

# MOHAMMED SATHAK A J COLLEGE OF ENGINEERING

Siruseri IT park, OMR, Chennai - 603103

## LESSON PLAN

### Department of INFORMATION TECHNOLOGY

|                     |                    |                              |           |
|---------------------|--------------------|------------------------------|-----------|
| Name of the Subject | Big Data Analytics | Name of the handling Faculty | V.Saranya |
| Subject Code        | CS8091             | Year / Sem                   | III/VI    |
| Acad Year           | 2021-202           | Batch                        | 2019-2023 |

### Course Objective

- To know the fundamental concepts of big data and analytics.
- To explore tools and practices for working with big data
- To learn about stream computing.

### Course Outcome

- CO1. Identify the big data tools and its analysis techniques.
- CO2. Correlating the data by utilizing clustering and classification methods.
- CO3. Find the different mining algorithms and recommendation systems for large volumes of data
- CO4. Estimate analytics on data streams.
- CO5. Implement the NoSQL databases and management.

| Sl. No. | Topic(s)                                 | T / R* | Period<br>s<br>Requir | Mode of<br>Teaching<br>(BB / PPT / | Blooms<br>Level (L1-<br>L6) | CO  | PO      |
|---------|------------------------------------------|--------|-----------------------|------------------------------------|-----------------------------|-----|---------|
|         |                                          | Book   |                       |                                    |                             |     |         |
| UNIT-I  |                                          |        |                       |                                    |                             |     |         |
| 1       | Evolution of Big data                    | T1     | 1                     | BB                                 | L1                          | CO1 | PO1-PO2 |
| 2       | Best Practices for Big data Analytics    | T1     | 1                     | BB                                 | L1                          | CO1 | PO1     |
| 3       | Understanding Big Data Storage           | T1     | 1                     | BB                                 | L2                          | CO1 | PO1-PO2 |
| 4       | HDFC                                     | T1     | 1                     | BB                                 | L2                          | CO1 | PO2     |
| 5       | Map Reduce Programming Model             | T1     | 1                     | BB                                 | L2                          | CO1 | PO2     |
| 6       | Characteristics of Big Data Applications | T1     | 1                     | BB,PPT                             | L3                          | CO1 | PO2     |
| 7       | Validating                               | T1     | 1                     | BB,PPT                             | L3                          | CO1 | PO2     |
| 8       | Big data characteristics                 | T1     | 1                     | BB,PPT                             | L1                          | CO1 | PO1     |
| 9       | MapReduce and YARN                       | T1     | 1                     | PPT                                | L2                          | CO1 | PO1-PO3 |

Suggested Activity: Assignment / Case Studies / Tuorials/ Quiz / Mini Projects / Model Developed/others Planned if any

### UNIT II CLUSTERING AND CLASSIFICATION

|    |                                          |       |   |        |    |     |         |
|----|------------------------------------------|-------|---|--------|----|-----|---------|
| 10 | Overview of Clustering                   | T1    | 1 | BB     | L1 | CO2 | PO1     |
| 11 | K-means                                  | T1    | 1 | BB     | L1 | CO2 | PO1     |
| 12 | Use Cases                                | T1    | 1 | BB     | L1 | CO2 | PO2     |
| 13 | Determining the Number of Clusters       | T1    | 1 | BB     | L1 | CO2 | PO2     |
| 14 | Diagnostics                              | T1    | 1 | BB     | L2 | CO2 | PO2     |
| 15 | Reasons to Choose and Cautions           | T1,W2 | 1 | BB,PPT | L2 | CO2 | PO2     |
| 16 | Decision Tree Algorithms                 | T1,W2 | 1 | BB,PPT | L2 | CO2 | PO2     |
| 17 | Decision Trees in R                      | T1,R1 | 1 | BB,PPT | L3 | CO2 | PO1-PO3 |
| 18 | Bayes' Theorem – Naïve Bayes Classifier. | T1,R1 | 1 | PPT    | L3 | CO2 | PO1-PO3 |

Suggested Activity: Assignment / Case Studies / Tuorials/ Quiz / Mini Projects / Model Developed/others Planned if any

Evaluation method

### UNIT III ASSOCIATION AND RECOMMENDATION SYSTEM

|    |                                        |       |   |        |    |     |     |
|----|----------------------------------------|-------|---|--------|----|-----|-----|
| 19 | Advanced Analytical Theory and Methods | T1    | 1 | BB     | L1 | CO3 | PO1 |
| 20 | Apriori Algorithm                      | T1    | 1 | BB     | L2 | CO3 | PO1 |
| 21 | Evaluation of Candidate Rules          | T1    | 1 | BB     | L2 | CO3 | PO2 |
| 22 | Applications of Association Rules      | T1    | 1 | BB     | L2 | CO3 | PO2 |
| 23 | Collaborative Recommendation           | T1    | 1 | BB     | L2 | CO3 | PO2 |
| 24 | Content Based Recommendation           | T1    | 1 | BB,PPT | L2 | CO3 | PO2 |
| 25 | Knowledge Based Recommendation         | R1,W1 | 1 | BB,PPT | L2 | CO3 | PO2 |

|    |                                         |       |   |     |    |     |         |
|----|-----------------------------------------|-------|---|-----|----|-----|---------|
| 26 | Hybrid Recommendation Approaches        | R2,W2 | 1 | PPT | L3 | CO3 | PO1-PO3 |
| 27 | Finding Association& finding similarity | T1,W1 | 1 | BB  | L3 | CO3 | PO1-PO3 |

Suggested Activity: Assignment / Case Studies / Tuorials/ Quiz / Mini Projects / Model Developed/others Planned if any

## UNIT IV STREAM MEMORY

|    |                                                 |       |          |    |     |         |
|----|-------------------------------------------------|-------|----------|----|-----|---------|
| 28 | Introduction to Streams Concepts                | T1    | 1 BB     | L1 | CO4 | PO1     |
| 29 | Stream Data Model and Architecture              | T1    | 1 BB     | L1 | CO4 | PO1     |
| 30 | Stream Computing                                | T1    | 1 BB     | L2 | CO4 | PO1     |
| 31 | Filtering Streams                               | T1    | 1 BB     | L2 | CO4 | PO2     |
| 32 | Decaying Window                                 | T1    | 1 BB     | L2 | CO4 | PO2     |
| 33 | Real time Analytics Platform(RTAP) applications | T1    | 1 BB     | L4 | CO4 | PO12    |
| 34 | Stock Market Predictions                        | T1    | 1 BB,PPT | L2 | CO4 | PO1-PO3 |
| 35 | Using Graph Analytics for Big Data              | R1,W2 | 1 BB,PPT | L3 | CO4 | PO1-PO3 |
| 36 | Graph Analytics                                 | T1,W2 | 1 PPT    | L3 | CO4 | PO1-PO3 |

Suggested Activity: Assignment / Case Studies / Tuorials/ Quiz / Mini Projects / Model Developed/others Planned if any

1. What are the various expectations of the Big data platform?
2. various types of graph analytics algorithmic approaches

[illegible]

|                                 |                                                                                           |   |   |   |   |                |   |   |   |   |           |     |   |   |
|---------------------------------|-------------------------------------------------------------------------------------------|---|---|---|---|----------------|---|---|---|---|-----------|-----|---|---|
| CO5                             | 3                                                                                         | 2 | 1 | 0 | 0 | 1              | 0 | 0 | 0 | 0 | 0         | 0   | 3 | 2 |
| Avg                             | 3                                                                                         | 2 | 1 | 0 | 0 | 0.2            | 0 | 0 | 0 | 0 | 0         | 0.2 | 3 | 2 |
| Justification for CO-PO mapping |                                                                                           |   |   |   |   |                |   |   |   |   |           |     |   |   |
| CO1                             | Identify and analysis big data tools and its analysis techniques                          |   |   |   |   |                |   |   |   |   |           |     |   |   |
| CO2                             | Correalting the data by utilizing clustering and classification methods.                  |   |   |   |   |                |   |   |   |   |           |     |   |   |
| CO3                             | Find the different mining algorithms and recommendation systems for large volumes of data |   |   |   |   |                |   |   |   |   |           |     |   |   |
| CO4                             | Estimate analytics on data streams.                                                       |   |   |   |   |                |   |   |   |   |           |     |   |   |
| CO5                             | Implement the NoSQL databases and management.                                             |   |   |   |   |                |   |   |   |   |           |     |   |   |
| 3                               | High level                                                                                |   |   |   | 2 | Moderate level |   |   |   | 1 | Low level |     |   |   |
| Name & Sign of Subject Expert : |                                                                                           |   |   |   |   |                |   |   |   |   |           |     |   |   |
| Head of the Department :IT      |                                                                                           |   |   |   |   |                |   |   |   |   |           |     |   |   |
| Format No :231                  |                                                                                           |   |   |   |   |                |   |   |   |   |           |     |   |   |

MSAJCE