BE(Comp.) sem VII R-19 c-scheme 04/06/25

Time: 03 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) Mention four characteristics of big data and explain in detail. [5]

b) Explain Shuffle & Sort phase and Reducer phase in Map Reduce. [5]

c) Demonstrate how business problems have been successfully solved faster, cheaper and more effectively considering NoSQL Google's Bigtable case study. Also illustrate the business drivers and the findings in it. [5]

d) List down all six constraints that must be satisfied for representing a stream by buckets using DGIM algorithm with examples. [5]

Q2 a) The project manager at XYZ Ltd., Ms. Meera, is responsible for maintaining details of all active projects. She has organized the project information in the following table: [10]

Project Id	Project Name	Budget	Status
1	CRM Implementation	120000	In Progress
2	Cloud Infrastructure	180000	Completed
3	Network Upgrade	60000	Not Started
4	E-Commerce Platform	220000	Completed
5	Data Analytics	90000	In Progress

i) Create a Data frame in R for the above project data and display the output.

ii) Ms. Meera has recently approved 2 new projects and wants to add their information. The new projects are as follows:

Project Id	Project Name	Budget	Status
6	UX Research	160000	Not Started
7	Cloud Integration	190000	Not Started

Update the Data frame to include the new projects and demonstrate the final output.

b) Write a short note on variations of NoSQL architectural patterns. [10]

Q.P. Code:- 87070

Program Code:- 1T00737

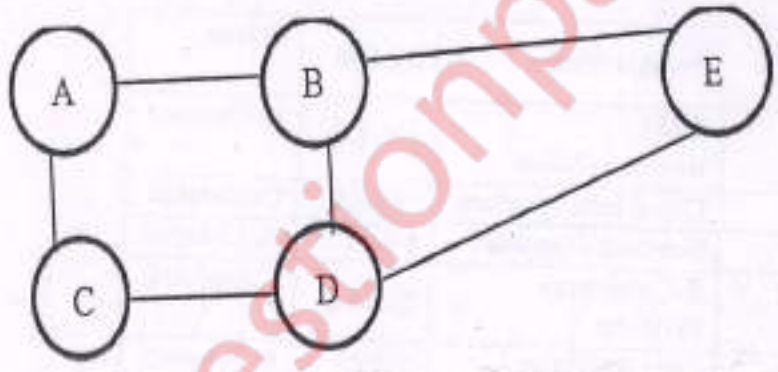
Q3 a) Suppose the stream is S = {10, 12, 8, 15, 6, 9, 14, 7}. Let hash functions $h(x) = 5x + 11 \bmod 32$ for some a and b, treat result as a 5-bit binary integer. Show how the Flajolet- Martin algorithm will estimate the number of distinct elements in this stream. [10]

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

Q4 a) Write a map reduce pseudo code to solve the word count problem. Apply map reduce working on the following document:
"Big data is powerful. Big data drives decisions." [10]

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

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) Describe the components of Hadoop ecosystem with the help of a diagram. [10]

b) What is recommendation system? How is classification algorithm used in recommendation system? [10]

****

06.06.2025

BE(COMP.)/Sem-VII/R-19/C scheme/NLP

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.

1 Attempt any FOUR [20]

a Compare Derivational & Inflectional morphology

b What is the output of Morphological Analysis for Regular Verb, Irregular verb, Singular noun, Plural noun.

c What are the limitations of Hidden Markov Model (HMM) and MaxEnt Model for POS Tagging.

d Explain pre-processing steps generally used in NLP.

e Explain following Syntactic and Semantic Constraints on Co reference
1) Number Agreement 2) Person & Case Agreement

2 a Explain concepts of Bi-gram and n-gram with formula. [10]

For following corpus, apply Bi-gram model

Training Corpus:

<s> I am Sam </s> <s> Sam I am </s> <s> Sam I like </s>

<s> Sam I do like </s> <s> do I like Sam </s>

1. What is the most probable next word predicted by the model for the following word sequences?

(a) <s> Sam . . . (b) <s> Sam I do . . . (c) <s> Sam I am Sam . . .

(d) <s> do I like . . .

2. Which of the following sentences is better, i.e., gets a higher probability with this model?

(e) <s> Sam I do I like </s>

(f) <s> Sam I am </s>.

(g) <s> I do like Sam I am </s>

b Explain different stages of NLP. Also explain generic NLP system. [10]

QP code: 83588

Prog. code: 1T00737

3 a i) Why there is need of word sense disambiguation [10]

ii) Explain Naive Bayes Supervised algorithm for Word sense Disambiguation

b Explain Shift Reduce Parser in NLP with example [10]

4 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>	

[10]

Foe given above corpus, S indicates start of the statement and E indicates end of the statement

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 How Anaphora Resolution is performed with Hobbs and Centering Algorithm [10]

5 a For a given grammar using CYK or CKY algorithm parse the statement [10]

"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 the significance of regular expression in NLP. [10]

6 Write Short Note [20]

a Explain Semi-supervised method (Yarowsky) Unsupervised (Hyperlex) [10]

b Explain Question Answering System with Algorithmic approach [10]

****

BE(COMP)/SEM VII/ B.C / R-19-Cscheme 1 10/05/25.

(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. Explain the concept of an orphaned block.
b. Write a program in solidity to check whether the number is prime or not.
c. Explain the concept of double spending with a suitable example.
d. Differentiate between hot wallet and cold wallet.
e. Explain mining pool and its difficulty.

Q2. Attempt the following:
a. State and explain different types of cryptocurrencies. (10)
b. Explain Hyperledger Fabric v1 Architecture. (10)

Q3. Attempt the following:
a. Write a program in solidity to implement multi-level inheritance. (10)
b. Describe the architecture of Ethereum. (10)

Q4. Attempt the following:
a. Differentiate between PoW, PoS, PoB & PoET. (10)
b. Explain Fallback function in Solidity with an example. (10)

Q5. Attempt the following:
a. Differentiate between public, private and consortium blockchain. (10)
b. Explain types of test network (10)

Q6. Write short notes on (any 2): (20)
a. Ethereum Vitual Machine
b. RAFT consensus algorithm
c. Ripple
d. UTXO model of Bitcoin

****

Q. P. Code :- 83709

Program Code: 1T00737

E(COMP.) / SEM-VII / R-19 C Scheme / MIS / 12/6/25.

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.

1		Attempt any FOUR	[20]
	a	What are the features of Executive Support System?	
	b	Define Information security with an example.	
	c	Define topology and its types with advantages and disadvantages.	
	d	Give an Overview of System Development?	
	e	Describe the tools that augment the traditional SDLC.	
2	a	List down the types of support provided by Accounting IS, finance IS, production/operations management (POM) IS, marketing IS, and human resources IS.	[10]
	b	Analyse the main reasons of Computer Crimes.	[10]
3	a	What do you mean by office automation system.	[10]
	b	Briefly describe the benefits of social commerce to customers.	[10]
4	a	Explain CRM. Describe the different types of CRM with example.	[10]
	b	Write note on mobile-commerce.	[10]
5	a	Describe the privacy issues affected by IT.	[10]
	b	Give examples of B2B and B2C Business Models and contribution of MIS to control these models.	[10]
6	a	What is Decision Support System. Explain the application of DSS.	[10]
	b	What is Cloud Computing? Explain its models?	[10]

****

Q.P. Code :- 82233

Program Code :- IT00737

BE (Comp) Sem VII R'19 c'-scheme 02.06.2025

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. Solve any **four** from following. [20]
a. What are the issues in Machine learning?
b. Explain Regression line, Scatter plot, Error in prediction and Best fitting line.
c. Explain the concept of margin and support vector.
d. Explain following performance metrics along with an example Accuracy, Precision, Recall and F1 score.
e. Explain Logistic Regression

Q2. a. Explain the steps of developing Machine Learning applications. [10]
b. Write short note on Linear Discriminant projection along with an example. [10]

Q3. a. Demonstrate CART method along with an example. [10]
b. Following table shows the midterm and final exam grades obtained for students in a database course. Using linear regression to predict the final exam grade of a student who received 86 in the midterm exam. [10]

Midterm exam (X)	72	50	81	74	94	86	59	83	86	33	88	81
Final exam (Y)	84	53	77	78	90	75	49	79	77	52	74	90

Q4. a. Explain the Random Forest algorithm in detail. [10]
b. Explain the different ways to combine the classifiers. [10]

Q5. a. Describe Multiclass classification. [10]
b. Demonstrate MST algorithm along with example. [10]

Q6. Write detailed note on following. (**Any two**) [20]
a. Performance Metrics for Classification
b. Principal Component Analysis for Dimension Reduction
c. DBSCAN algorithm.

****

QP code
85954

Prog. Code
ITO0737