

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]

