



National University of Engineering (UNI)
School of Computer Science
Syllabus 2026-I

1. COURSE

CS272. Databases II (Mandatory)

2. GENERAL INFORMATION

2.1 Course	: CS272. Databases II
2.2 Semester	: 5 th Semester.
2.3 Credits	: 3
2.4 Hours	: 1 HT; 4 HP;
2.5 Duration of the period	: 16 weeks
2.6 Type of course	: Mandatory
2.7 Learning modality	: Face to face
2.8 Prerequisites	: CS271. Data Management. (4 th Sem)

3. PROFESSORS

Meetings after coordination with the professor

4. INTRODUCTION TO THE COURSE

This course is a continuation of Data Management I, focusing on the internal mechanisms of database management systems (DBMS) and modern database technologies. It covers query optimization, indexing structures, and transactional integrity in depth. Additionally, it introduces students to non-relational models (NoSQL) and distributed architectures, which are essential for handling large-scale data in contemporary cloud environments.

5. GOALS

- Understand the internal storage structures and indexing techniques.
- Analyze and optimize query execution plans.
- Comprehend advanced transaction protocols and recovery mechanisms.
- Evaluate and implement solutions using NoSQL database models.
- Understand the principles of distributed and parallel databases.

6. COMPETENCES

AG-C08) Problem Analysis: Identifies, formulates, and analyzes complex computing problems. (Usage)

AG-C12) Uses systemic approaches to select, develop, apply, integrate, and manage secure computing technologies to achieve user objectives. (Usage)

7. TOPICS

Unit 1: Procesamiento de Consultas (12 hours)	
Competences Expected: AG-C08,AG-C12	
Topics	Learning Outcomes
• Estructuras de página • Estructuras de índice: – árboles B+ – Índices hash: estáticos y dinámicos – Creación de índices en SQL • Estructuras de archivos: – Archivos de montón (<i>heap</i>) – Archivos hash • Algoritmos para operadores de consulta: – Ordenamiento externo – Selección – Proyección; con y sin eliminación de duplicados – Uniones naturales: Bucle anidado, Mezcla-ordenada, Unión hash – Análisis de eficiencia de algoritmos • Transformaciones de consultas • Optimización de consultas: – Rutas de acceso – Construcción de plan de consulta – Estimación de selectividad – Planes solo de índice • Procesamiento de Consultas en Paralelo (por ejemplo, exploración paralela, unión paralela, agregación paralela) • Afinamiento/rendimiento de base de datos: – Selección de índices – Impacto de índices en el rendimiento de consultas – Desnormalización	• Describir el propósito y organización de las estructuras de índice tanto de árbol B+ como hash [Describir] • Componer un comando SQL para crear un índice (de cualquier tipo) [Componer] • Especificar los pasos para los diversos algoritmos de operadores de consulta: ordenamiento externo, proyección con eliminación de duplicados, unión por mezcla-ordenada, unión hash, unión por bucle anidado por bloques [Analizar] • Derivar el tiempo de ejecución (en solicitudes de E/S) para cada uno de los algoritmos anteriores [Analizar] • Transformar una consulta en álgebra relacional a su equivalente apropiado para una ejecución canalizada de profundidad izquierda [Rediseñar] • Calcular estimaciones de selectividad para una operación de selección y/o unión dada [Computar/Calcular] • Describir cómo modificar una estructura de índice para facilitar una operación solo de índice para una relación dada [Describir] • Para un escenario dado decidir qué índices soportar para la ejecución eficiente de un conjunto de consultas [Evaluar (valorar)] • Describir cómo los DBMS aprovechan el paralelismo para acelerar el procesamiento de consultas dividiendo el trabajo entre múltiples procesadores o nodos [Describir]
Readings : [SKS19], [RG03]	

Unit 2: Internos de DBMS (12 hours)	
Competences Expected: AG-C08,AG-C12	
Topics	Learning Outcomes
• Gestión de Búfer de DB • Gestión de Transacciones: – Niveles de Aislamiento – ACID – Serializabilidad – Transacciones Distribuidas • Control de Concurrency: – Bloqueo en 2 Fases – Estrategias de manejo de bloqueos mutuos – Modelos de consistencia basados en quórum • Gestor de Recuperación: – Relación con el Gestor de Búfer • Control de Concurrency: – Control de concurrencia optimista – Control de concurrencia por marca de tiempo • Gestor de Recuperación: – Registro por escritura anticipada – Sistema de recuperación ARIES (Análisis, REDO, UNDO)	• Describir cómo un DBMS gestiona su Grupo de Búferes [Describir] • Describir las cuatro propiedades para un gestor de transacciones correcto [Describir] • Esbozar el principio de serializabilidad [Esquematar/Esbozar]
Readings : [SKS19], [GUW08]	

Unit 3: Bases de Datos Distribuidas/Computación en la Nube (12 hours)	
Competences Expected: AG-C08,AG-C12	
Topics	Learning Outcomes
• DBMS Distribuido: – Almacenamiento de datos distribuido – Procesamiento de consultas distribuido – Modelo de transacción distribuida – Soluciones homogéneas y heterogéneas – Bases de datos distribuidas cliente-servidor • DBMS Paralelo: – Arquitecturas de DBMS paralelo: memoria compartida, disco compartido, nada compartido – Aceleración y escalamiento, por ejemplo, uso del modelo de procesamiento MapReduce – Replicación de datos y modelos de consistencia débil	• Describir los componentes clave de un DBMS distribuido, incluyendo almacenamiento de datos distribuido, procesamiento de consultas y gestión de transacciones [Describir] • Analizar las ventajas y desventajas entre arquitecturas de DBMS paralelo: memoria compartida, disco compartido y nada compartido [Analizar] • Describir estrategias de replicación de datos y modelos de consistencia débil en sistemas de bases de datos distribuidas [Describir]
Readings : [SKS19], [ÖV20]	

Unit 4: Sistemas NoSQL (12 hours)	
Competences Expected: AG-C08,AG-C12	
Topics	Learning Outcomes
• ¿Por qué NoSQL? (por ejemplo, Desajuste de impedancia entre Aplicación [CRUD] y RDBMS) • Modelo de datos Clave-Valor y Documento • Sistemas de almacenamiento (por ejemplo, sistemas Clave-Valor, Lagos de Datos (<i>Data Lakes</i>)) • Modelos de Distribución (Fragmentación y Replicación) • Bases de Datos de Grafos • Modelos de Consistencia (Actualización y Lectura, consistencia de quórum, teorema CAP) • Modelo de procesamiento (por ejemplo, Map-Reduce, map-reduce multi-etapa, map-reduce incremental) • Estudios de Caso: Sistema de almacenamiento en la nube (por ejemplo, S3); Bases de datos de grafos; "Cuándo no usar NoSQL"	• Desarrollar un caso de uso para el uso de NoSQL sobre RDBMS [Crear] • Describir las características definitorias detrás de los modelos de datos basados en Clave-Valor y Documentos [Describir]
Readings : [SF12], [Kle17]	

Unit 5: Bases de Datos Semiestructuradas y No Estructuradas (12 hours)	
Competences Expected: AG-C08,AG-C12	
Topics	Learning Outcomes
- Datos no estructurados vectorizados (texto, video, audio, etc.) y almacenamiento vectorial: - Vectorizador TF-IDF con n-grama - Word2Vec - Base de datos de arreglos o manejo de tipo de datos arreglo - Bases de datos semiestructuradas (por ejemplo, JSON): - Almacenamiento: - Codificación y compresión de tipos de datos anidados - Indexación: - árbol B, índice de salto, filtro Bloom - Índice invertido y compresión de mapa de bits - Indexación por curva de llenado de espacio para datos geo-semiestructurados - Procesamiento de consultas para casos de uso OLTP y OLAP: - Compromisos de inserción, selección, actualización/eliminación - Estudios de caso sobre Postgres/JSON, MongoDB y Snowflake/JSON	- Describir representaciones vectorizadas de datos no estructurados (por ejemplo, TF-IDF, Word2Vec) y sus modelos de almacenamiento [Describir] - Aplicar técnicas de indexación (por ejemplo, árbol B, índice invertido, filtro Bloom) a datos semiestructurados almacenados en formato JSON [Aplicar] - Analizar las ventajas y desventajas del procesamiento de consultas en bases de datos semiestructuradas para casos de uso OLTP y OLAP [Analizar]
Readings : [Kle17], [MRS08], [Mik+13], [Agr+23], [Mon25]	

8. WORKPLAN

8.1 Methodology

Individual and team participation is encouraged to present their ideas, motivating them with additional points in the different stages of the course evaluation.

8.2 Theory Sessions

The theory sessions are held in master classes with activities including active learning and roleplay to allow students to internalize the concepts.

8.3 Practical Sessions

The practical sessions are held in class where a series of exercises and/or practical concepts are developed through problem solving, problem solving, specific exercises and/or in application contexts.

9. EVALUATION SYSTEM

***** EVALUATION MISSING *****

10. BASIC BIBLIOGRAPHY

- [RG03] Raghu Ramakrishnan and Johannes Gehrke. *Database Management Systems*. 3rd. McGraw-Hill, 2003.
- [GUW08] Hector Garcia-Molina, Jeffrey D. Ullman, and Jennifer Widom. *Database Systems: The Complete Book*. 2nd. Pearson, 2008.
- [MRS08] Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schütze. *Introduction to Information Retrieval*. Cambridge University Press, 2008.

- [SF12] Pramod J. Sadalage and Martin Fowler. *NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence*. Addison-Wesley, 2012.
- [Mik+13] Tomas Mikolov et al. “Efficient Estimation of Word Representations in Vector Space”. In: *arXiv preprint* (2013). arXiv:1301.3781.
- [Kle17] Martin Kleppmann. *Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable Systems*. O’Reilly Media, 2017.
- [SKS19] Abraham Silberschatz, Henry F. Korth, and S. Sudarshan. *Database System Concepts*. 7th. McGraw-Hill, 2019.
- [ÖV20] M. Tamer Özsu and Patrick Valduriez. *Principles of Distributed Database Systems*. 4th. Springer, 2020.
- [Agr+23] Divy Agrawal et al. “Toward a Systems Architecture for AI-Enabled Applications”. In: *Communications of the ACM* 66.6 (2023), pp. 54–63. DOI: 10.1145/3580316.
- [Mon25] MongoDB, Inc. *MongoDB Manual*. MongoDB, Inc. <https://www.mongodb.com/docs/>. 2025.