

Q.P.Code 23CS0910

R23

H.T.No.

SIDDHARTH INSTITUTE OF ENGINEERING & TECHNOLOGY:: PUTTUR
(AUTONOMOUS)

B.Tech. III Year II Semester Regular Examinations April-2026
BIG DATA ANALYTICS & AI APPLICATIONS
(Common to CSM & CAI)

Time: 3 Hours

Max. Marks: 70

PART-A

(Answer all the Questions 10 x 2 = 20 Marks)

- | | | | | | |
|---|---|--|-----|----|----|
| 1 | a | List the characteristics of big data. | CO1 | L1 | 2M |
| | b | What is Big Data Analytics? | CO1 | L1 | 2M |
| | c | What is in-memory computing in Spark? | CO2 | L1 | 2M |
| | d | Define Data Frame in Apache Spark. | CO2 | L1 | 2M |
| | e | Define MLlib. Mention any two features of MLlib. | CO3 | L1 | 2M |
| | f | Define accuracy and precision in model evaluation. | CO3 | L1 | 2M |
| | g | What is AIoT in smart cities? | CO4 | L1 | 2M |
| | h | What is Hyper-personalization in retail analytics? | CO4 | L1 | 2M |
| | i | Define Explainable AI (XAI). | CO5 | L1 | 2M |
| | j | State any two ethical issues in Big Data AI. | CO6 | L1 | 2M |

PART-B

(Answer all Five Units 5 x 10 = 50 Marks)

UNIT-I

- | | | | | | |
|---|---|--|-----|----|----|
| 2 | a | Explain Big Data and its importance in modern organizations. | CO1 | L2 | 5M |
| | b | Describe the 5 V's of Big Data with suitable examples. | CO1 | L2 | 5M |

OR

- | | | | | | |
|---|---|--|-----|----|----|
| 3 | a | Explain the role of Zookeeper and Oozie in Hadoop ecosystem. | CO1 | L2 | 5M |
| | b | Discuss about Key value stores and Document oriented Database. | CO1 | L2 | 5M |

UNIT-II

- | | | | | | |
|---|---|--|-----|----|----|
| 4 | a | Illustrate the different deployment modes of Spark on Hadoop. | CO2 | L2 | 5M |
| | b | Explain Resilient Distributed Datasets (RDDs) and their characteristics. | CO2 | L2 | 5M |

OR

- | | | | | | |
|---|---|---|-----|----|----|
| 5 | a | Describe the architecture and working of Apache Hive. | CO2 | L2 | 5M |
| | b | Discuss about Apache Pig and the need for Pig Latin. | CO2 | L2 | 5M |

UNIT-III

- | | | | | | |
|---|--|--|-----|----|-----|
| 6 | | Describe the complete data preprocessing pipeline for big data machine learning with examples. | CO3 | L2 | 10M |
|---|--|--|-----|----|-----|

OR

- | | | | | | |
|---|---|--|-----|----|----|
| 7 | a | Explain the concept of distributed training in machine learning. | CO3 | L2 | 5M |
| | b | Compare synchronous and asynchronous training methods. | CO3 | L5 | 5M |

UNIT-IV

- | | | | | | |
|---|---|--|-----|----|----|
| 8 | a | Illustrate about the key challenges in implementing AI-based predictive maintenance. | CO4 | L3 | 5M |
| | b | Explain how Machine Learning enables real-time fraud detection in banking. | CO4 | L2 | 5M |

OR

- | | | | | | |
|---|--|---|-----|----|-----|
| 9 | | Describe the integration of AI and IoT (AIoT) in smart cities with suitable examples. | CO4 | L2 | 10M |
|---|--|---|-----|----|-----|

UNIT-V

- | | | | | |
|----|---|---|-----|----|
| 10 | a | Explain about how TensorFlow integrates with Apache Spark for large-scale deep learning | CO5 | L2 |
| | b | Describe the role of Spark Data Frames in preprocessing Big Data for deep learning. | CO5 | L2 |

OR

- | | | | | |
|----|---|---|-----|----|
| 11 | a | Analyze about Feature Importance, Attribution, and Visualization approaches in XAI. | CO5 | L4 |
| | b | Explain about how Edge Computing supports low-latency AI applications. | CO6 | L2 |

*** END ***