Learn Research
Innovate

UNITED INSTITUTE OF TECHNOLOGY

(An Autonomous Institution)
(Approved by AICTE | Affiliated to Anna University |
Accredited by NAAC with A+ Grade | Certified by ISO 9001:2015)
Periyanaickenpalayam, Coimbatore – 641020

BACHELOR OF TECHNOLOGY

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

QUESTION BANK

III YEAR

ODD SEMESTER

ACADEMIC YEAR 2024 – 2025

INDEX

HEAD OF THE DEPARTMENT

ACOE **PRINCIPAL** **CHAIRMAN**

AD3501

DEEP LEARNING

UNIT I

DEEP NETWORKS BASICS

Linear Algebra: Scalars — Vectors — Matrices and tensors; Probability Distributions — Gradient-basedOptimization – Machine Learning Basics: Capacity — Overfitting and underfitting –Hyperparametersand validation sets — Estimators — Bias and variance — Stochastic gradient descent — Challenges Motivating deep learning; Deep Networks: Deep feedforward networks; Regularization —Optimization.

Q.No	Question	CO	BTL	Marks
	PART A			
1.	Define Deep learning.	1	1	2
2.	Compare between supervised and unsupervised learning.	1	2	2
3.	What is Stochastic Gradient descent?	1	1	2
4.	Compare between Bias and Variance.	1	2	2
5.	Compare between Regularization and optimization.	1	2	2
6.	Define tensors.	1	1	2
7.	What is Optimization?	1	1	2
8.	Define Deep feedforward networks?	1	1	2
1.	i)Discuss the Bias and Variance. (8) ii)Discuss Overfitting and Underfitting with an example.(8)	1	6	16
2.	Explain the operations of Deep Forward Network with a diagram.	1	5	16
3.	Explain the Challenges motivating in deep learning.	1	5	16
4.	Explain linear algebra.	1	5	16

UNIT II

CONVOLUTIONAL NEURAL NETWORKS

Convolution Operation — Sparse Interactions — Parameter Sharing — Equivariance — Pooling —Convolution Variants: Strided — Tiled — Transposed and dilated convolutions; CNN Learning:Nonlinearity Functions — Loss Functions — Regularization — Optimizers –Gradient Computation.

Q.No	Question	CO	BTL	Marks
	PART A			
1.	What are sparse interactions in a convolutional neural network?	2	1	2
2.	Present an outline of the pooling layer in a convolutional neural network.	2	2	2
3.	Define loss function.	2	1	2
4.	Compare between non linearity functions and loss functions .	2	2	2
5.	Define Gradient Computation.	2	1	2
6.	Compare between Transposed and Dilated convolutions.	2	2	2
7.	What is Pooling?	2	1	2
8.	Define CNN.	2	1	2
	PART B			
1.	Elaborate Convolutional Neural Network Outline Transposed and dilated convolutional with an example	2	6	16
2.	Discuss non linearity in a convolutional neural network? Explain with an example	2	6	16
3.	Examine Gradient computations in detail. Explain with an example.	2	5	16
4.	list types of convolution in detail.	2	4	16

UNIT III

RECURRENT NEURAL NETWORKS

Unfolding Graphs — RNN Design Patterns: Acceptor — Encoder –Transducer; Gradient Computation— Sequence Modeling Conditioned on Contexts — Bidirectional RNN — Sequence to Sequence RNN– Deep Recurrent Networks — Recursive Neural Networks — Long Term Dependencies; Leaky Units:skip connections and dropouts; Gated Architecture: LSTM.

Q.No	Question	CO	BTL	Marks
	PART A			
1.	Define a Recurrent neural network.	3	1	2
2.	What is LSTM?How does it differ from RNN ?	3	1	2
3.	Define Dropouts.	3	1	2
4	What is Transducer?	3	1	2
5.	List out the advantage of introducing depth in Deep recurrent Networks.	3	1	2
6.	Compare between Bi-directional RNN & Uni-directional RNN.	3	2	2
7.	Define Leaky Units.	3	1	2
8.	What is Assess explicit memory.	3	1	2
	PART -B			
1.	What is a Bidirectional Recurrent Neural Network?Explain the Architecture of Bidirectional Recurrent Neural Network with an example.	3	5	16
2.	What is Long Short Term Memory? Compare and contrast LSTM and gated recurrent units .	3	5	16
3.	Explain RNN design patterns in detail.	3	5	16
4.	What is Long Short Term Memory? Compare and contrast LSTM and gated recurrent units .	3	5	16

UNIT IV

MODEL EVALUATION

Performance metrics — Baseline Models — Hyperparameters: Manual Hyperparameter — Automatic Hyperparameter — Grid search — Random search — Debugging strategies.

Q.No	Question	CO	BTL	Marks
	PART A			
1.	lustrate Performance Metrics for Classifications	4	2	2
2.	In machine learning, how each task or problem is divided ?	4	1	2
3.	Define F-Scores.	4	1	2
4.	Define Recall or Sensitivity.	4	1	2
5.	When to Use Accuracy?	4	1	2
6.	Define Confusion Matrix.	4	1	2
7.	What is Random Search?	4	1	2
8.	What is a baseline model in Deep Learning ?	4	1	2
	PART B			
1.	Discuss the various Performance metrics to evaluate a deep learning model with a diagram.	4	6	16
2.	What are Hyperparameters? Discuss the steps to perform hyperparameter tuning.	4	6	16
3.	Explain in detail the Grid Search and Random Search.	4	5	16
4.	Evaluate in detail the different Debugging strategies.	4	5	16

UNIT V

AUTOENCODERS AND GENERATIVE MODELS

Autoencoders: Undercomplete autoencoders — Regularized autoencoders — Stochastic encodersand decoders —Learning with autoencoders; Deep Generative Models: Variational autoencoders –Generative adversarial networks.

Q.No	Question	CO	BTL	Marks
	PART A			
1.	List application Of Generative Adversarial Networks (GANs).	5	1	2
2.	Define Denoising Autoencoder.	5	1	2
3.	Define Generative Adversarial Networks	5	1	2
4.	Outline the architecture of Autoencoder in Deep Learning	5	2	2
5.	Define Discriminator Loss(JD).	5	1	2
6.	What are the different ways to constrain the network in autoencoders?	5	1	2
7.	What are regularized Autoencoders?	5	1	2
8.	What is a Stochastic Encoder ?	5	1	2
	PART B			
1.	Justify your answer that how autoencoders are suitable compared to principal compact analysis for dimensionality reduction .	5	5	16
2.	Explain about Stochastic encoders and decoders	5	5	16
3.	What are Generative adversarial networks?Explain the architecture of GAN with a diagram.	5	5	16
4.	i)Elaborate short notes Sparse Autoencoders.(8) ii)Elaborate Denoising Auto encoders. (8)	5	6	16

----------END--------

CW3551

DATA AND INFORMATION SECURITY

UNIT I

INTRODUCTION

History, What is Information Security?, Critical Characteristics of Information, NSTISSC Security Model, Components of an Information System, Securing the Components, Balancing Security and Access, The SDLC, The Security SDLC

Q.No	Question	CO	BTL	Marks
	PART A			
1.	Define Information Security.	1	1	2
2.	What are the critical characteristics of information?	1	1	2
3.	What is the purpose of cryptography in information security?	1	1	2
4.	What are the three states of information in the McCumber Cube?	1	1	2
5.	What does the NSTISSC security model focus on?	1	1	2
6.	What is a computer virus?	1	1	2
7.	How does networking contribute to an information system?	1	1	2
8.	What is the role of firewalls in network security?	1	1	2
	PART B			
1.	Discuss the importance of information security in organizations and the challenges faced.	1	4	16
2.	Explain in detail the critical characteristics of information security with examples	1	4	16
3.	Design a security policy that ensures both security and usability	1	6	16
4.	Compare and contrast traditional SDLC and Security SDLC.	1	5	16

UNIT II

SECURITY INVESTIGATION

Need for Security, Business Needs, Threats, Attacks, Legal, Ethical and Professional Issues - An Overview of Computer Security - Access Control Matrix, Policy-Security policies, Confidentiality policies, Integrity policies and Hybrid policies

Q.No	Question	CO	BTL	Marks
	PART A			
1.	Why is security needed in an organization?	2	1	2
2.	What are the three main objectives of security?	2	1	2
3.	Mention any two types of cyber threats.	2	1	2
4.	Define SQL injection.	2	1	2
5.	What is a Denial-of-Service (DoS) attack?	2	1	2
6.	What is the importance of security policies in an organization?	2	1	2
7.	What is the difference between a subject and an object in access control?	2	1	2
8.	What is an access control mechanism?	2	1	2
	PART B			
1.	Explain the **need for security** in today's digital world and discuss the risks associated with weak security.	2	4	16
2.	Explain different **types of threats** in cybersecurity with real-world examples.	2	4	16
3.	Explain security policies in detail and discuss their significance in information security.	2	4	16
4.	Compare discretionary access control (DAC), mandatory access control (MAC), and role-based access control (RBAC).	2	4	16

UNIT III

DIGITAL SIGNATURE AND AUTHENTICATION

Digital Signature and Authentication Schemes: Digital signature-Digital Signature Schemes and their Variants- Digital Signature Standards-Authentication: Overview- Requirements Protocols - Applications - Kerberos -X.509 Directory Services

Q.No	Question	CO	BTL	Marks
	PART A			
1.	How does a digital signature provide authentication?	3	1	2
2.	What is the RSA-based digital signature scheme?	3	1	2
3.	Define the ElGamal digital signature scheme	3	1	2
4.	What is the difference between authentication and authorization?	3	2	2
5.	What is multi-factor authentication (MFA)?	3	2	2
6.	What is the purpose of challenge-response authentication?	3	2	2
7.	Define token-based authentication.	3	2	2
8.	What is a one-time password (OTP)?	3	1	2
	PART B			
1.	Explain the working of a digital signature and discuss how it ensures integrity, authentication, and non-repudiation.	3	4	16
2.	Compare RSA, ElGamal, and DSA digital signature schemes	3	4	16
3.	Explain the Kerberos authentication system with a detailed diagram. Compare X.509 with other digital certificate standards and analyze its importance in secure communications.	3	4	16
4.	Compare X.509 with other digital certificate standards and analyze its importance in secure communications.	3	6	16

UNIT IV

E-MAIL AND IP SECURITY

E-mail and IP Security: Electronic mail security: Email Architecture -PGP – Operational Descriptions- Key management- Trust Model- S/MIME.IP Security: Overview- Architecture - ESP, AH Protocols IPSec Modes – Security association - Key management.

Q.No	Question	CO	BTL	Marks
	PART A			
1.	What is the purpose of **PGP (Pretty Good Privacy)?**	4	1	2
2.	Define email encryption.	4	1	2
3.	Name two protocols used in email communication	4	1	2
4.	Define SMTP (Simple Mail Transfer Protocol).	4	1	2
5.	How does PGP ensure message confidentiality?	4	1	2
6.	What is key management in IPSec?	4	1	2
7.	Define Authentication Header (AH) protocol in IPSec.	4	1	2
8.	Define Internet Key Exchange (IKE).	4	1	2
	PART B			
1.	Explain email architecture and describe the functioning of SMTP, IMAP, and POP3.	4	3	16
2.	Compare PGP and S/MIME in terms of encryption, authentication, and usability.	4	4	16
3.	Discuss the role of Security Associations (SAs) in IPSec	4	5	16
4.	Explain the working of S/MIME and how it enhances email security.	4	4	16

UNIT V

WEB SECURITY

Web Security: Requirements- Secure Sockets Layer- Objectives-Layers -SSL secure communication-Protocols - Transport Level Security. Secure Electronic Transaction-Entities DS Verification-SET processing

Q.No	Question	CO	BTL	Marks
	PART A			
1.	List two major threats to web security.	5	1	2
2.	What are the primary functions of SSL?	5	1	2
3.	How does SSL protect against eavesdropping?	5	1	2
4.	What is the Change Cipher Spec Protocol?	5	1	2
5.	What is the role of symmetric encryption in SSL?	5	2	2
6.	What is the SSL Handshake Protocol?	5	2	2
7.	How does the payment gateway function in SET?	5	2	2
8.	What is the role of hashing in digital signatures?	5	2	2
	PART B			
1.	Compare SSL and TLS in terms of security and functionality	5	4	16
2.	Analyze different security mechanisms used to protect web applications.	5	4	16
3.	Describe the SSL handshake process with a step-by-step explanation.	5	5	16
4.	Explain the SET transaction processing with a detailed step-by-step diagram	5	5	16

---------END-------

CS3551

DISTRIBUTED COMPUTING

UNIT I

INTRODUCTION

Introduction: Definition-RelationtoComputerSystemComponents–Motivation–Message-Passing Systems versus Shared Memory Systems – Primitives for Distributed Communication – Synchronous versus Asynchronous Executions – Design Issues and Challenges; A Model of Distributed Computations: A Distributed Program – A Model of Distributed Executions – Models of Communication Networks – Global State of a Distributed System.

Q.No	Question	CO	BTL	Marks
	PART A			
1.	What are the main components of a computer system related to distributed systems?	1	1	2
2.	Differentiate between message-passing systems and shared memory systems	1	2	2
3.	List any two advantages of a distributed system	1	1	2
4.	What are the primitives used for distributed communication?	1	1	2
5.	Differentiate between synchronous and asynchronous execution.	1	2	2
6.	What is the significance of the global state in a distributed system?	1	2	2
7.	State any two challenges in designing a distributed system.	1	1	2
8.	What are the models of communication networks in a distributed system?	1	1	2
	PART B			
1.	Discuss the motivation for using distributed systems. What are the advantages and challenges?	1	5	16
2.	Describe the primitives for distributed communication and their significance in real-world applications.	1	3	16
3.	Differentiate between synchronous and asynchronous execution in distributed systems with examples	1	4	16
4.	Design a simple distributed system architecture for a cloud-based application and discuss its components.	1	6	16

UNIT II

LOGICAL TIME AND GLOBAL STATE

Logical Time: Physical Clock Synchronization: NTP – A Framework for a System of Logical Clocks – Scalar Time – Vector Time; Message Ordering and Group Communication: Message Ordering Paradigms – Asynchronous Execution with Synchronous Communication – Synchronous Program Order on Asynchronous System – Group Communication – Causal Order – Total Order; Global State and Snapshot Recording Algorithms: Introduction–System Modeland Definitions–Snapshot Algorithms for FIFO Channels

Q.No	Question	CO	BTL	Marks
	PART A			
1.	What is clock drift in a distributed system?	1	1	2
2.	Differentiate between physical clocks and logical clocks.	1	2	2
3.	What is the happened-before (→) relation in Lamport timestamps?	1	2	2
4.	What are the types of message ordering in distributed systems?	1	2	2
5.	What are the types of message ordering in distributed systems?	1	2	2
6.	Define causal order in message passing	1	1	2
7.	What is a snapshot algorithm in a distributed system?	1	1	2
8.	Mention any two properties of FIFO channels in distributed systems.	1	2	2
	PART B			
1.	Describe Lamport's logical clock algorithm with an example. How does it ensure event ordering?	1	3	16
2.	Explain asynchronous execution with synchronous communication in distributed systems	1	4	16
3.	Compare and contrast different message ordering paradigms (FIFO, causal order, total order) in distributed systems	1	5	16
4.	Explain the Chandy-Lamport snapshot algorithm for recording global states in FIFO channels.	1	4	16

UNIT III

DISTRIBUTED MUTEX AND DEADLOCK

Distributed Mutual exclusion Algorithms: Introduction–Preliminaries–Lamport's algorithm –Ricart- Agrawala's Algorithm — Token-Based Algorithms – Suzuki-Kasami's Broadcast Algorithm; Deadlock Detection in Distributed Systems: Introduction – System Model – Preliminaries – Models of Deadlocks – Chandy-Misra-Haas Algorithm for the AND model and OR Model

Q.No	Question	CO	BTL	Marks
	PART A			
1.	Define Lamport's algorithm for mutual exclusion.	1	1	2
2.	What is the key difference between Lamport's and Ricart-Agrawala's algorithm?	1	2	2
3.	State two advantages of token-based mutual exclusion algorithms	1	2	2
4.	How does Suzuki-Kasami's algorithm ensure mutual exclusion?	1	1	2
5.	Define deadlock in distributed systems.	1	1	2
6.	What is a wait-for graph (WFG) in deadlock detection?	1	1	2
7.	Differentiate between AND and OR models in deadlock detection	1	2	2
8.	What are the methods to handle deadlocks in distributed systems?	1	2	2
	PART B			
1.	Describe Ricart-Agrawala's algorithm in detail. How does it improve Lamport's algorithm?	1	4	16
2.	Compare and contrast token-based and timestamp-based mutual exclusion algorithms	1	5	16
3.	Discuss different approaches to handling deadlocks in distributed systems	1	5	16
4.	Explain the working of the Chandy-Misra-Haas algorithm for the OR model with an example.	1	4	16

UNIT IV

CONSENSUS AND RECOVERY

Consensus and Agreement Algorithms: Problem Definition–Overview of Results–Agreement in a Failure-Free System(Synchronous and Asynchronous) –Agreement in Synchronous Systems with Failures; Check pointing and Rollback Recovery: Introduction–Background and Definitions–Issues in Failure Recovery–Checkpoint-based Recovery–Coordinated Check pointing Algorithm-Algorithm for Asynchronous Check pointing and Recovery

Q.No	Question	CO	BTL	Marks
	PART A			
1.	What is the difference between synchronous and asynchronous agreement?	1	2	2
2.	Define the role of failure detectors in achieving consensus.	1	2	2
3.	What is Byzantine agreement in distributed systems?	1	1	2
4.	What are the main challenges in reaching consensus in asynchronous systems?	1	1	2
5.	Differentiate between coordinated and uncoordinated checkpointing	1	2	2
6.	What is rollback propagation?	1	1	2
7.	What is a stable storage in the context of rollback recovery?	1	1	2
8.	Define asynchronous checkpointing	1	1	2
	PART B			
1.	Explain the consensus problem in distributed systems. Why is it difficult to achieve in asynchronous systems?	1	4	16
2.	Discuss Byzantine agreement and its solutions in a distributed system.	1	4	16
3.	Describe the challenges and techniques used in checkpoint-based rollback recovery	1	3	16
4.	Discuss the issues in failure recovery and how rollback recovery handles them.	1	5	16

UNIT V

CLOUD COMPUTING

Definition of Cloud Computing – Characteristics of Cloud – Cloud Deployment Models – Cloud Service Models – Driving Factors and Challenges of Cloud – Virtualization – Load Balancing – Scalability and Elasticity – Replication – Monitoring – Cloud Services and Platforms: Compute Services – Storage Services – Application Services

Q.No	Question	CO	BTL	Marks
	PART A			
1.	List any two characteristics of cloud computing	5	2	2
2.	What are the three main cloud service models?	5	1	2
3.	Differentiate between public and private clouds.	5	2	2
4.	List any two driving factors of cloud computing.	5	1	2
5.	Mention two challenges of cloud computing.	5	1	2
6.	What is virtualization in cloud computing?	5	2	2
7.	What is the role of load balancing in cloud computing?	5	2	2
8.	What is replication in cloud computing?	5	1	2
	PART B			
1.	Explain the cloud service models (IaaS, PaaS, SaaS) with suitable examples	5	3	16
2.	Compare and contrast private, public, hybrid, and community cloud models	5	5	16
3.	Explain virtualization in cloud computing and its role in resource optimization	5	3	16
4.	Evaluate different cloud computing platforms and their services (AWS, Azure, Google Cloud).	5	5	16

--------END-------

CCS334
BIG DATA ANALYTICS

UNIT I

UNDERSTANDING BIG DATA

Introduction to big data – convergence of key trends – unstructured data – industry examples of big data – web analytics – big data applications– big data technologies – introduction to Hadoop – open source technologies – cloud and big data – mobile business intelligence – Crowd sourcing analytics – inter and trans firewall analytics

Q.No	Question	CO	BTL	Marks
	PART A			
1.	Define Big data.	1	1	2
2.	List out the key trends in big data.	1	2	2
3.	Define Web analytics.	1	1	2
4.	Define HDFS	1	1	2
5.	State the difference between inter and trans firewall analytics.	1	2	2
6.	Define Crowdsourcing	1	1	2
7.	What is the Internal Firewall?	1	1	2
8.	State the difference between Big data and cloud computing	1	2	2
	PART B			
1.	Generalize the characteristics of big data applications and explain how the big data use cases leverages the benefits and values.	1	5	16
2.	Write brief notes about Web analytics Explain in detail about the big data technologies	1	3	8 8
3.	With a neat sketch explain Apache Hadoop Ecosystem.	1	4	16
4.	Explain about mobile business intelligence with an example Explain in detail about inter and trans firewall analytics.	1	5	16

UNIT II

NOSQL DATA MANAGEMENT

Introduction to NoSQL – aggregate data models – key-value and document data models – relationships – graph databases – schemaless databases – materialized views – distribution models – master-slave replication – consistency - Cassandra – Cassandra data model – Cassandra examples – Cassandra clients.

Q.No	Question	CO	BTL	Marks
	PART A			
1.	What is the primary Advantage of aggregate data models in NoSQL databases?	2	1	2
2.	How do graph databases differ from other NoSQL databases?	2	2	2
3.	What are materialized views in the context of NoSQL databases?	2	2	2
4.	State CAP Theorem.	2	2	2
5.	List the advantages of Sharding.	2	2	2
6.	Why is Cassandra known for its high availability and fault tolerance?	2	2	2
7.	What does it mean for a database to be schemaless?	2	1	2
8.	What are the primary key characteristics in a Cassandra data model?	2	1	2
	PART B			
1	Explain in detail about all the Aggregate Data Model in NOSQL databases	2	5	16
2	Explain in detail about Schema less databases. Explain details about materialized views in NOSQL	2	5	8 8
3.	Elaborate the Distribution model with a neat sketch.	2	6	16
4.	Elaborate the architecture and Data model of Cassandra with a neat sketch.	2	6	16

UNIT III

MAP REDUCE APPLICATIONS

MapReduce workflows – unit tests with MRUnit – test data and local tests – anatomy of MapReduce job run – classic Map-reduce – YARN – failures in classic Map-reduce and YARN – job scheduling – shuffle and sort – task execution – MapReduce types – input formats – output formats

Q.No	Question	CO	BTL	Marks
	PART A			
1.	Define Map Reduce.	3	1	2
2.	List out failures in classic map reduce.	3	2	2
3.	Define the term MR unit	3	1	2
4.	What is fair Scheduler?	3	1	2
5.	What are the limitations of Map Reduce.	3	1	2
6.	Define the term YARN.	3	1	2
7.	Differentiate YARN and Map Reduce.	3	2	2
8.	What is Text Input Format?	3	1	2
	PART B			
1.	Explain in detail about YARN architecture. Write about failures in classic Map-reduce.	3	2	16
2.	How are failures managed in MapReduce and YARN, and what mechanisms ensure the reliability and fault tolerance of MapReduce jobs in the face of node or task failures?	3	4	16
3.	What are the key considerations in job scheduling for MapReduce, and how do fair scheduling and capacity scheduling algorithms work to optimize resource allocation?	3	4	16
4.	Explain detail about anatomy of MapReduce job run.	3	5	16

UNIT IV

BASICS OF HADOOP

Data format – analyzing data with Hadoop – scaling out – Hadoop streaming –Hadoop pipes – design of Hadoop distributed file system (HDFS) – HDFS concepts –Java interface – data flow – Hadoop I/O – data integrity – compression – serialization –Avro – file-based data structures - Cassandra – Hadoop integration.

Q.No	Question	CO	BTL	Marks
	PART A			
1.	Define the term scaling out.	4	1	2
2.	Why do we need Hadoop streaming?	4	2	2
3.	Define the term name node and data node.	4	1	2
4.	Write down the advantages of Hadoop.	4	1	2
5.	What is data locality optimization?	4	2	2
6.	List out types of Hadoop data formats.	4	1	2
7.	Define serialization.	4	1	2
8.	What is Cassandra and its uses?	4	1	2
	PART B			
1.	What is Hadoop streaming? Explain the concept by using the diagram. Explain in detail about the Hadoop I/O system.	4	2	8 8
2.	With a neat sketch explain Hadoop distributed file system Architecture. Explain in detail about serialization in Hadoop.	4	5	8 8
3.	Explain in detail about Avro with an example. Write brief notes on Cassandra and its functions in big data.	4	3	8 8
4.	Explain in detail HDFS concepts in Hadoop and Java interface.	4	5	16

UNIT V

HADOOP RELATED TOOLS

Hbase – data model and implementations – Hbase clients – Hbase examples – praxis. Pig – Grunt – pig data model – Pig Latin – developing and testing Pig Latin scripts. Hive – data types and file formats – HiveQL data definition – HiveQL data manipulation – HiveQL queries.

Q.No	Question	CO	BTL	Marks
	PART A			
1.	Write down the Hive QL queries.	5	2	2
2.	List the features of Hbase.	5	1	2
3.	What is the difference between HBase and Hive?	5	2	2
4.	List down the Hive DDL commands.	5	2	2
5.	Write down the Hive QL queries.	5	2	2
6.	What is Pig, HBase, Hive?	5	1	2
7.	Define Pig Latin.	5	1	2
8.	What is CRUD operation?	5	1	2
	PART B			
1	Explain in detail about Hbase data model Hbase clients with an example	5	2	16
2	Explain in detail about Pig data model. Write a brief about Pig Latin scripts	4	4	8 8
3.	Explain in detail about HiveQL queries.	5	5	16
4.	Explain in detail about Hive data types and file formats.	5	5	16

---------- END --------

CCS335

CLOUD COMPUTING

UNIT I

CLOUD ARCHITECTURE MODELS AND INFRASTRUCTURE

Cloud Architecture: System Models for Distributed and Cloud Computing – NIST Cloud Computing Reference Architecture – Cloud deployment models – Cloud service models; Cloud Infrastructure: Architectural Design of Compute and Storage Clouds – Design Challenges

Q.No	Question	CO	BTL	Marks
	PART A			
1.	What are the five essential characteristics of cloud computing according to NIST?	1	1	2
2.	List the three main service models defined by the NIST Cloud Computing Reference Architecture	1	1	2
3.	Name the four main cloud deployment models.	1	1	2
4.	What is a hybrid cloud, and how is it beneficial?	1	2	2
5.	Which cloud deployment model is most cost-effective for small businesses? Why?	1	2	2
6.	Define scalability in the context of cloud computing.	1	2	2
7.	Why is elasticity important for handling dynamic workloads in cloud systems?	1	2	2
8.	How do latency and bandwidth issues affect cloud computing performance?	1	2	2
	PART B			
1.	How do latency and bandwidth issues affect cloud computing performance?	1	5	16
2.	Discuss the NIST Cloud Computing Reference Architecture and its significance in cloud computing.	1	5	16
3.	Analyze the role of cloud deployment models in meeting diverse business needs, with a focus on public, private, hybrid, and community clouds.	1	5	16
4.	Illustrate the importance of cloud service models in enabling efficient resource utilization and scalability, and discuss their impact on the IT industry.	1	5	16

UNIT II

VIRTUALIZATION BASICS

Virtual Machine Basics – Taxonomy of Virtual Machines – Hypervisor – Key Concepts – Virtualization structure – Implementation levels of virtualization – Virtualization Types: Full Virtualization – Para Virtualization – Hardware Virtualization – Virtualization of CPU, Memory and I/O devices.

Q.No	Question	CO	BTL	Marks
	PART A			
1.	Define a virtual machine.	2	2	2
2.	Define virtualization.	2	2	2
3.	What is full virtualization?	2	2	2
4.	Define para-virtualization.	2	2	2
5.	List the implementation levels of virtualization	2	1	2
6.	What is the difference between Guest OS and Host OS?	2	2	2
7.	Name the two types of hypervisors.	2	1	2
8.	What is a hypervisor?	2	2	2
	PART B			
1.	What is a Hypervisor? Explain its types and the significance of each type in virtualization.	2	2	16
2.	Describe the different implementation levels of virtualization.	2	2	16
3.	Differentiate between Full Virtualization and Para Virtualization. Provide advantages and disadvantages of each.	2	2	16
4.	Explain how virtualization is achieved for CPU, memory, and I/O devices. Illustrate with relevant examples	2	2	16

UNIT III

VIRTUALIZATION INFRASTRUCTURE AND DOCKER

Desktop Virtualization – Network Virtualization – Storage Virtualization – System-level of Operating Virtualization – Application Virtualization – Virtual clusters and Resource Management – Containers vs. Virtual Machines – Introduction to Docker – Docker Components – Docker Container – Docker Images and Repositories.

Q.No	Question	CO	BTL	Marks
	PART A			
1.	What is desktop virtualization?	3	2	2
2.	Define storage virtualization	3	2	2
3.	What is system-level virtualization in operating systems?	3	2	2
4.	What is application virtualization?	3	2	2
5.	What is the main purpose of network virtualization?	3	2	2
6.	What is Docker?	3	2	2
7.	What is a Docker container?	3	2	2
8.	What is a Docker image?	3	2	2
	PART B			
1.	Discuss the concept of Desktop Virtualization.	3	2	16
2.	Explain Storage Virtualization and its role in data management.	3	2	16
3.	Explain the concept of Virtual Clusters and Resource Management in cloud computing	3	2	16
4.	Explain Docker Containers, Images, and Repositories.	3	2	16

UNIT IV

CLOUD DEPLOYMENT ENVIRONMENT

Google App Engine – Amazon AWS – Microsoft Azure; Cloud Software Environments – Eucalyptus – OpenStack.

Q.No	Question	CO	BTL	Marks
	PART A			
1.	Summarize the Service Offerings by AWS	4	2	2
2.	Depict the benefits of OpenStack Compute	4	2	2
3.	What do you mean by open cloud ecosystem	4	2	2
4.	Write the procedure to deploy the application in Google App Engine	4	2	2
5.	What is MS Azure	4	2	2
6.	What are the benefits of using Amazon AWS	4	2	2
7.	Identify the key services provided by MZ Azure	4	1	2
8.	What is OpenStack	4	2	2
	PART B			
1.	Discuss Amazon AWS and MS Azure	4	2	16
2.	Draw and explain the architecture of Eucalyptus	4	4	16
3.	What is Google App Engine? Describe the major building blocks and functional modules of the Google Cloud Platform with a diagram	4	3	16
4.	Breakdown the architecture of OpenStack and explain how its components interact.	4	4	16

UNIT V

CLOUD SECURITY

Virtualization System-Specific Attacks: Guest hopping – VM migration attack – hyperjacking. Data Security and Storage; Identity and Access Management (IAM) - IAM Challenges - IAM Architecture and Practice.

Q.No	Question	CO	BTL	Marks
	PART A			
1.	What is Hyper jacking attack	5	2	2
2.	List out the IAM challenges	5	1	2
3.	What is guest hopping in virtualization security	5	2	2
4.	What are the common virtualization system specific attacks	5	2	2
5.	What is the purpose of identify and access Management	5	2	2
6.	Define VM Migration attack	5	2	2
7.	Describe the role of IAM in cloud security	5	2	2
8.	List out the key components of data security and storage in cloud environments	5	1	2
	PART B			
1.	List the virtualization system specific attacks and explain any two of them	5	4	16
2.	Write a note about guest hopping and VM Migration attacks. Provide real time case studies for the same	5	6	16
3.	Write a detailed note on cloud security	5	2	16
4.	What is Identity and Access Management? Describe its architecture. Depict the procedure to carry out IAM in AWS cloud platform.	5	5	16

--------- END ---------

CCS341

DATA WAREHOUSING

UNIT I

INTRODUCTION TO DATA WAREHOUSE

Data warehouse Introduction - Data warehouse components- operational database Vs data warehouse – Data warehouse Architecture – Three-tier Data Warehouse Architecture - Autonomous Data Warehouse- Autonomous Data Warehouse Vs Snowflake - Modern Data Warehouse

Q.No	Question	CO	BTL	Marks
	PART A			
1.	What is a Data Warehouse?	1	1	2
2.	List the components of a Data Warehouse	1	1	2
3.	What are the layers in Three-Tier Data Warehouse Architecture?	1	1	2
4.	What is the difference between an Operational Database and a Data Warehouse?	1	1	2
5.	Describe the difference between Autonomous Data Warehouse and Snowflake.	1	2	2
6.	How does the Three-Tier Data Warehouse Architecture enhance performance?	1	2	2
7.	Differentiate between Snowflake's architecture and traditional data warehouses.	1	2	2
8.	What is the difference between an Operational Database and a Data Warehouse?	1	1	2
	PART B			
1.	Analyze the key components of a data warehouse and explain how they contribute to efficient data management and decision-making.	1	4	16
2.	Analyze the significance of Three-Tier Data Warehouse Architecture and explain how it improves scalability, performance, and data security.	1	3	16
3.	Evaluate the advantages and challenges of implementing a modern cloud-based data warehouse compared to a traditional on-premises data warehouse.	1	5	16
4.	Evaluate the significance of Snowflake's architecture in modern data warehousing and justify whether it can fully replace traditional data warehouse solutions.	1	5	16

UNIT II

ETL AND OLAP TECHNOLOGY

What is ETL – ETL Vs ELT – Types of Data warehouses - Data warehouse Design and Modeling - Delivery Process - Online Analytical Processing (OLAP) - Characteristics of OLAP - Online Transaction Processing (OLTP) Vs OLAP - OLAP operations- Types of OLAP- ROLAP Vs MOLAP Vs HOLAP.

Q.No	Question	CO	BTL	Marks
	PART A			
1.	What is ETL in Data Warehousing?	2	1	2
2.	List the types of Data Warehouses.	2	1	2
3.	What is Data Warehouse Design and Modeling?	2	1	2
4.	What is OLAP?	2	1	2
5.	What is the difference between OLTP and OLAP?	2	1	2
6.	Differentiate between ETL and ELT in terms of performance and data handling	2	2	2
7.	Differentiate between MOLAP, ROLAP, and HOLAP.	2	2	2
8.	Explain the role of Delivery Process in a Data Warehouse.	2	2	2
	PART B			
1.	Analyze the ETL process in data warehousing and explain how each step (Extract, Transform, Load) impacts the efficiency and accuracy of the data warehouse.	2	4	16
2.	Analyze the different types of data warehouses and explain how they address various business needs.	2	4	16
3.	Compare ROLAP, MOLAP, and HOLAP architectures and analyze their advantages and disadvantages in OLAP systems.	2	4	16
4.	Evaluate the effectiveness of various data warehouse design and modeling approaches, including star schema, snowflake schema, and fact constellation	2	5	16

UNIT III

META DATA, DATA MART AND PARTITION STRATEGY

Meta Data – Categories of Metadata – Role of Metadata – Metadata Repository – Challenges for Meta Management - Data Mart – Need of Data Mart- Cost Effective Data Mart- Designing Data Marts- Cost of Data Marts- Partitioning Strategy – Vertical partition – Normalization – Row Splitting– Horizontal Partition

Q.No	Question	CO	BTL	Marks
	PART A			
1.	What is Metadata in data warehousing?	3	1	2
2.	List the main categories of Metadata.	3	1	2
3.	What is a Data Mart?	3	1	2
4.	List the different partitioning strategies used in data marts.	3	1	2
5.	Define Normalization in the context of data warehouses.	3	2	2
6.	Explain the difference between Business Metadata and Technical Metadata.	3	2	2
7.	Explain how cost-effective data marts can be designed.	3	2	2
8.	What is Vertical Partitioning in data warehousing?	3	1	2
	PART B			
1.	Analyze the categories of metadata and explain how each category plays a role in enhancing the efficiency of data warehouse operations.	3	4	16
2.	Compare and analyze vertical and horizontal partitioning strategies in data warehouses and explain their impact on query performance and data management.	3	4	16
3.	Evaluate the need for cost-effective data marts and suggest strategies for designing low-cost, scalable data marts.	3	5	16
4.	Evaluate the process of designing data marts and assess how schema design (e.g., star schema vs. snowflake schema) impacts data mart performance and usability.	3	5	16

UNIT IV

DIMENSIONAL MODELING AND SCHEMA

Dimensional Modeling- Multi-Dimensional Data Modeling – Data Cube- Star Schema- Snowflake schema- Star Vs Snowflake schema- Fact constellation Schema- Schema Definition - Process Architecture- Types of Data Base Parallelism – Datawarehouse Tools

Q.No	Question	CO	BTL	Marks
	PART A			
1.	What is Dimensional Modeling?	4	1	2
2.	What is a Data Cube in data warehousing?	4	1	2
3.	What is the difference between Star Schema and Snowflake Schema?	4	1	2
4.	List the types of Database Parallelism.	4	1	2
5.	Name some common Data Warehouse Tools.	4	1	2
6.	Differentiate between Intra-query Parallelism and Inter-query Parallelism.	4	2	2
7.	Explain the purpose of OLAP tools in data warehouses.	4	2	2
8.	Why is Snowflake Schema more normalized than Star Schema?	4	2	2
	PART B			
1.	Analyze the importance of Dimensional Modeling in a data warehouse and explain how facts and dimensions improve data analysis.	4	4	16
2.	Compare and analyze Star Schema and Snowflake Schema in terms of structure, performance, and usability.	4	4	16
3.	Analyze the concept of Database Parallelism and explain how different types of parallelism enhance data warehouse performance	4	5	16
4.	Evaluate the role of Fact Constellation Schema in modern data warehouse design and its effectiveness in handling large-scale business data.	4	5	16

UNIT V

SYSTEM & PROCESS MANAGERS

Data Warehousing System Managers: System Configuration Manager- System Scheduling Manager - System Event Manager - System Database Manager - System Backup Recovery Manager - Data Warehousing Process Managers: Load Manager – Warehouse Manager- Query Manager – Tuning – Testing

Q.No	Question	CO	BTL	Marks
	PART A			
1.	What is the role of a System Configuration Manager in a data warehousing system?	5	1	2
2.	Define System Event Manager.	5	1	2
3.	What is the function of a System Backup and Recovery Manager?	5	1	2
4.	What is the role of a Warehouse Manager in data warehousing?	5	1	2
5.	What is the function of a Query Manager?	5	1	2
6.	What is Testing in a data warehousing system?	5	1	2
7.	Differentiate between Load Manager and Warehouse Manager in a data warehouse process.	5	2	2
8.	How does a System Database Manager enhance query efficiency?	5	2	2
	PART B			
1.	Analyze the roles and responsibilities of various Data Warehousing System Managers in ensuring the efficient functioning of a data warehouse.	5	4	16
2.	Compare and analyze the roles of Query Manager, Tuning, and Testing in enhancing query performance and ensuring data warehouse reliability.	5	4	16
3.	Evaluate the importance of System Managers (Configuration, Scheduling, Event, Database, and Backup Recovery Managers) in maintaining the performance, security, and scalability of a data warehouse.	5	5	16
4.	Evaluate the role of process automation through Scheduling Manager and Event Manager in improving data warehouse operational efficiency and reducing manual intervention.	5	5	16

---------END-------