SIDDHARTH GROUP OF INSTITUTIONS:: PUTTUR
(AUTONOMOUS)
Siddharth Nagar, Narayanavanam Road – 517583

QUESTION BANK (DESCRIPTIVE)

Subject with Code: BDA (19CS5023)   **Course & Branch**:M.Tech - CSE

Year & Sem: II-M.Tech & I-Sem   **Regulation:** R19

UNIT –I
INTRODUCTION TO BIG DATA

1.	List out the main characteristics of big data architecture with a neat schematic diagram.	[L2][CO1]	**[10M]**
2.	Explain in detail about the challenges of conventional system.	[L5][CO1]	**[10M]**
3.	Illustrate in detail about web data and evolution of analytic scalability.	[L2][CO1]	**[10M]**
4.	How would you show your understanding of the tools, trends and technology in big data.	[L1][CO1]	**[10M]**
5	a. What are the best practices in big data analytics?	[L1][CO1]	**[5M]**
	b. Explain the techniques used in big data analytics.	[L2][CO1]	**[5M]**
6.	Explicate in detail about Modern Data Analytic Tools.	[L4][CO1]	**[10M]**
7.	a. Compare the reason for data analysis and data reporting.	[L1][CO1]	**[5M]**
	b. Examine in detail the trends and technology in big data analytics.	[L4][CO1]	**[5M]**
8.	Discuss the use of big data analytics in business with suitable real world example.	[L6][CO1]	**[10M]**
9.	Explain in detail about modern data analysis tools.	[L4][CO1]	**[10M]**
10.	Explain web data and evolution of analytic scalability in detail.	[L5][CO1]	**[10M]**

UNIT –II
HADOOP FRAMEWORK

1.	Explain generalize the list of tools related to Hadoop framework	[L2][CO2]	**[10M]**
2.	Evaluate the details of distributed file system in Hadoop.	[L5][CO2]	**[10M]**
3.	Discuss the following features of Apache Hadoop in detail with diagram as necessary.	[L6][CO2]	**[10M]**
4.	Discuss in detail about Map Reduce distribution system.	[L6][CO2]	**[10M]**
5.	Explain the complexity theory for Map-reduce with reducer size and replication rate.	[L4][CO2]	**[10M]**
6.	Explain the architecture for Map Reduce function with neat diagram	[L2][CO2]	**[10M]**
7.	Apply the techniques used in Big Data analytics and Vector Multiplication system	[L3][CO2]	**[10M]**
8.	Generalize how the data flow takes places in Map Reduce framework.	[L3][CO2]	**[10M]**
9.	Illustrate Map Reduce framework in detail. Draw the architecture diagram for physical organization of compute nodes.	[L2][CO2]	**[10M]**
10.	Examine in detail the trends and technology in Big Data analytics.	[L4][CO2]	**[10M]**

UNIT –III
DATA ANALYSIS

1.	Explain in detail about regression modeling algorithms in data analytics.	[L5][CO3]	**[10M]**
2.	Apply the proof of any two statistical methods with suitable example.	[L3][CO3]	**[10M]**
3.	Recall the multivariate analysis with Classification Algorithm.	[L1][CO3]	**[10M]**
4.	Experiment the implementation of SVM model with suitable examples.	[L3][CO3]	**[10M]**
5.	Explicate the kernel methodology with advantages.	[L4][CO3]	**[10M]**
6.	Determine the Association rule mining rule with apriori algorithms in detail.	[L5][CO3]	**[10M]**
7.	Illustrate the K-means algorithms in clustering technique.	[L2][CO3]	**[10M]**
8.	Determine the list of model based cluster methods used in data analytics.	[L5][CO3]	**[10M]**
9.	Elaborate the clustering high dimensional data model with diagram.	[L6][CO3]	**[10M]**
10.	Importance of predictive analytics with suitable examples.	[L5][CO3]	**[10M]**

UNIT –IV
MINING DATA STREAMS

1.	Describe the stream data model concepts with architecture of stream data.	[L4][CO4]	**[10M]**
2.	Illustrate the Sampling data in a stream model with examples.	[L4][CO4]	**[10M]**
3.	Outline the list of mining data streams concepts.	[L5][CO4]	**[10M]**
4.	Discuss in detail about the mining timing series data model with samples	[L6][CO4]	**[10M]**
5.	Elaborate the Real time analytics platform system and implementation.	[L6][CO4]	**[10M]**
6.	Compare the list of applications in data stream mining system.	[L2][CO4]	**[10M]**
7.	Choose the data streaming case studies of mining applications.	[L1][CO4]	**[10M]**
8.	Model for real time sentiment analytics system with suitable example.	[L3][CO4]	**[10M]**
9.	Recall the concepts of sentiment analytics in prediction system.	[L1][CO4]	**[10M]**
10.	Experiment with stock market predictions for data mining streams.	[L3][CO4]	**[10M]**

UNIT –V
BIG DATA FRAMEWORKS

1.	Classify the List of NoSQL data models with diagram in detail.	[L4][CO5]	**[10M]**
2.	Recall the concepts of Hbase model and implementation system.	[L1][CO5]	**[10M]**
3.	Illustrate the Cassandra data model with satiable example.	[L4][CO5]	**[10M]**
4.	Explicate in detail Pig architecture in relation to Hadoop ecosystem.	[L5][CO5]	**[10M]**
5.	Developing and testing pig Latin script running model program.	[L3][CO5]	**[10M]**
6.	Outline the Pig Latin structure and application flow statement.	[L5][CO5]	**[10M]**
7.	Discuss in detail about Cassandra clients with Hadoop ecosystem.	[L6][CO5]	**[10M]**
8.	Demonstrate the Hive architecture and data types in file formats.	[L2][CO5]	**[10M]**
9.	Summarize the HiveQL data manipulation in its file types.	[L2][CO5]	**[10M]**
10.	How to select the data retrieval queries in HiveQL queries system.	[L1][CO5]	**[10M]**

Prepared by:
Dr.J.Manikandan, Professor/CSE