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.

Q1a) Explain the applications of Natural Language processing.   **5M**

Q1b) Illustrate the concept of tokenization and stemming in Natural Language processing. **5M**

Q1c) Discuss the challenges in part of speech tagging`   .   **5M**

Q1d) Describe the semantic analysis in Natural Language processing.   **5M**

Q2a) Explain inflectional and derivational morphology with an example   **10M**

Q2b) Illustrate the working of Porter stemmer algorithm   **10M**

Q3a) Explain hidden markov model for POS based tagging.   **10M**

Q3b) Demonstrate the concept of conditional Random field in NLP   **10M**

Q4a) Explain the Lesk algorithm for Word Sense Disambiguation.   **10M**

Q4b) Demonstrate lexical semantic analysis using an example   **10M**

Q5a) Illustrate the reference phenomena for solving the pronoun problem   **10M**

Q5b) Explain Anaphora Resolution using Hobbs and Cantering Algorithm   **10M**

Q6a) Demonstrate the working of machine translation systems   **10M**

Q6b) Explain the Information retrieval system   **10M**

———————————

(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 Gas and Ethers in detail.

b. What is the fundamental difference between a hot wallet and a cold wallet in the context of blockchain and cryptocurrency storage?

c. Explain the concept of an orphaned block.

d. Describe how solidity supports multiple inheritance with an example.

e. Compare Bitcoin and Ethereum.

Q2. **Attempt the following:**

a. Differentiate between public, private and consortium blockchain. (10)

b. Differentiate between PoW, PoS, PoB & PoET. (10)

Q3. **Attempt the following:**

a. Explain Merkle Tree with the help of an example. (10)

b. What is mining difficulty and how is it calculated in a proof-of-work? Explain with

an example. (10)

Q4. **Attempt the following:**

a. Write and elaborate a code in solidity to explain visibility and activity qualifiers. (10)

b. Explain view function and pure function in solidity with suitable examples. (10)

Q5. **Attempt the following:**

a. Explain state machine replication with suitable example. (10)

b. Explain RAFT consensus algorithm with a suitable example. (10)

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

a. Role of smart contracts in decentralized finance (DeFi)

b. Ripple

c. Ethereum Virtual Machine (EVM)

d. Mining pool and its methods

****

55193

Duration: - 3 Hours   **Marks: 80 Marks**

NB: - Question 1 is compulsory
       Solve any four questions from Question no. 1.
       Solve any three questions from the remaining.

1 **a.** Describe the differences between information and data in the context of **20 (4x5)**
       Information Retrieval.
   b. Explain the process of Structured text retrieval model.
   c. List the key components of the vector space model.
   d. What are different types of queries in Information Retrieval?
   e. Explain multimedia indexing approach?

2 **a.** What is Information Retrieval Model? Explain the taxonomy of Information **10**
       Retrieval Models
   b. Discuss the relative advantages and disadvantages of the Classic Information **10**
       Retrieval compared to the Alternative Probabilistic Models. Is the probabilistic
       model always superior to the Classic Information Retrieval?

3 **a.** Explain in detail how information retrieval systems utilize user feedback to **10**
       enhance the search results.
   b. Explain the role of a suffix tree in the indexing of Information Retrieval, **10**
       Pattern matching, Structured queries with example

4 **a.** Explain the function and advantages of using parametric and zone indexes in **10**
       enhancing information retrieval.
   b. What is the significance tf-idf weight? Can the tf-idf weight of a term in a **10**
       document exceed 1? Why?

5 **a.** Compare and contrast evaluation of ranked and unranked Retrieval Results? **10**
   b. Explain the process of query processing in distributed information retrieval **10**
       systems.

6   Write short notes on any two   **20**
 a. Inexact top K document retrieval
 b. Automatic local analysis vs Automatic global analysis
 c. Latent Semantic Indexing Model
 d. Flat browsing vs hypertext browsing model.

_____

Duration: 3hrs   **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	List the various components of CBIS.	[5]
	b	What are the features of Executive Support System?	[5]
	c	Define Information security with an example.	[5]
	d	Are Blogs and Wikis different? Justify with application of each.	[5]
	e	How is E-commerce supported by MIS? Give one case to describe same.	[5]
2	a	Highlight the Economic impacts of IS. Give example.	[10]
	b	What do you mean by CAAS, SAAS, IASS ? Give the application of each of these.	[10]
3	a	Contrast to bring out the advantages and disadvantages of Completive environment in an organization.	[10]
	b	Discuss how privacy issue can impact transborder data flows?	[10]
4	a	What are types IS? Explain with example.	[10]
	b	Identify advantages and drawbacks of businesses implementation which uses an enterprise resource planning system.	[10]
5	a	Briefly describe the risks of social computing to business giving suitable examples.	[10]
	b	Is security an ethical responsibility? Justify with a case study.	[10]
6	a	Analyze the key benefits of cloud computing	[10]
	b	How the quality of data is ensured in an organization?	[10]

****

55908   **Page 1 of 1**

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 the following   [20]
A   Explain any five business applications of Machine learning.
B   What is dimensionality reduction? Explain how it can be utilized for classification and clustering task in Machine learning.
C   Explain performance evaluation metrics for binary classification with suitable example.
D   Explain Gini index along with an example.
E   Explain the concept of k fold cross validation.

Q2   A   Write a short note on issues in Machine Learning.   [10]
B   Compare Bagging and Boosting with reference to ensemble learning. Explain how these   [10]
methods help to improve the performance of the machine learning model.

Q3   A   Consider the example below where the mass, y (grams), of a chemical is related to the   [10]
time, x (seconds), for which the chemical reaction has been taking place according to the
table. Find the equation of the regression line. Also explain performance evaluation
measures for regression.

Time, x (seconds)	5	7	12	16	20
Mass, y (grams)	40	120	180	210	240

B   What is Density based clustering? Explain the steps used for clustering task using   [10]
Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm.

Q4   A   Explain Clustering with minimal spanning tree along with example.   [10]
B   Consider the dataset given below with 3 features Color, Wig, Num. Ears and one output   [10]
variable Emotion

Color	G	G	G	B	B	R	R	R	R
Wig	Y	N	N	N	N	N	N	N	Y
Num. Ears	2	2	2	2	2	2	2	2	3
Emotion	S	S	S	S	H	H	H	H	H

i)   Find root node of decision tree using GINI index
ii)   Explain techniques can be used to handle over fitting in decision trees?

Q5   A   Consider the use case of Email spam detection. Identify and explain the suitable machine   [10]
learning technique for this task.

B   Explain the Dimensionality reduction technique Linear Discriminant Analysis and its   [10]
real-world applications.

Q6   A   Define following terminologies with reference to Support vector machine: [10]
Hyper plane, Support Vectors, Hard Margin, Soft Margin, Kernel

B   Explain Ensemble learning algorithm Random Forest and its use cases in real world   [10]
applications.

****

57037   **Page 1 of 1**

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) Explain how big data problems are handled by Hadoop system. [5]

    b) Mention four characteristics of big data and explain in detail. [5]

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

    d) Explain the concept of bloom filter with an example. [5]

Q2 a) What is graph store? Give an example where a graph store can be used to effectively [10] solve a particular business problem.

    b) Write a map reduce pseudo code for word count problem. Illustrate with an example [10] showing all the steps.

Q3 a) Suppose the stream is S = {4, 2, 5 ,9, 1, 6, 3, 7}. Let hash functions h(x) = 3x + [10] 7mod 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.

    b) Describe applications of data visualization. [10]

Q4 a) Explain selection and projection relational algebraic operation using MapReduce. [10]

    b) Explain DGIM algorithm for counting ones in a stream with example. [10]

Q5 a) Determine communities for the given social network graph using Girvan- Newman [10] algorithm.
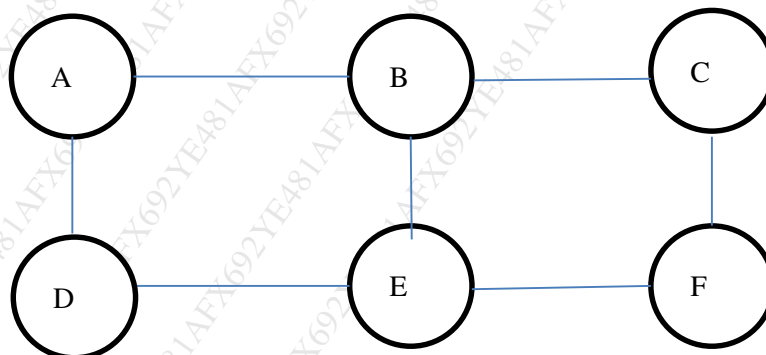
b) Consider the following data frame given below: [10]

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 5 by using [ ] brackets and demonstrate the output.
   ii. Create a subset where the course column is less than 4 or the class equals to 1 by using subset () function and demonstrate the output.

Q6 a) i. Write a script to create a dataset named data1 in R containing the following text. [10]

   Text: 2, 3, 4, 5, 6.7, 7, 8.1, 9

   ii. Explain the various functions provided by R to combine different sets of data.

   b) Describe collaborative filtering in recommendation system. [10]

_____