

# Advance Data Analytics (AS5523)

(Credit hours: 3-0-0-0)

## COMPONENT 1

### UNIT 1

Association mining, classification, and Clustering: Revision. Data Streams mining, Social Network Analysis, Graph mining

### UNIT 2

Mining algorithms for large data, mining big data, Hadoop, map-reduce, HDFS, Apache Spark, seminars

## COMPONENT 2

### UNIT 3

Mining sequence pattern in TD, mining, time-series data mining WWW, seminars

### UNIT 4

Advanced Machine Learning: deep learning, probabilistic learning, seminars

## TEXT/REFERENCE BOOKS:

- *Data Mining: Practical Machine Learning Tools and Techniques (2Ed)* by Ian H. Witten and Eibe Frank. Morgan Kaufmann (2005)
- *Mining data with Complex Structures* by Fedja Hadzic, Henry Tan and Tharam S. Dillon. Springer (2011)
- *Modern Information Retrieval* by Yates R. B. and Neto B. R. Pearson Education (2005)
- *Python Data Analytics: Data Wrangling with Pandas, NumPy, and Jupyter (3Ed)* by Wes McKinney. O'Reilly (2022)
- *Learning Spark: Lightning-Fast Data Analytics (2Ed)* by Jules S. Damji, Brooke Wenig, Tathagata Das and Denny Lee. O'Reilly (2020)