

IV B. TECH I SEMESTER REGULAR EXAMINATIONS, NOVEMBER - 2023
INSIGHTS OF BIG DATA

(CSE – ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING)

Time: 3 hours

Max. Marks: 70

Note : Answer ONE question from each unit (5 × 14 = 70 Marks)

~~~~~

UNIT-I

- 1. a) Explain briefly about 3V’s in Big Data Analytics. [6M]
- b) Justify how Big data platform are better than traditional database management systems with examples. [8M]

(OR)

- 2. a) Identify various characteristics of big data analytics. [6M]
- b) Discuss the Classification of Big Data analytics. [8M]

UNIT-II

- 3. a) Explain in detail about YARN and list out its key features. [8M]
- b) Differentiate RDBMS and Hadoop. [6M]

(OR)

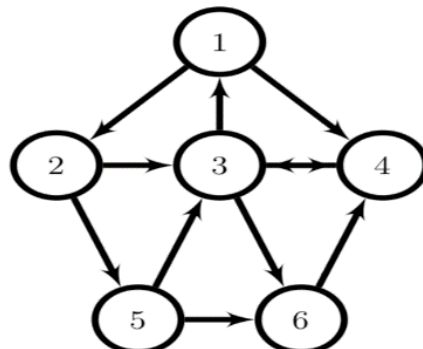
- 4. a) Outline the evolution of Hadoop platform and discuss the role of Google, Apache and Yahoo in each stages of development. [7M]
- b) Explain the Hadoop Eco-system in detail. [7M]

UNIT-III

- 5. a) Build a MapReduce program for counting sequence for following Sentence.**Input:** “Computer science is the study of computation, automation, and information. Computer science spans theoretical disciplines to practical disciplines.” [10M]
- b) Explain in detail about Phases of Map( ) and Reduce ( ) Functions with Example. [4M]

(OR)

- 6. a) Apply the Page Rank algorithm and compute the page ranks of each website for the given below web graph. (Assume the initial ranks of each web page as 1/6). [10M]



- b) Why Combiner is also called as Local reducer? Justify. [4M]

UNIT-IV

7. a) Describe Pig Philosophy. [6M]  
b) Consider the student Relation with the following fields Student [8M]  
(Sid:int, name:chararray, branch:chararray, CGPA:float) and Apply  
the following Pig operations for the above relation.  
i. LOAD  
ii. FILTER  
iii. FOREACH  
iv. DUMP

(OR)

8. a) Discuss the Relational Operations of PIG with example. [7M]  
b) Explain the Anatomy of PIG. [7M]

UNIT-V

9. a) Explain in detail about Hive architecture and discuss its [8M]  
advantages over other components with similar functionalities.  
b) Discuss about the Hive datatypes. [6M]

(OR)

10. a) Create the following using HQL: [8M]  
i. Create a database  
ii. Create a relation Employee (EID, FName, LName, DOB,  
Salary, Address) with DB properties  
iii. Create a static partition on the above relation based on  
salary.  
b) Differentiate PIG and HIVE. [6M]

\* \* \* \* \*