# GUJARAT TECHNOLOGICAL UNIVERSITY

**Program Name: Diploma in Engineering**
**Level: Diploma**
**Branch: Computer Engineering / Computer Science and Engineering**
**Subject Code: DI04000031**
**Subject Name: Introduction to No SQL**

| | |
|---|---|
| **W.E. F. Academic Year:** | 2025-26 |
| **Semester:** | 4th |
| **Category of the Course:** | PCC |

| | |
|---|---|
| **Prerequisite:** | • Basic knowledge of database concepts and relational databases<br>• Familiarity with computer programming<br>• Awareness of data management needs in AI/ML |
| **Rationale:** | This course introduces the concepts and applications of NoSQL databases, essential for modern data management. Covering MongoDB, Cassandra, Neo4j, and Redis, it equips students with skills in installation, data modeling, querying, and optimization. With an emphasis on AI-driven and real-time applications, learners will explore practical use cases such as recommendation systems, fraud detection, and caching, preparing them to apply NoSQL solutions effectively in industry and research. |

**Course Outcome:**
After Completion of the Course, Student will be able to:

| No | Course Outcomes | RBT Level |
|---|---|---|
| **01** | Select and justify the appropriate NoSQL database type for a given application scenario. | Apply |
| **02** | Apply MongoDB features for effective data modeling and management. | Apply |
| **03** | Design and implement distributed datasets using Cassandra's data model and CQL. | Apply |
| **04** | Utilize Neo4j to solve relationship-based problems with graph queries. | Apply |
| **05** | Implement Redis data structures for caching and performance optimization. | Apply |

**Revised Bloom's Taxonomy (RBT)*

**Teaching and Examination Scheme:**

| Teaching Scheme (in Hours) | | | Total Credits L+T+ (PR/2) | Assessment Pattern and Marks | | | | Total Marks |
|---|---|---|---|---|---|---|---|---|
| | | | | Theory | | Tutorial / Practical | | |
| L | T | PR | C | ESE (E) | PA (M) | PA (I) | ESE (V) | |
| 0 | 1 | 4 | 3 | 0 | 0 | 20 | 30 | 50 |

**Course Content:**

| Unit No. | Content | No. of Hours | % of Weightage |
|---|---|---|---|
| **1.** | **Introduction to NoSQL:**<br>• Concepts: Why NoSQL? Limitations of RDBMS for modern applications.<br>• The CAP Theorem (only basics with analogy)<br>• Overview of NoSQL types: Document, Key-Value, Wide-Column, Graph.<br>• AI-focused use cases for different NoSQL databases | **8** | **13** |
| **2.** | **Document Databases: MongoDB**<br>• Introduction & Core Concepts (Documents, Collections, BSON)<br>• MongoDB Architecture, Installing MongoDB and using mongosh<br>• Basics and CRUD Operations**:** Creating databases and collections**,** Data types in MongoDB (ObjectId, String, Array, Embedded Document)**,** Insert, Find, Update, Delete operations**,** Query operators and filters ($gt, $lt, $in, $and, $or) Sorting, limiting results, and projections<br>• Data Modeling in MongoDB: Embedded vs. Referenced documents<br>• Indexing and Query Optimization (basic examples)<br>• Aggregation Framework: $match, $group, $sum, $avg<br>• Conceptual overview of replication and sharding (no configuration required) | **18** | **30** |

| | | | |
|---|---|---|---|
| **3.** | **Wide-Column Databases: Cassandra**<br>• Introduction to Cassandra & its uses.<br>• Installing Cassandra and connecting with *cqlsh* (basic demo).<br>• Data Model: Keyspaces, Tables, Columns, Partitions<br>• Primary keys: Partition keys and clustering columns<br>• Cassandra Query Language (CQL): Create, Alter, Drop Keyspaces & Tables, Data types (int, text, timestamp, etc.)<br>• CRUD Operations with CQL: Insert, Update, Delete, Select with filters<br>• Simple queries: WHERE, ORDER BY, LIMIT<br>• Conceptual overview of cluster health and nodetool (awareness level only). | **12** | **20** |
| **4.** | **Graph Databases: Neo4j**<br>• Basics of graph databases & theory (conceptual; focus on applications, not mathematical graph proofs), Installing Neo4j<br>• Cypher Query Language Basics, Graph operations & queries<br>• Graph algorithms (only examples, not implementation): recommender systems, fraud detection,<br>• Real-world AI use cases (LinkedIn, Netflix, Banking) | **12** | **20** |
| **5.** | **Key-Value Databases: Redis**<br>• Overview of Redis and its data structures (Strings, List, Sets, Hashes)<br>• AI-focused use cases of Redis.<br>• Basic Commands and Operations<br>• Redis Transactions and Caching Strategies<br>• Using Redis in Real-World Scenarios<br>• Integration with Other Technologies (Simplified) | **10** | **17** |
| | **Total** | **60** | **100** |

**Suggested Specification Table with Marks (Theory):**

| Distribution of Theory Marks (in %) | | | | | |
|---|---|---|---|---|---|
| **R Level** | **U Level** | **A Level** | **N Level** | **E Level** | **C Level** |
| **0** | **0** | **0** | **0** | **0** | **0** |

*Where R: Remember; U: Understanding; A: Application, N: Analyze and E: Evaluate C: Create (as per Revised Bloom's Taxonomy)*

**References/Suggested Learning Resources:**

**(a) Books:**

1. MongoDB: The Definitive Guide, Shannon Bradshaw, Eoin Brazil, Kristina Chodorow, O'Reilly Media, 2019, 978-1-491-95446-1, 978-1-491-95441-6, 978-1-491-95442-3, 978-1-491-95443-0
2. NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence, Pramod J. Sadalage & Martin Fowler, Addison-Wesley Professional, 2013, 978-0-321-82662-6, 0-321-82662-0, 978-0-133-03612-1, 0-133-03612-X
3. Cassandra: The Definitive Guide: Distributed Data at Web Scale, Jeff Carpenter & Eben Hewitt, O'Reilly Media, 2020, 978-1-098-11516-6, 1-098-11516-3
4. Graph Databases: New Opportunities for Connected Data, Ian Robinson, Jim Webber, Emil Eifrem, O'Reilly Media, 2015, 978-1-491-93089-2, 1491930896
5. Redis in Action, Josiah L. Carlson, Manning Publications, 2013, 9781617290855, 1617290858

**(b) Open source software and website:**

1. **IBM – What Is NoSQL?: Overview of NoSQL databases and their benefits.**
   https://www.ibm.com/topics/nosql-databases
2. **Coursera – Introduction to NoSQL: Beginner-friendly lecture on NoSQL concepts.**
   https://www.coursera.org/lecture/nosql-databases/introduction-to-nosql-VdRNp
3. **GeeksforGeeks – Introduction to NoSQL: Explains NoSQL basics with examples.**
   https://www.geeksforgeeks.org/introduction-to-nosql/
4. **JavaTpoint – NoSQL Databases: Tutorials on various NoSQL database types.**
   https://www.javatpoint.com/nosql-databases
5. **DataCamp – NoSQL Tutorial Using MongoDB: Hands-on tutorial focusing on MongoDB.**
   https://www.datacamp.com/tutorial/nosql-tutorial
6. **Google Cloud – What Is NoSQL?: Explains NoSQL types and real-world use cases.**
   https://cloud.google.com/discover/what-is-nosql
7. **Dataquest – Introduction to NoSQL: Covers four main NoSQL types with examples.**
   https://www.dataquest.io/blog/introduction-to-nosql/
8. **Couchbase – NoSQL Explained: Insights into NoSQL databases and their structure.**

https://www.couchbase.com/resources/why-nosql/

9. **W3Schools – MongoDB Tutorial: Practical guide to MongoDB with exercises.**
https://www.w3schools.com/mongodb/
10. **Redis – What Is NoSQL?: Focus on NoSQL concepts using Redis.**
https://redis.io/nosql/what-is-nosql/
11. **Instaclustr – "NoSQL tutorial: Taking your first steps in the NoSQL world"** — A general NoSQL tutorial and guide
https://www.instaclustr.com/education/nosql-database/nosql-tutorial-taking-your-first-steps-in-the-nosql-world/

**Suggested Course Practical List:**

1. Introduction and Types of NoSQL Databases
   Study and compare different types of NoSQL databases—Document, Key-Value, Wide-Column, and Graph. Identify their unique characteristics, advantages, and application areas.
2. Introduction and Installation of MongoDB
   Install MongoDB Community Edition and MongoDB Shell (mongosh). Configure the environment and verify installation through version checks and database connectivity.
3. Basic CRUD Operations with MongoDB
   Create and manage databases and collections. Perform Insert, Find, Update, and Delete operations using query filters, sorting, projection, and logical operators.
4. Introduction and Setup of Cassandra
   Install Apache Cassandra and connect using cqlsh. Explore system keyspaces, define keyspaces, and understand the role of partition keys and clustering columns.
5. Data Modeling and Simple Queries with Cassandra
   Create tables using appropriate primary keys. Perform CRUD operations and queries using WHERE, ORDER BY, and LIMIT. Demonstrate data retrieval from distributed partitions.
6. Real-Time AI Dataset Implementation using Cassandra
   Import an AI-related dataset (e.g., sensor data or recommendation dataset) into Cassandra. Perform data distribution and simple analytical queries to simulate real-time applications.
7. Introduction to Neo4j Graph Databases
   Install and configure Neo4j Desktop or Aura. Explore graph components—nodes, relationships, and properties—using the Neo4j Browser interface.
8. Basic Graph Queries and Implementations with Neo4j
   Create nodes and relationships. Perform MATCH, WHERE, and RETURN queries. Demonstrate AI-driven use cases such as friend recommendations or fraud detection.
9. Redis Basics: Introduction and Key-Value Operations

Install Redis and explore basic data types (Strings, Lists, Sets, Hashes). Execute commands for storing, retrieving, and managing key-value pairs.

**List of Laboratory/Learning Resources Required:**

1. **Computer System Configuration**: Desktop or Laptop with minimum specifications:
   - Processor: Intel i3 (or equivalent) and above
   - RAM: 8 GB (recommended for database and virtualization tasks)
   - Storage: Minimum 256 GB SSD / HDD
   - Operating System: Windows 10/11, Ubuntu Linux, or macOS
2. **Software Requirements**
   - **MongoDB** – NoSQL Document Database
     - Download: https://www.mongodb.com/
     - Tools: MongoDB Community Server, MongoDB Shell (mongosh), MongoDB Compass
   - **Apache Cassandra** – NoSQL Wide-Column Store
     - Download: https://cassandra.apache.org/
     - Tools: CQLSH (Cassandra Query Language Shell)
   - **Redis** – In-Memory Key-Value Data Store
     - Download: https://redis.io/
     - Tools: Redis CLI and RedisInsight (GUI)
   - **Neo4j** – Graph Database
     - Download: https://neo4j.com/
     - Tools: Neo4j Desktop / Neo4j Browser / Cypher Query Interface

**Suggested Project List:**

1. **Online Bookstore Inventory System (MongoDB)**
   Design a document-based system to manage book details, stock levels, authors, and pricing using MongoDB collections and aggregation.
2. **Blog Management Platform (MongoDB)**
   Build a content management system to store posts, comments, and user data with indexing and search optimization using MongoDB.
3. **E-commerce Product Catalog (MongoDB)**
   Develop a product catalog with nested documents for categories, attributes, and user reviews to demonstrate embedded vs. referenced data modeling.
4. **Session Management System (Redis)**
   Implement user session storage and retrieval using Redis for login tracking, user activity, and fast key-value operations.

5. **Real-Time Leaderboard (Redis)**
   Create a leaderboard system for online games or contests using Redis sorted sets to maintain real-time rankings and updates.
6. **IoT Sensor Data Storage (Cassandra)**
   Store and query large volumes of time-stamped IoT sensor data using Cassandra for scalable and distributed data handling.
7. **Time-Series Data Analytics (Cassandra)**
   Perform analytics on time-series datasets such as temperature, energy usage, or stock prices using partitioning and clustering keys in Cassandra.
8. **Social Network Analysis (Neo4j)**
   Model and query social connections using Neo4j graph relationships to identify mutual friends, communities, and influencers.
9. **Recommendation Engine (Neo4j)**
   Build a recommendation system (e.g., movies or products) using Neo4j's graph traversal and similarity queries for AI-driven personalization.
10. **Travel Booking and Recommendation System (MongoDB + Redis + Neo4j)**
    Integrate multiple NoSQL databases: use MongoDB for booking data, Redis for caching, and Neo4j for suggesting travel destinations or routes based on relationships.

**Beyond these projects, students should:**

- Practice **CRUD operations** and **data modeling** in MongoDB, Cassandra, Redis, and Neo4j.
- Explore simple **AI and data analytics ideas**, such as recommendations or trend analysis using NoSQL datasets.
- Develop small applications to **interact with databases programmatically** (using Python or Node.js).
- Observe how data is **stored, queried, and visualized**, and implement basic **reports or dashboards** for insights.
- Adds **clear one-line purpose/descriptions** for each project.
- Introduces **AI and analytics relevance** where suitable.
- Ensures **balanced coverage** of all four NoSQL types.
- Improves consistency in grammar and structure (every project starts with an action verb).
- The last section ties projects to **learning outcomes** and skill development.

**Suggested Activities for Students:**

In addition to classroom and laboratory sessions, students are encouraged to participate in the following **co-curricular and experiential learning activities**. These activities are designed to strengthen conceptual understanding, promote teamwork, and enhance employability skills related to NoSQL database technologies.

**Recommended Activities**

a) **Hands-on Practical Sessions in a Controlled Lab Environment**
- Conduct structured experiments on database creation, queries, and data operations using MongoDB, Cassandra, Neo4j, and Redis.
- Perform exercises beyond the standard lab list, such as query optimization or multi-database integration.

b) **Database Manipulation and Performance Exercises**
- Design and manipulate datasets to analyze query execution times, indexing effects, and schema flexibility.
- Compare query performance between relational and NoSQL databases.

c) **Mini Application Development Using NoSQL Databases**
- Develop small-scale applications such as user management, product catalogs, or chat systems.
- Use APIs or programming languages (Python, Node.js, or Java) to connect and interact with NoSQL databases.

d) E**xploration of Simple AI / Data Analytics Applications**
- Implement simple recommendation systems, trend analysis, or clustering using NoSQL datasets.
- Demonstrate how unstructured data can be leveraged in AI and machine learning pipelines.

e) **Study and Presentation of Real-World Case Studies**
- Analyze industry case studies (e.g., Netflix using Cassandra, LinkedIn using Neo4j, or e-commerce systems using MongoDB).
- Present findings on database selection rationale, scalability, and performance outcomes.

**********