GUJARAT TECHNOLOGICAL UNIVERSITY

Program Name: Master of Engineering

Level: PG

Branch: Internet of Things

Subject Code : ME02062041

Subject Name : Data Analytics Techniques

w. e. f. Academic Year:	2024-25
Semester:	2
Category of the Course:	Professional Elective Course

Prerequisite:	Fundamentals of Data Science
Rationale:	The course covers the majority of the existing and evolving open source technology stack for real-time data analytics techniques. The students will get to know about all the real-time solution aspects, from the source to the presentation to persistence. Through this course, the students will be equipped with a clear understanding of how to solve challenges on real-time data. The course cover topics such as how to set up components, basic executions, integrations, advanced use cases, alerts, and monitoring. The student will be exposed to the popular tools used in data analytics today such as Apache Spark, Apache Flink, and Storm. By the end of the course, the students will have a solid understanding of all the aspects of data analytics, and will know how to deploy the solutions in production environments in the best possible manner

Course Outcome:

After Completion of the Course, Student will able to:

No	Course Outcomes	RBT Level*
01	Understand the basic data analytics streaming architecture, service configuration and coordination	UN
02	Apply data flow management techniques/tools in data streaming analysis	AP
03	Demonstrate various data analytics techniques/processing using various tools	AP
04	Analyze of various techniques/systems for storage of streaming data.	AN
05	Evaluate various delivering streaming metrics for various applications.	EL

*RM: Remember, UN: Understand, AP: Apply, AN: Analyze, EL: Evaluate, CR: Create

GUJARAT TECHNOLOGICAL UNIVERSITY

Program Name: Master of Engineering

Level: PG

Branch: Internet of Things

Subject Code : ME02062041

Subject Name : Data Analytics Techniques

Teaching and Examination Scheme:

Teaching Scheme			Total Credits	Assessment Pattern and Marks				Total Marks
				Theory		Practical		
L	T	PR	C	ESE (E)	PA(M)	ESE (V)	PA (I)	
03	00	02	04	70	30	30	20	150

Sr No	Course Content	No of Hours	% weightage
1	**Unit I: Introduction & Designing Real-Time Streaming Architecture:** Sources of streaming data, why streaming data is different, infrastructure and algorithms, Real-time architecture components, features of a real-time architecture, languages for real-time programming, a real-time architecture checklist	7	15
2	**Unit-II: Service Configuration and Coordination:** Motivation for configuration and coordination systems, maintaining distributed state, Apache ZooKeeper	6	15
3	**Unit-III: Data-Flow Management in Streaming Analysis:** Distributed data flows, Apache Kafka: high-throughput distributed messaging, Apache Flume: distributed log collection,	7	15
4	**Unit-IV: Processing Streaming Data:** Distributed streaming data processing, processing data with Storm, Processing data with Samza	9	20
5	**Unit-V: Storing Streaming Data:** Consistent Hashing, NoSQL storage systems, other storage technologies, Choosing a technology, warehousing	7	15
6	**Unit VI: Delivering Streaming Metrics:** Streaming web applications, Visualizing data, mobile streaming applications	9	20
	TOTAL	**45**	**100**

Course Content:

Recent Development in Data Analytics (Study from latest research papers published in Scopus indexed/web of science journals - 04 hours during laboratory hours)

Reference Book:

- Real-Time Analytics Techniques to Analyze and Visualize Streaming Dat, by Byron Ellis, Wiley, 2016
- Practical Real-Time Data Processing and Analytics, by Shilpi Saxena, Saurabh Gupta, Packt, 2017

Suggested Course Practical List:

- The practical work will be carried out based on the content covered during the academic session.

List of Laboratory/Learning Resources Required:

- List of Software: Python, JavaScript, Go, R, MATLAB
- List of Open Source Tools/Simulator: Apache Storm, Apache ZooKeeper, Apache Kafka, Apache Flume, Apache Samza
- List of Useful websites/MOOCs:
    - https://nptel.ac.in/courses/110106064
    - https://nptel.ac.in/courses/106104189

****