

SYNERGY INSTITUTE OF ENGINEERING AND TECHNOLOGY, DHENKANAL

NBA Accredited

Assignment 1

Near NH-55, Banamali Prasad – 759001

Full Marks – 40

Duration 1 week from notification

Roll no-

Registration no-

Name:

Subject with Code : Big Data Integration and Management (BDIM) PCAC2009

Course & Branch : B.Tech : CE, EE, CS

Year & Semester: 2nd & IV

	Section- A Answer All questions	
1	What do you mean by ‘big data Integration’? Explain briefly.	[2marks][CO1][L1]
2	What is Value in Big Data? Explain briefly.	[2marks][CO1][L1]
3	What do you mean by Data explosion? Explain	[2marks][CO1][L2]
4	What do you mean by a NoSQL Model? Explain briefly.	[2marks][CO1][L2]
	Section-B Answer All Questions	
1	What is HDFS? Give a clear and detailed explanation	[6marks][CO1][L2]
2	What are the main challenges in designing a Big Data system? Explain briefly.	[6marks][CO1][L2]
	Section-C Answer All Questions	
1	What do you mean by Processing in Big Data infrastructure? Explain the clearly.	[10 Marks][CO1][L2]
2	What is a Relational Database (RDBMS)? Explain its concepts and application.	[10 Marks][CO1][L2]
	Additional Questions	
1	What is a Data source in Big Data? Give examples	[2marks][CO1][L2]
2	What do you mean by Distributed Systems? Explain briefly	[2marks][CO1][L2]
3	What is the role of Data Models in Big Data storage and retrieval? Explain briefly	[2marks][CO1][L2]

SYNERGY INSTITUTE OF ENGINEERING AND TECHNOLOGY, DHENKANAL

NBA Accredited

Assignment 2

Near NH-55, Banamali Prasad – 759001

Full Marks – 40

Duration 1 week from notification

Roll no-

Registration no-

Name:

Subject with Code : Big Data Integration and Management (BDIM) PCAC2009

Course & Branch : B.Tech : CE, EE, CS

Year & Semester: 2nd & IV

	Section- A Answer All questions	
1	What do you mean by Data Architecture in Big Data?	[2marks][CO2][L1]
2	What is the role of Networking in Big Data infrastructure?	[2marks][CO2][L1]
3	What is HP Vertica? How is it useful in Big Data analytics?	[2marks][CO2][L2]
4	What is Redis? Why is it significant in Big Data?	[2marks][CO2][L2]
	Section-B Answer All Questions	
1	What is Impala, and how does it enhance Big Data querying? Compare it with SparkSQL ? Explain	[6marks][CO2][L2]
2	Explain Big Data Integration & its importance with examples.	[6marks][CO2][L2]
	Section-C Answer All Questions	
1	Describe MongoDB and its role in Big Data processing. How does it integrate with Spark?	[10 Marks][CO2][L2]
2	Explain Neo4j and its significance in handling graph-based Big Data applications	[10 Marks][CO2][L2]
	Additional Questions	
1	Explain the challenges in designing a Big Data system and their possible solutions.	[2marks][CO2][L2]
2	Describe the differences between traditional and big data management systems with suitable examples.	[2marks][CO2][L2]
3	What are the different techniques for retrieving Big Data, and how do they impact performance?	[2marks][CO2][L2]

SYNERGY INSTITUTE OF ENGINEERING AND TECHNOLOGY, DHENKANAL

NBA Accredited

Assignment 3

Near NH-55, Banamali Prasad – 759001

Full Marks – 40

Duration 1 week from notification

Roll no-

Registration no-

Name:

Subject with Code : Big Data Integration and Management (BDIM) PCAC2009

Course & Branch : B.Tech : CE, EE, CS

Year & Semester: 2nd & IV

	Section- A Answer All questions	
1	What do you mean by Machine Learning?	[2marks][CO3][L1]
2	Why is Machine Learning important for Big Data?	[2marks][CO3][L1]
3	What are the different types of Machine Learning?	[2marks][CO3][L2]
4	What is Classification in Machine Learning?	[2marks][CO3][L2]
	Section-B Answer All Questions	
1	What are the challenges of applying Machine Learning to Big Data?	[6marks][CO3][L2]
2	Explain the differences between Supervised, Unsupervised, and Reinforcement Learning.	[6marks][CO3][L2]
	Section-C Answer All Questions	
1	Discuss the importance of Data Exploration and Data Preparation in Machine Learning.	[10 Marks][CO3][L2]
2	Explain Regression in Machine Learning with an example.	[10 Marks][CO3][L2]
	Additional Questions	
1	What are the different types of Clustering techniques? Explain with examples	[2marks][CO3][L2]
2	What are the different techniques for retrieving Big Data, and how do they impact performance	[2marks][CO3][L2]

SYNERGY INSTITUTE OF ENGINEERING AND TECHNOLOGY, DHENKANAL

NBA Accredited

Assignment 4

Near NH-55, Banamali Prasad – 759001

Full Marks – 40

Duration 1 week from notification

Roll no-

Registration no-

Name:

Subject with Code : Big Data Integration and Management (BDIM) PCAC2009

Course & Branch : B.Tech : CE, EE, CS

Year & Semester: 2nd & IV

	Section- A Answer All questions	
1	What do you mean by Clustering in Machine Learning?	[2marks][CO4][L1]
2	What is the role of Feature Selection in Machine Learning?	[2marks][CO4][L1]
3	What is the Decision Tree algorithm?	[2marks][CO4][L2]
4	What is Apache Spark MLlib, and how is it useful for Big Data?	[2marks][CO4][L2]
	Section-B Answer All Questions	
1	How does Random Forest improve classification tasks? Explain with an example.	[6marks][CO4][L2]
2	What are the key steps in evaluating a Machine Learning model?	[6marks][CO4][L2]
	Section-C Answer All Questions	
1	Discuss the role of Distributed Computing in Machine Learning.	[10 Marks][CO4][L2]
2	How does Apache Spark support scalable Machine Learning?	[10 Marks][CO4][L2]
	Additional Questions	
1	Explain the advantages of using Spark MLlib for Machine Learning	[2marks][CO4][L2]
2	Explain how Machine Learning techniques can be used for anomaly detection in Big Data.	[2marks][CO4][L2]