average rate of 300 per hour. 10

- Find out the probability that none passes in a given minute.
- What is the expected number of passing in two minutes?
- Find the probability that this expected number found above actually pass through in a given two-minute period.

- b) Find the simple linear regression equation for the data given below: 10

X	Y
2	21
4	27
6	29
8	64
10	86
12	92

- 6 a) Explain any 3 numerical measures for : 10

- Measures of variability
- Measures of location
- Measures of distribution shape

- b) Difference between 10

- Parametric and non-parametric test
- Discrete and Continuous probability distribution.

Time: 3 hours

Max. Marks: 80

- N.B. (1) Question one is Compulsory.  
(2) Attempt any 3 questions out of the remaining.  
(3) Assume suitable data if required.
- 

Q.1 Answer any 4

(a) Define following terms. 05

- i. Project
- ii. Critical path
- iii. Earned value
- iv. Process
- v. Scope

(b) Compare Agile and traditional software development models 05

(c) Explain cohesion and coupling. 05

(d) What is the golden rule for User interface Design 05

(e) Explain W5HH Principle. 05

Q.2 a) What is the importance of requirement analysis? Explain different Requirement engineering tasks. 10

b) Identify any two Risks in online examination System. Perform risk assessment and prepare RMMM Plan for any one risk. 10

Q.3 a) Prepare SRS for Railway Reservation System. 10

b) Explain following testing strategies 10

- i. Unit Testing
- ii. Integration Testing
- iii. Validation Testing
- iv. System Testing

Q.4 a) Draw CFG and calculate cyclomatic complexity for the given PDL

if(c1 or c2 )

do

s1;

s2;

while (c3)

else

while(c4) s3;

s4;

10

b) Explain Software Design concepts. 10

Q. 5 a) Explain with diagram CMM model

10

b) Draw AON diagram and find the critical path. Find the total float time for each path and list down the critical and non-critical activities. Find the total duration of the project?

10

Activity	Predecessor activity	Duration(days)
A	None	2
B	A	4
C	A	3
D	B	2
E	C	3
F	D	6
G	D	5
H	F,G,E	4

Q.6 Write Short notes on any 4.

20

- i. FTR
- ii. PMBOK Knowledge Areas
- iii. Scrum Model
- iv. SCM Process
- v. Software Reengineering

\*\*\*\*\*

(3 Hours)

(Total Marks: 80)

- N.B.:**
1. Question No. 1 is compulsory.
  2. Answer any three out of the remaining questions.
  3. Assume suitable data if necessary.
  4. Figures to the right indicate full marks.

**Q1. Attempt any FOUR** [20]

- [A] List and explain different key roles for successful data analytics?
- [B] What is Stepwise regression? Explain its types.
- [C] Explain Term Frequency-Inverse Document Frequency (TF-IDF) with a suitable example.
- [D] Difference between Matplotlib and Seaborn library.
- [E] Explain components of time series?

**Q2. Attempt the following** [20]

- [A] Explain different phases in data analytics lifecycle.
- [B] Explain ARIMA model in detail. Also state its Pros and Cons.

**Q3. Attempt the following** [20]

- [A] Explain in detail seven practice areas of text analytics.
- [B] Explain different types of data visualisation in Python programming language.

**Q4. Attempt the following** [20]

- [A] The number of bacterial cells (y) per unit volume in a culture at different hours (x) is given below:

<b>x</b>	0	1	2	3	4	5	6	7	8	9
<b>y</b>	43	46	82	98	123	167	199	213	245	272

Fit lines of regression of y on x and x on y. Also, estimate the number of bacterial cells after 15 hours.

[B] How Exploratory Data Analysis (EDA) is performed in R?

**Q5. Attempt the following**

**[20]**

[A] Enlist and explain the steps of text analysis.

[B] What is Logistic Regression? What are the similarities and differences between linear regression and logistic regression?

**Q6. Write short notes on:**

**[20]**

[A] Box-Jenkins Methodology

[B] Generalized Linear Model

[C] Pandas library

[D] Data import and export in R

\*\*\*\*\*



Time: 3 Hrs

Marks: 80.

Note:

1. Question 1 is compulsory.
2. Attempt any 3 questions out of the remaining questions.

Q1. Attempt any Four.

- a. Define the security Mechanisms and attacks for OSI security architecture. 05
- b. List various types of security that are applied on database. 05
- c. Define web security. Explain the role of cookies. 05
- d. Explain the different types of firewalls and layer in which it operate. 05
- e. Explain the different modes of block ciphers. 05
- f. Differentiate between Digital signature and digital certificate. 05

Q2.

- a. Explain various authentication protocols detail. 10
- b. List and explain in detail security features of operating system. 10

Q3.

- a. Explain the role of SSH. What is the difference between HTTP and HTTPS? 10
- b. List different poly-alphabetic substitution ciphers. Use Hill cipher to encrypt the text "short". The key to be used is "hill". 10

Q4.

- a. Define digital signature.  
P=7 and Q=17 are two prime numbers. If E=5 as public key,  
find the cipher text for 5. Also decrypt the cipher text back to get plain text. 10
- b. Explain MD-5 hash function. Compare with SHA 256. 10

Q5.

- a. Explain Penetration testing in detail. 10
- b. Explain Needham Schroeder Authentication protocol. 10

Q6. Attempt any Four.

- a. List and explain security requirements of a website 10
- b. Explain DES algorithms in detail. 10

\*\*\*\*\*

(3 Hours)

[Total Marks: 80]

N.B.: (1) Question No. 1 is **compulsory**.(2) Solve any **three** questions out of remaining **five**.(3) Figures to **right** indicate **full** marks.(4) Assume suitable **data** where **necessary**.

1.
  - (a) Define hackers, crackers and Phreakers. [5]
  - (b) Discuss in detail the challenges in handling digital evidences. [5]
  - (c) What volatile data can be obtained from the investigation of routers. [5]
  - (d) Describe the staircase digital investigation process model. [5]
2.
  - (a) Explain the role of the following tools in digital forensics: i) netstat ii) psloggedon iii) tcptrace iv) netcat v) cryptcat [10]
  - (b) Explain process of live data collection from Windows system in detail. [10]
3.
  - (a) What is an incident response (IR)? Describe phases of IR process in detail. [10]
  - (b) Describe the various types of network monitoring techniques for data and evidence collection. [10]
4.
  - (a) What is Mobile Forensics? What are different Mobile Forensic tools? Explain [10]
  - (b) What is the significance of command-line utility dd in unix? Enumerate the steps for Simple Forensics imaging with dd, [10]
5.
  - (a) What is network forensics? What is the standard procedure used for network forensics. [10]
  - (b) Explain guidelines for incident report writing. Give one report-writing example. [10]
6. **Write short notes on (any 4)**
  - (a) Incident response methodology. [5]
  - (b) Preventive measures for cybercrime. [5]
  - (c) Digital Forensic Methodology. [5]
  - (d) Role of Windows registry in collecting forensic evidence. [5]
  - (e) E-mail forensic investigation methods. [5]