COURSE CONTENT

CLOUD DATA ENGINEERING								
VIII Semester: CSE (CS)								
Course Code	**Category**	**Hours / Week**			**Credits**	**Maximum Marks**		
		L	**T**	**P**	**C**	**CIA**	**SEE**	**Total**
ACCD31	**Elective**	3	0	0	3	40	60	100
Contact Classes: 48	**Tutorial Classes: Nil**	**Practical Classes: Nil**				**Total Classes: 48**		
Prerequisites: High-Performance Computing								

I. COURSE OVERVIEW:

Cloud data engineering involves designing, building, and managing scalable data pipelines and infrastructure in the cloud. The cloud provides a flexible, cost-efficient environment where data can be ingested, stored, processed, and analyzed at scale. It provides access to large volumes of information while ensuring security and scalability in a cloud environment; essentially, it's about organizing and handling data in the cloud for optimal analysis and insights.

II. COURSE OBJECTIVES:

The students will try to learn:

I The understanding of cloud computing concepts and data access at scale
II The study of cloud service models to store, process, and analyze data
III The workflow management tools to organize and handle data.
IV The applications of data engineering pipelines for key data engineering tasks.

III. COURSE OUTCOMES:

After successful completion of the course, students should be able to:

CO1 Apply Cloud Computing Concepts and Data Engineering Pipelines Techniques for key data engineering tasks.
CO2 Understand various models in cloud.
CO3 Understand several data processing techniques
CO4 Apply data storage concepts to analyse data visualization
CO5 Analyze various Cloud technologies
CO6 Understand basic concepts in cloud computing and data engineering

IV. COURSE CONTENT:

MODULE – I: INTRODUCTION TO CLOUD COMPUTING (10)

Basic Concepts of Cloud Computing Computer Network Basics. Concepts of Distributed Systems. Concepts of Cloud Computing and its Necessity. Cloud Service Providers in use and their Significance. Cloud Infrastructure Cloud Pros and Cons. Cloud Delivery Models. Cloud Deployment Models

MODULE – II: INTRODUCTION TO DATA ENGINEERING (10)

Overview of Data Engineering: Role, importance, and challenges, Data Engineering vs. Data Science vs. Data Analytics, Data Lifecycle: Ingestion, storage, processing, analysis, and visualization.

MODULE – III: DATA STORAGE TECHNOLOGIES (10)

Relational Databases: SQL databases, schema design, normalization, and indexing. NoSQL Databases: Document, key-value, column-family, and graph databases

Data Warehouses: Concepts, star and snowflake schemas, OLAP, and data marts. Distributed File Systems: Hadoop HDFS, Google Cloud Storage, Amazon S3

MODULE - IV: DATA PROCESSING TECHNOLOGIES (09)

Batch Processing: MapReduce, Apache Spark, data parallelism, Stream Processing: Apache Kafka, Apache Flink, event-driven architectures , In-Memory Computing: Redis, Apache Ignite, Memcache

MODULE – V: CLOUD SERVICE MODELS (9)

Characteristics of cloud computing. Cloud Deployment Model: Public cloud, Private cloud, Community cloud, Hybrid cloud. Cloud Service Models: IaaS, PaaS, SaaS. Cloud Economics

V. TEXTBOOKS:

1. T. Erl, R. Puttini, and Z. Mahmood, Cloud Computing: Concepts, Technology & Architecture, Prentice Hall, 1st edition., 2013.
2. J. Reis and M. Housley, Fundamentals of Data Engineering: Plan and Build Robust Data Systems, O'Reilly Media, 1st edition., 2022.

VI. REFERENCE BOOKS:

1. M. Kleppmann, Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable Systems, O'Reilly Media, 1st edition., 2017.
2. I. Foster and D. B. Gannon, Cloud Computing for Science and Engineering, The MIT Press, 1st edition, 2017.

VII, ELECTRONIC REFERENCES:

1. https://www.geeksforgeeks.org
2. https://www.coursera.org
3. https://www.techtarget.com
4. https://www.digitalocean.com
5. https://lumenalta.com

VIII. MATERIALS ONLINE:

1. Course template
2. Tutorial question bank
3. Tech-talk topics
4. Definitions and terminology
5. Assignments
6. Model question paper – I
7. Model question paper – II
8. Lecture notes
9. PowerPoint presentation
10. E-Learning Readiness Videos (ELRV)