

**SIDDHARTH INSTITUTE OF ENGINEERING & TECHNOLOGY:: PUTTUR
(AUTONOMOUS)**

Siddharth Nagar, Narayanavanam Road – 517583

**QUESTION BANK (DESCRIPTIVE)**Subject with Code: **Big Data Analytics (20CS5024)**Course & Branch: **M.Tech - CSE**Regulation: **R20**Year & Sem: **II M.Tech & I - Sem****UNIT –I
INTRODUCTION TO BIG DATA**

1	List out the main characteristics features of big data architecture with a neat schematic diagram.	[L1][CO1]	[12M]
2	Explain in detail about the challenges of conventional system.	[L2][CO1]	[12M]
3	Illustrate in detail about web data and evolution of analytic scalability.	[L3][CO1]	[12M]
4	How would you show your understanding of the tools, trends and technology in big data.	[L1][CO1]	[12M]
5	a) What are the best practices used in big data analytics?	[L1][CO1]	[6M]
	b) Explain the techniques used in big data analytics.	[L2][CO1]	[6M]
6	a) List out various applications of Big data	[L1][CO1]	[6M]
	b) Discuss about Structure of Big data briefly.	[L2][CO1]	[6M]
7	Explicate in detail about Modern Data Analytic Tools.	[L4][CO1]	[12M]
8	a) Compare the reason for data analysis and data reporting.	[L4][CO1]	[6M]
	b) Examine the risks of Big Data Analytics.	[L1][CO1]	[6M]
9	Discuss the use of big data analytics in business with suitable real world example.	[L2][CO1]	[12M]
10	Explain web data and evolution of analytic scalability in detail.	[L2][CO1]	[12M]

**UNIT –II
HADOOP FRAMEWORK**

1	Generalize the list of tools related to Hadoop framework	[L6][CO2]	[12M]
2	Evaluate the details of distributed file system in Hadoop.	[L5][CO2]	[12M]
3	Discuss the features of Apache Hadoop in detail with diagram as necessary.	[L2][CO2]	[12M]
4	Summarize		
	i) HDFS concepts ii) Hadoop YARN.	[L5][CO2] [L5][CO2]	[6M] [6M]
5	Discuss in detail about Map Reduce distribution system.	[L2][CO2]	[12M]
6	Explain the complexity theory for Map-reduce with reducer size and replication rate.	[L4][CO2]	[12M]

7	Explain the architecture for Map Reduce function with neat diagram.	[L2][CO2]	[12M]
8	Apply the techniques used in Big Data analytics and Vector Multiplication system	[L3][CO2]	[12M]
9	Generalize how the data flow takes places in Map Reduce framework.	[L6][CO2]	[12M]
10	Illustrate Map Reduce framework in detail. Draw the architecture diagram for physical organization of compute nodes.	[L3][CO2]	[12M]

UNIT –III DATA ANALYSIS

1	Explain in detail about regression modeling algorithms in data analytics.	[L2][CO3]	[12M]
2	Apply the proof of any two statistical methods with suitable example.	[L3][CO3]	[12M]
3	Recall the multivariate analysis with Classification Algorithm.	[L1][CO3]	[12M]
4	Experiment the implementation of SVM model with suitable examples.	[L3][CO3]	[12M]
5	Explicate the kernel methodology with advantages.	[L4][CO3]	[12M]
6	a Discuss briefly about Cluster Analysis and types of Data in Cluster Analysis.	[L2][CO4]	[6M]
	b Evaluate Data analysis using R.	[L5][CO4]	[6M]
7	Determine the Association rule mining rule with apriori algorithms in detail.	[L5][CO4]	[12M]
8	Illustrate the K-means algorithms in clustering technique.	[L3][CO4]	[12M]
9	Elaborate the clustering high dimensional data model with diagram.	[L6][CO4]	[12M]
10	Summarize the following		
	i) Partitioning Methods	[L5][CO4]	[6M]
	ii) Hierarchical Methods	[L5][CO4]	[6M]

UNIT –IV MINING DATA STREAMS

1	Describe the stream data model concepts with architecture of stream data.	[L1][CO5]	[12M]
2	Illustrate the Sampling data in a stream model with examples.	[L3][CO5]	[12M]
3	Outline the list of mining data streams concepts.	[L5][CO5]	[12M]
4	Discuss in detail about the mining timing series data model with examples.	[L2][CO5]	[12M]
5	Elaborate the Real time analytics platform system and implementation.	[L6][CO5]	[12M]
6	Compare the list of applications in data stream mining system.	[L2][CO5]	[12M]
7	Choose the data streaming case studies of mining applications.	[L3][CO5]	[12M]
8	Sketch a model for real time sentiment analytics system with suitable example.	[L3][CO5]	[12M]
9	Recall the concepts of sentiment analytics in prediction system.	[L1][CO5]	[12M]
10	Experiment with stock market predictions for data mining streams.	[L3][CO5]	[12M]

UNIT –V
BIG DATA FRAMEWORKS

1		Classify the List of NoSQL data models with diagram in detail.	[L4][CO6]	[12M]
2		Recall the concepts of Hbase model and its implementation system.	[L1][CO6]	[12M]
3	a	Explain Aggregate Data Models in Big data.	[L4][CO6]	[6M]
	b	Demonstrate Hbase Clients and Hadoop Integration.	[L2][CO6]	[6M]
4		Illustrate the Cassandra data model with suitable example.	[L3][CO6]	[12M]
5	a	Explicate in detail Pig architecture in relation to Hadoop ecosystem.	[L4][CO6]	[8M]
	b	Summarize Pig – Grunt briefly.	[L5][CO6]	[4M]
6		Outline the Pig Latin structure and application flow statement.	[L5][CO6]	[12M]
7		Discuss in detail about Cassandra clients with Hadoop ecosystem.	[L2][CO6]	[12M]
8	a	Demonstrate the Hive architecture and data types in file formats.	[L3][CO6]	[8M]
	b	Explain in brief about HiveQL Data Definition.	[L2][CO6]	[4M]
9		Summarize the HiveQL data manipulation in its file types.	[L5][CO6]	[12M]
10		How to select the data retrieval queries in HiveQL queries system.	[L1][CO6]	[12M]

Prepared by:
Dr.R.Elankavi, Associate Professor/CSE/SIETK