

--	--	--	--	--	--	--	--	--	--



GIET UNIVERSITY, GUNUPUR – 765022

Ph.D. (Second Semester) Examinations, November – 2023

WPPECS2011 – Big Data Analytics

(CSE)

Time: 3 hrs

Maximum: 70 Marks

The figures in the right hand margin indicate marks.

Answer ANY FIVE Questions

(14 x 5 = 70 Marks)

	Marks
1.a. Enumerate and provide comprehensive explanations for various applications of big data.	7
b. Investigate the diverse applications and benefits of crowdsourcing analytics in the field of data analytics.	7
2.a. Elaborate on the concept of mobile business intelligence, providing a detailed exploration of its functionalities and applications.	7
b. Elaborate on the operational mechanisms of the MapReduce framework, breaking down the Map, Combiner, Shuffle and Sort, and Reducer phases, using a common example for better understanding.	7
3.a. Illuminate the defining characteristics of a NoSQL database, offering insights into its key features.	7
b. Explain working of following phases of Map Reduce with one common example. (i) Map Phase (ii) Combiner Phase (iii) Shuffle and Sort Phase (iv) Reducer Phase	7
4.a. Provide a brief overview of the Hadoop Ecosystem and its association with file-based data structures, highlighting key components and functionalities.	7
b. Explain the Hadoop Distributed File System (HDFS) commands—'get,' 'cp,' and 'chown'—including their syntax, and present at least one example for each command.	7
5.a. Provide an in-depth explanation of Job Scheduling in MapReduce, detailing the procedures for (i) The Fair Scheduler and (ii) The Capacity Scheduler.	7
b. Write Map Reduce code for counting occurrences of specific words in the input text file(s). Also write the commands to compile and run the code.	7
6.a. Delve into a detailed discussion of Hadoop YARN, with a focus on its role and how it addresses failures in classic MapReduce.	7
b. Discuss the concept of regions in HBase, delving into the storage of big data using HBase and detailing the fundamental principles associated with these aspects.	7
7 a. Offer a comprehensive discussion on Cassandra, including an exploration of Cassandra clients and their functionalities, providing detailed insights into this NoSQL database.	7
b. Describe the working of the Hive architecture, breaking down the operational steps and augmenting the explanation with a diagram to enhance understanding.	7
8 a. Elaborate on the architecture of Apache Pig, elucidating its components and illustrating the system's structure through a diagram.	7
b. List out the applications of big data and explain them in detail.	7

---End of Paper---