Register No : _____

THE KAVERY ENGINEERING COLLEGE

(An Autonomous Institution)

UG/PG DEGREE END SEMESTER THEORY EXAMINATIONS, APRIL-MAY 2026

Fifth Semester

Computer Science and Engineering

CCS334 - BIG DATA ANALYTICS

(Common To Artificial Intelligence and Data Science, Information Technology)

Regulations - 2021

Duration : 3 Hrs ***Maximum : 100 Marks***

PART – A(ANSWER ALL QUESTIONS) (Marks 10*2=20)

Q.No	*Questions*	*BTL*	*CO*	*Marks*
1.	List the main characteristics of Big Data.	L1	1	2
2.	Differentiate the inter and trans firewall analytics.	L2	1	2
3.	Define graph database.	L1	2	2
4.	Infer Master-Slave replication.	L2	2	2
5.	Define the term MRUnit.	L1	3	2
6.	Compare YARN with the classic MapReduce framework in Hadoop.	L2	3	2
7.	Write the purpose of Hadoop pipes.	L1	4	2
8.	Define serialization.	L1	4	2
9.	Compare RDBMS and Hbase.	L2	5	2
10.	Outline the purpose of HiveQL queries in the Hive ecosystem.	L2	5	2

PART – B (ANSWER ALL QUESTIONS) (Marks 5*16 = 80)

Q.No			*Questions*	*BLT*	*CO*	*Marks*
11.	a	i	Describe the main features of a big data in detail.	L2	1	8
		ii	Explain about industrial example for Big data in detail.	L2	1	8
			Or			
	b	i	Discuss in detail about any three big data technologies.	L2	1	8
		ii	Elaborate the mobile business intelligence with an example.	L2	1	8
12.	a	i	Examine Aggregate Data Model in NOSQL databases.	L4	2	8
		ii	Evaluate the detailed implementation of Schema-less databases.	L4	2	8
			Or			
	b	i	Examine the architectural differences between a general distributed data model and Cassandra's distributed model. Explain how Cassandra addresses challenges in distributed systems with examples."	L4	2	8
		ii	Analyze about the Cassandra clients with an example.	L4	2	8
13.	a	i	Elaborate in detail about Map Reduce workflows.	L2	3	8
		ii	Describe the components involved in the Anatomy of a MapReduce job run.	L2	3	8
			Or			
	b	i	Discuss about shuffling and sorting in Hadoop MapReduce.	L2	3	8
		ii	Classify and explain the MapReduce types in details.	L2	3	8
14.	a		Contrast the architecture of Hadoop distributed file system (HDFS) and explain its function and components.	L4	4	16
			Or			

	b	i	Analyze the components of the Hadoop I/O system and explain how they interact to ensure efficient data read/write operations. Support your answer with diagrams	L4	4	8
		ii	Analyse the impact of seamless Hadoop integration on data processing and analytics.	L4	4	8
15.	a	i	Evaluate the Hbase data model and its implementations with an example.	L5	5	8
		ii	Evaluate the effectiveness of Pig Latin for large-scale data transformation tasks. Assess its strengths and limitations in script development and testing compared to other data processing tools.	L5	5	8
			Or			
	b	i	Estimate the suitability of various Hive data types and file formats for different big data use cases. Discuss their impact on query performance, storage efficiency, and schema evolution.	L5	5	8
		ii	Summarize the salient points on data manipulation in Hive using HiveQL.	L5	5	8