Course Code	UDS21S03J	Course Name	DATA ENGINEERING FOR ENTERPRISE	Course Category	S	Skill Enhancement Course	L	T	P	C
							4	0	2	5

Pre-requisite Courses	Nil	Co-requisite Courses	Nil	Progressive Courses	Nil
Course Offering Department	Computer Applications	Data Book / Codes/Standards	Nil		

Course Learning Rationale (CLR):	The purpose of learning this course is to,	Learning			Program Learning Outcomes (PLO)														
		1	2	3	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
CLR-1 :	To introduce the participants to the fundamental concepts of bigdata, its tools and technologies, their working and frameworks																		
CLR-2 :	To introduce the participants to the fundamental concepts of internet of things, a system of interconnected computers, digital machines, devices etc.																		
CLR-3 :	To enumerate all the business challenges involved in the data engineering process.																		
CLR-4 :	To work with the Data mapping, Data Integration, Data Validation, Governance, Quality systems, with their tools and technologies.																		
CLR-5 :	The Primary Objective of this unit is to work with various structured, unstructured, semi-structured, sensor and machine datasets and process with the right tools, technologies, and libraries available.																		
CLR-6 :	To use all the tools and technologies to collect data in real-time and create a data pipeline for increasing the business agility and optimization.																		

Course Learning Outcomes (CLO):	At the end of this course, learners will be able to:	Level of Thinking (Bloom)	Expected Proficiency (%)	Expected Attainment (%)	Fundamental Knowledge	Application of Concepts	Link with Related Disciplines	Procedural Knowledge	Skills in Specialization	Ability to Utilize Knowledge	Skills in Modeling	Analyze, Interpret Data	Investigative Skills	Problem Solving Skills	Communication Skills	Analytical Skills	ICT Skills	Professional Behavior	Life Long Learning
CLO-1 :	Have a firm understanding of Big data from academic an industry perspective. They will have a solid understanding of big data principle, tools, techniques and frameworks.	2	85	80	H	H	H	M	H	H	H	H	H	H	M	H	H	H	H
CLO-2 :	Have a firm understanding of defining the role Big data and IoT plays in building scalable AI Products.	3	85	80	H	H	H	M	H	H	H	H	H	H	M	H	H	H	H
CLO-3 :	Have solid hands-on skills, knowledge and expertise in Data gathering, Data collection, Data Mapping, Data Conversion, Data Quality, Data Validation with domain-specific components	3	85	80	H	H	H	M	H	H	H	H	H	H	M	H	H	H	H
CLO-4 :	Have solid hands-on skills, knowledge and expertise in Collecting data from different enterprise systems and process them efficiently	3	85	80	H	H	H	M	H	H	H	H	H	H	M	H	H	H	H
CLO-5 :	Able to reading, process, and write data from Big Data and IIoT platforms using the right tools and techniques involved	3	85	80	H	H	H	M	H	H	H	H	H	H	M	H	H	H	H
CLO-6 :	Get a firm understanding of the fundamental concepts involved in data integration from multiple data sources and processing data in memory.	3	85	80	H	H	H	M	H	H	H	H	H	H	M	H	H	H	H

Note: All our curriculum, study materials, assignments, quizzes, lab works, and learning resources are personalized and dynamically generated using machine learning models based on the learner's learning ability. Users can review our learning curriculum only through our intelligent learning management platform (iLMSP), and our learning resources and lab infrastructures are available only in the digital form on our cloud infrastructures.

Duration (hour)		18	18	18	18	18
S-1	SLO-1	**Unit 1: Introduction to Big Data**	Business Benefits of Big Data and IIoT together	Data Conversion	Data Validation tools	Data Type and Structure
	SLO-2	Big Data Overview	Big Data tools and support for the industry	Data Quality	Data Validation techniques	Data Source Systems ✓ Oracle ✓ SAP ✓ Twitter ✓ Hadoop ✓ Images ✓ Videos
S-2	SLO-1	Big Data Defined from Academic and Industry Perspective	Approaches to support Big Data and IIoT	Data Validation	Data Governance overview	Data processing tools, techniques and libraries
	SLO-2	Examples of Big Data	**Unit 4: Data Engineering Challenges**	Data Governance	Data Governance in Action	Benefits and Challenges of Data processing from different sources
S-3	SLO-1	Types of Big Data	Data Challenges	Data Pipeline	Data Governance tools	**Unit 10: Working with Big Data and Iiot Platforms - Reading, Processing, Writing, Deleting Data**
	SLO-2	Characteristics of Big Data	Data Type	Building your Data Engineering Architecture	Data Governance techniques	Big Data Sources ✓ Twitter ✓ Facebook ✓ Hadoop ✓ Images ✓ Videos ✓ Reading data from Twitter ✓ Reading data from Facebook ✓ Processing data from Twitter

					✓ Processing data from Facebook ✓ IIoT Data Sources ✓ Sensors ✓ Machine ✓ Industrial Control systems ✓ Website logs ✓ Wearables ✓ Location
S-4	SLO-1	The V's of Big Data	Data format	Unit 6: Enterprise Systems	Data Quality overview	Reading data from Sensors, Website logs, Wearables
	SLO-2	Advantages of Big Data Processing	Data Structure	Data Sources	Data Quality in Action	Processing data from Sensors, Website logs, Wearables
S-5 & S-6	SLO-1	**Lab 1 :** Install a Virtual machine to setup the Hadoop environment and its ecosystems.	**Lab 4 :** **Write a Map Reduce program that mines temperature data. temperature sensors collects data each hour from many locations across the world, gather a large volume of log data**	**Lab 7:** **Connect of Facebook using suitable tools, techniques and libraries, import Customer videos and preprocessing them**	**Lab 10 : Checking with the Data quality**	**Lab 13: Working with Big Data and Iiot Platforms - Reading, Processing, Writing, Deleting Data**
	SLO-2					
S-7	SLO-1	Big Data Technologies	Data Integration from multiple systems	Data Source Systems	Data Quality tools	**Unit 11: Design and Develop a Real-Time Data Collection and Data Streaming Pipeline**
	SLO-2	Big Data Framework	Data Quality Challenges	Relational Databases – Oracle, SQL Server	Data Quality techniques	Data Collection pipeline overview
S-8	SLO-1	**Unit 2: Internet of Things Overview**	Completeness	Datawarehouses – Informatica	**Unit 8: Data Mapping Dictionaries**	Data Collection from different sources
	SLO-2	IoT Overview	Accuracy	DataLake – AWS Redshift	Data Dictionary overview	Data Streaming pipeline overview, working, Architecture, Frameworks
S-9	SLO-1	IoT Defined from Academic and Industry Perspective	Integrity	DataLakehouse - AWS Redshift	Business of Data Dictionary	Data Streaming tools and techniques
	SLO-2	History of IoT	Consistency	**Unit 7: Data Sources - Mapping, Integration, Validation, Governance and Quality**	Data Standards	Business Benefits and challenges of Data Streaming frameworks, Real-time analytics,

S	SLO					
						Data Streaming and real-time analytics
S-10	SLO-1	Working of IoT	Conformity	Data Mapping overview	Data File and Formats	**Unit 12: Design and Develop Data Integration and In-memory Data Processing Pipeline**
	SLO-2	Business Challenges	Timeliness	Data Mapping in Action	Data Templates	Data Integration from multiple sources
S-11 & S-12	SLO-1	**Lab 2 :** **Install Hadoop in the below three operating modes:** ✓ **Standlaone** ✓ **Psudo Distributed** ✓ **Fullly distributed**	**Lab 5 :** **Install Hive and create, alter, and drop tables, views, functions, databases and indexes.**	**Lab 8:** **Connect to Facebook and Twitter using suitable tools, techniques and libraries, Intergrate the collected product and customer data**	**Lab 11: Understanding Data File and Formats and Data Templates**	**Lab 14: Design and Develop Data Integration and In-memory Data Processing Pipeline**
	SLO-2					
S-13	SLO-1	Business Benefits	Data Security Challenges	Data Mapping tools	Naming Conventions	In-Memory Data Processing Overview, pipeline, working, Architecture and frameworks
	SLO-2	IoT Technologies	Data Pipeline Maintenance	Data Mapping techniques	Data Mapping Dictionary Usage	In-Memory Data Processing pipeline tools and techniques
S-14	SLO-1	IoT Framework	Data Governance Issues	Data Integration overview	Documentation	Business Benefits and challenges of In-Memory Data Processing pipeline
	SLO-2	**Unit 3: Big Data and IIoT in AI Products and Solutions**	**Unit 5: Data Engineering Architecture**	Data Integration in Action	Communication	**Unit 13:Working with Sensor and machine data**
S-15	SLO-1	Big Data and IoT best together	Components of Data Engineering Process	Data Integration tools	Application design	IoT devices overview, Sensor, Machine Data
	SLO-2	Big Data and IoT in a nutshell	Data Collection	Data Integration techniques	System Analysis	IoT Sensor Data Processing, architecture, Business Benefits, Business Shallenges Data
S-16	SLO-1	Relation Big Data and IIoT	Data Integration	Data Validation overview	**Unit 9: Working with Different Data Sources - Twitter, Hadoop, Oracle, SAP**	Data Analysis Techniques for IoT Sensor Data Processing, Machine Data Processing
	SLO-2	Role of Big Data in IIoT	Data Mapping	Data Validation in Action	Data Sources	IoT Machine Data Processing architecture, Business Benefits, Business Challenges of Machine Data Processing architecture

S-17 & S-18	SLO-1	Lab 3: Perform the below Hadoop Management Tasks: ✓ Add and Delete Dirctories ✓ Add and Delete files	Lab 6: Connect of twitter using suitable tools, techniques and libraries, import product images and preprocessing them	Lab 9: Working with validation	Lab 12: Working with Different Data Sources - Twitter, Hadoop, Oracle, SAP	Lab 15: Working with Data Analysis Techniques for IoT Sensor Data Processing, Machine Data Processing
	SLO-2					

Learning Resources	1. Data Science and Engineering at Enterprise Scale by Jerome Nilmeier Released April 2019 Publisher(s): O'Reilly Media, Inc	2. Enterprise Big Data Engineering, Analytics, and Management, Martin Atzmueller (University of Kassel, Germany), Samia Oussena (University of West London, UK) and Thomas Roth-Berghofer (University of West London, UK)

Learning Assessment

	Bloom's Level of Thinking	Continuous Learning Assessment (50% weightage)								Final Examination (50% weightage)	
		CLA – 1 (10%)		CLA – 2 (10%)		CLA – 3 (20%)		CLA – 4 (10%) #			
		Theory	Practice	Theory	Practice	Theory	Practice	Theory	Practice	Theory	Practice
Level 1	Remember	20%	15%	20%	15%	20%	15%	20%	15%	20%	15%
	Understand										
Level 2	Apply	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%
	Analyze										
Level 3	Evaluate	10%	15%	10%	15%	10%	15%	10%	15%	10%	15%
	Create										
	Total	100 %		100 %		100 %		100 %		100 %	

CLA – 4 can be from any combination of these: Assignments, Seminars, Tech Talks, Mini-Projects, Case-Studies, Self-Study, MOOCs, Certifications, Conf. Paper etc.,

Course Designers

Experts from Industry	Experts from Higher Technical Institutions	Internal Experts
Mr.Jothi, Periyasamy , Chief AI Architect DeepSphere.AI, CA, USA	Dr.S.Gopinathan, Associate Professor, University of Madras, Chennai	Dr.S.Albert Antony Raj, SRMIST
		Dr. Thilagavathy, SRMIST

Course Code	UMI20S01L	Course Name	My India Project	Course Category	S	Skill Enhancement course	L	T	P	C
							0	0	0	1

Pre-requisite Courses	Nil	Co-requisite Courses	Nil	Progressive Courses	Nil	
Course Offering Department		Computer Applications	Data Book / Codes/Standards	Nil		

(Assessment Method – Fully Internal)

Assessment Tools	Marks
Review – I (Activities)	50
Review – II (Project report and Presentation)	50
Total	100