COMP / VII R-19 14/6/23

Paper / Subject Code: 42171 / MACHINE LEARNING

Duration: 3hrs [Max Marks:80]

N.B. : (1) Question No 1 is Compulsory.
(2) Attempt any three questions out of the remaining five.
(3) All questions carry equal marks.
(4) Assume suitable data, if required and state it clearly.

Q1 Attempt any **four** from following. [20]

A How to choose the right ML algorithm?

B Explain Regression line, Scatter plot, Error in prediction and Best fitting line.

C Explain the concept of feature selection and extraction.

D Explain K-means algorithm.

E Explain the concept of Logistic Regression

Q2 A Explain any five applications of Machine Learning. [10]

B Explain Multivariate Linear regression method [10]

Q3 A Create a decision tree using Gini Index to classify following dataset for profit. [10]

Age	Competition	Type	Profit
old	Yes	software	down
old	No	software	Down
old	No	hardware	Down
mid	Yes	software	Down
mid	Yes	hardware	Down
mid	No	hardware	Up
mid	No	software	Up
new	Yes	software	Up
new	No	hardware	Up
new	no	software	Up

B Find SVD for $A = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}$ [10]

Q4 A Explain the Random Forest algorithm in detail. [10]

B Explain the concept of bagging and boosting [10]

Q5 A Describe Multiclass classification. [10]

B Explain the concept of Expectation Maximization Algorithm. [10]

Q6 Write detailed note on following. **(Any two)** [20]

A Linear Regression

B Linear Discriminant Analysis for Dimension Reduction

C DBSCAN

COMP / VII R-19 16/6/23

Paper / Subject Code: 42172 / BIG DATA ANALYTICS

Time: 3 Hours **Marks: 80**

Note: 1. Question 1 is compulsory
2. Answer any three out of the remaining five questions.
3. Assume any suitable data wherever required and justify the same.

Q1 a) Distinguish between Name node and Data node. **[5]**

b) List and explain the core business drivers behind the NoSQL movement. **[5]**

c) Mention four characteristics of big data. Elaborate these characteristics with respect to social media websites. **[5]**

d) List and explain the different issues and challenges in data stream query processing. **[5]**

Q2 a) What is a key-value store? What are the benefits of using a key-value store? **[10]**

b) Write a map reduce pseudo code to multiply two matrices. Apply map reduce working to perform following matrix multiplication. **[10]**

$$\begin{matrix} 1 & 2 \\ 3 & 4 \end{matrix} \times \begin{matrix} 6 & 7 \\ 8 & 9 \end{matrix}$$

Q3 a) Suppose the stream is S = {2, 1, 6, 1, 5, 9, 2, 3, 5}. Let hash functions h(x) = ax + b mod 16 for some a and b, treat result as a 4-bit binary integer. Show how the Flajolet- Martin algorithm will estimate the number of distinct elements, h(x) = 4x + 1 mod 16. **[10]**

b) Consider the following data frame given below: **[10]**

course	id	class	marks
1	11	1	56
2	12	2	75
3	13	1	48
4	14	2	69
5	15	1	84
6	16	2	53

i. Create a subset of course less than 3 by using [ ] brackets and demonstrate the output.

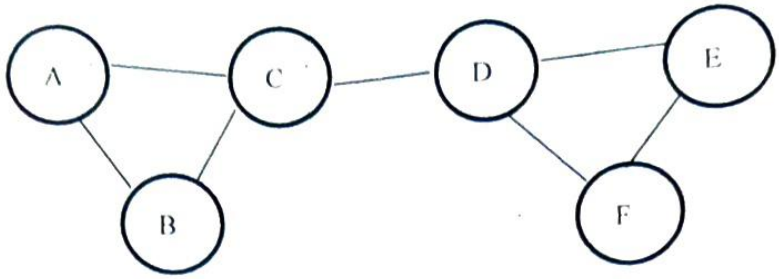
ii. Create a subset where the course column is less than 3 or the class equals to 2 by using subset () function and demonstrate the output.

Q4 a) Explain natural join and grouping and aggregation relational algebraic operation using MapReduce. **[10]**

b) With a neat sketch, explain the architecture of the data-stream management system. **[10]**

19089D1E17245103693DBDBD2501360B

Q5 a) Determine communities for the given social network graph using Girvan-Newman algorithm. 10



b) List and discuss various types of data structures in R. 10

Q6 a) i. The following table shows the number of units of different products sold on different days: 10

Product	Monday	Tuesday	Wednesday	Thursday	Friday
Bread	12	3	5	11	9
Milk	21	27	18	20	15
Cola Cans	10	1	33	6	12
Chocolate bars	6	7	4	13	12
Detergent	5	8	12	20	23

Create five sample numeric vectors from this data.

ii. Name and explain the operators used to form data subsets in R.

b) Define collaborative filtering. Using an example of an e-commerce site like flipkart or amazon describe how it can be used to provide recommendation to users. 10

19089D1E17245103693DBDBD2501360B

For given above corpus,

N: Noun [Martin, Justin, Will, Spot, Pat]

M: Modal verb [can , will]

V:Verb [watch, spot, pat]

Create Transition Matrix & Emission Probability Matrix

Statement is "Justin will spot Will"

Apply Hidden Markov Model and do POS tagging for given statements

b Describe in detail Centering Algorithm for reference resolution.

Q.5 10 marks each

a For a given grammar using CYK or CKY algorithm parse the statement

"The man read this book"

Rules:

S → NP VP	Det → *that* | *this* | *a* | *the*
S → Aux NP VP	Noun → *book* | *flight* | *meal* | *man*
S → VP	Verb → *book* | *include* | *read*
NP → Det NOM	Aux → *does*
NOM → Noun	
NOM → Noun NOM	
VP → Verb	
VP → Verb NP	

b Explain Porter Stemmer algorithm with rules

Q.6 10 marks each

a Explain information retrieval versus Information extraction systems

b Explain Maximum Entropy Model for POS Tagging

Time: 3 hours Max. Marks: 80

N.B. (1) Question No. 1 is compulsory

(2) Assume suitable data if necessary

(3) Attempt any three questions from the remaining questions

Q.1 **Solve any Four out of Five** **5 marks each**

a Explain the challenges of Natural Language processing.

b Explain how N-gram model is used in spelling correction

c Explain three types of referents that complicate the reference resolution problem.

d Explain Machine Translation Approaches used in NLP.

e Explain the various stages of Natural Language processing.

Q.2 **10 marks each**

a What is Word Sense Disambiguation (WSD)? Explain the dictionary based approach to Word Sense Disambiguation.

b Represent output of morphological analysis for Regular verb, Irregular verb, singular noun plural noun Also Explain Role of FST in Morphological Parsing with an example

Q.3 **10 marks each**

a Explain the ambiguities associated at each level with example for Natural Language processing.

b Explain Discourse reference resolution in detail.

Q.4 **10 marks each**

a

<S>	Martin	Justin	can	watch	Will	<E>
<S>	Spot	will	watch	Martin	<E>	
<S>	Will	Justin	spot	Martin	<E>	
<S>	Martin	will	pat	Spot	<E>	

Comp / VII 22/6/23

(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 the following (any 4): **(20)**

a. Distinguish between public, private, and consortium blockchain.

b. Explain the concept of double spending with a suitable example.

c. Compare hot wallets and cold wallets.

d. What is a Merkle tree? Explain the structure of a Merkle tree.

e. Write a program in solidity to find the second largest element in an array.

Q2. Attempt the following:

a. With a suitable diagram, explain the structure of a block header with a list of transactions. **(10)**

b. State and explain different types of cryptocurrencies. **(10)**

Q3. Attempt the following:

a. Describe the concept of state machine replication. How is a smart contract represented as a state machine? **(10)**

b. Explain Hyperledger Fabric v1 architecture. **(10)**

Q4. Attempt the following:

a. Describe the architecture on Ethereum. **(10)**

b. Write a program in solidity to implement single inheritance. **(10)**

Q5. Attempt the following:

a. Explain RAFT consensus mechanism for a private blockchain. **(10)**

b. Explain fixed and dynamic arrays in solidity with suitable examples. **(10)**

Q6. Write short notes on (any 2): **(20)**

a. Ripple

b. UTXO model of Bitcoin

c. Corda

d. Blockchain for DeFi

4BC58CAA29636565BDB38A0440E41A66

(3 Hours) **Total Marks: 80**

Note:

1. **Question No. 1 is compulsory.**
2. Attempt any **THREE** out of the remaining **FIVE** questions.
3. Assume suitable data if necessary.

1	Answer the following (any 4)	(20)
	a) Define the terms: Mitigation, Triage	5
	b) What is Climate Change? What are the effects of Global Warming?	5
	c) What is Disaster Scenario of India?	5
	d) What are Man-made hazards? Also discuss possibilities of chemical spills.	5
	e) Explain Risk and Vulnerability.	5
2	a) Elaborate the guidelines laid down by NDMA for disaster management in India.	10
	b) What are different types of flood? Enlist structural mitigation measures for flood.	10
3	a) Explain disasters related to industries. Discuss the socioeconomic impact of industrial disasters with a case study.	10
	b) Enlist and explain various terms and concepts with respect to earthquakes.	10
4	a) Explain role of various softwares in disaster management.	10
	b) What are different mitigation measures for Tsunami and Cyclones?	10
5	a) What is the role NGO in disaster management? Enlist major NGOs working on disaster management	10
	b) Appraise the role of Geo-informatics in disaster management	10
6	a) Explain various means of raising finance for mitigating and managing disasters	10
	b) Discuss the role of GIS and Remote Sensing in disaster management.	10

****

COMP / VII 26/6/23

Paper / Subject Code: 42184 / Cyber Security Laws

Duration: 3 hours **Max. Marks: 80**

N.B.: 1) Question No.**1** is **compulsory**.
2) Attempt any **THREE** questions out of remaining **FIVE** questions.
3) **Figures** to the **right** indicates **full** marks.
4) Assume suitable data if **necessary**.

Q1	**Attempt any FOUR of the following**	**20**
a	List General guidelines for password policies.	
b	Difference between virus and worm.	
c	How cybercrimes differs from most terrestrial crimes?	
d	What are different Security Risks for Organizations?	
e	What are Mobile Vulnerabilities?	
Q.2		
a	Discuss steps involved in planning of cyberattacks by criminal.	**10**
b	What is vishing attack? How it works? How to protect from vishing attack?	**10**
Q.3		
a	What is e-commerce? Discuss types of e-commerce.	**10**
b	Explain E-contracts and its different types.	**10**
Q.4		
a	What are basic security precautions to be taken to safeguard Laptops and Wireless devices? Explain.	**10**
b	What is Cybercrime? Who are Cybercriminals? Explain.	**10**
Q.5		
a	What is digital evidence? Where one can find it.	**10**
b	What are illegal activities observed in Cyber Cafe? What are safety and security measures while using the computer in Cyber Cafe?	**10**
Q.6	**Write short notes on any FOUR**	**20**
a	Cyberdefamation	
b	HIPAA	
c	Buffer overflow attack	
d	Steganography	
e	DDOS attack	
f	Trojan horse and backdoor	

****