

CARRERA: Ingeniería en Informática	CURSO LECTIVO: 2028
ASIGNATURA: Base de Datos Avanzada y Big Data	CURSO: 5° año – 1° semestre
DURACIÓN: Semestral	Hs. TOTALES: 48 Hs. Reloj Totales
SEMANAS: 16	Hs. TEÓRICAS: 32 Hs. Reloj Totales Hs. PRÁCTICAS: 16 Hs. Reloj Totales

PROFESOR PROTITULAR: Ricardo Di Pasquale

PROFESOR ASISTENTE: Luciana González

1. OBJETIVOS DE LA ASIGNATURA

Que el alumno logre:

- Introducir otros modelos de Bases de Datos diferentes al estándar relacional (SQL).
- Presentar los distintos enfoques de consumo de datos en función de la naturaleza de las aplicaciones.
- Introducir el ciclo de los datos en organizaciones orientadas a datos.
- Explorar las disciplinas relacionadas a los proyectos de inteligencia aplicada en el ciclo de los datos.
- Aplicar los conocimientos adquiridos de estadística, programación, estructura de datos, modelos y simulaciones, y bases de datos para extraer valor del activo más importante de las organizaciones orientadas a datos: sus datos.

2. COMPETENCIAS

Eje / Enunciado	Bajo	Medio	Alto
1. Especificación, proyecto y desarrollo de sistemas de información.			
2. Especificación, proyecto y desarrollo de sistemas de comunicación de datos.			
3. Especificación, proyecto y desarrollo de software.			
4. Proyecto y dirección en lo referido a seguridad informática.			
5. Establecimiento de métricas y normas de calidad de software.			
6. Procedimientos y certificaciones del funcionamiento, condición de uso o estado de sistemas de información, sistemas de comunicación de datos, software, seguridad informática y calidad de software.			
7. Dirección y control de la implementación, operación y mantenimiento de sistemas de información, sistemas de comunicación de datos, software, seguridad informática y calidad de software.			
8. Identificación, formulación y resolución de problemas de ingeniería en sistemas de información / informática.			
9. Concepción, diseño y desarrollo de proyectos de ingeniería en sistemas de información / informática.			X
10. Gestión, planificación, ejecución y control de proyectos de ingeniería en sistemas de información / informática.			
11. Utilización de técnicas y herramientas de aplicación en la ingeniería en sistemas de información / informática.			
12. Generación de desarrollos tecnológicos y/o innovaciones tecnológicas.			X
13. Desempeño en equipos de trabajo.			X
14. Comunicación efectiva.			X
15. Actuación profesional ética y responsable.			
16. Evaluación y actuación en relación con el impacto social de su actividad profesional en el contexto global y local.			
17. Aprendizaje continuo.			
18. Desarrollo de una actitud profesional emprendedora.			

3. UNIDADES TEMÁTICAS

Unidad Temática 1 - Bases de Datos Orientadas a Objetos y ORM en Aplicaciones OLTP:

Orientación a objetos. Diferencias de impedancia con el modelo relacional. ORM. Conceptos y estrategias de mapeo Objeto-Relacional. Modelo de Dominio. Ciclo de Vida de los Objetos.

Unidad 2 - Procesamiento por Lotes:

Características del Procesamiento "Batch". Referencia histórica. Frameworks para su tratamiento con tecnologías orientadas a objetos. Implementación de procesos Batch.

Unidad Temática 3 – Big Data:

No-SQL. Conjetura de Brewer (CAP). BASE vs ACID. Schemaless. Bases de Datos Orientadas a Documentos. MongoDB. Bases de Datos Clave-Valor. DynamoDB. Bases de Datos Orientadas a Grafos. File Systems distribuidos. Hadoop. HDFS. Map Reduce. Spark. Resilient Distributed Data Structures. Bases de datos Vectoriales.

Unidad Temática 4 – Data Science y Data Engineering::

[AWS agregar y Conceptos de Cloud]. Pipelines de procesamiento. Data Quality. Data Observability. Arquitecturas de referencia. Cloud Data Platforms. [Completar temas AWS]. Data engineering. Trade-off sesgo varianza. Diferencia Correlación – Causalidad. Modelo probabilístico. Taxonomía básica de los métodos en Ciencia de Datos. Métodos Supervisados. Clasificación. Regresión. Métodos No Supervisados. Modelos de Forecasting.

Unidad Temática 5 - Bases de datos para el soporte de toma de decisiones (Business Intelligence):

OLAP. Datawarehouse. Dashboards. Modelo dimensional. Cubos de información. Procesos de Extracción, Transformación y Carga (ETL)

Unidad Temática 6 - Tiempo real y Streaming:

Características de las aplicaciones de Tiempo Real. Tiempo casi real. Restricciones de los runtime para la ejecución en tiempo real. Bases de Datos orientadas a series temporales. Implementación. Kafka y modelos de colas.

Unidad Temática 7 - Otros modelos de bases de datos:

Bases de datos orientadas a documentos para aplicaciones OLTP. Bases de datos orientadas a grafos para aplicaciones OLTP y BI.

Unidad Temática 8 - Cloud Data Plataforms y Data Governance:

[Completar con apuntes]. Vector database.

4.

BIBLIOGRAFÍA

4.1

BIBLIOGRAFÍA GENERAL OBLIGATORIA

- V. Khononov, Learning Domain-Driven Design: Aligning Software Architecture and Business Strategy, O'Reilly, 2021. ISBN: 978-1098100131.
- D. Bechberger, J. Perryman, Exploring Graph Databases, Manning, 2021. ISBN: 978-1617299674 .
- R. Johnson, J. Hoeller, K. Donald et al., Spring Framework Reference Documentation, Spring, 2022. Disponible en <https://docs.spring.io/spring-framework/docs/current/reference/pdf/spring-framework.pdf>
- N. Marz, J. Warren, Big Data: Principles and best practices of scalable realtime data systems. Manning Publications, 2015. ISBN: 978-1617290343.
- M. Feick, N. Kleer, M. Kohn, Fundamentals of Real-Time Data Processing Architectures: Lambda and Kappa, SKILL 2018, Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2018.
- T. White, Hadoop: The Definitive Guide, 4th Edition. O'Reilly Media, Inc. 2015. ISBN: 9781491901632.
- J. Perrin, Spark in Action, Second Edition, Manning. 2020. ISBN 978- 1617295522.
- D. Zburivsky, L. Partner, Designing Cloud Data Platforms. Manning. 2021. ISBN: 978-1617296444.

4.2

BIBLIOGRAFÍA GENERAL COMPLEMENTARIA

- O. Gierke, T. Darimont, C. Strobl, M. Paluch, et al., Spring Data JPA - Reference Documentation. Spring. 2022. Disponible en <https://docs.spring.io/spring-data/jpa/docs/current/reference/html/>
- J. Becla, D. Wang, Lessons Learned from Managing a Petabyte, Conference on Innovative Data Systems Research. 2005. Disponible en <https://pdfs.semanticscholar.org/b8c1/358de9b3eb459b5e4bff40ec2f982daddf15.pdf>
- L. Ward , D. Syer , T. Risberg , R. Kasanicky, D. Garrette, W. Lund, M. Minella, C. Schaefer, G. Hillert, Spring Batch Reference. Spring. 2022. Disponible en <https://docs.spring.io/spring-batch/docs/current/reference/pdf/spring-batch-reference.pdf>
- C. Davis, Cloud Native Patterns: Designing change-tolerant software. Manning. 2019. ISBN: 978-1617294297.
- D. Scott, V. Gamov, D. Klein, Kafka in Action. Manning. 2022. ISBN: 978-1617295232.
- B. Wilson, Machine Learning Engineering in Action. Manning. 2022. ISBN: 978-1617298714.

5.

METODOLOGÍA DE ENSEÑANZA

Los elementos sobre los que se basa el proceso de Enseñanza y Aprendizaje en la materia

se basan en los siguientes pilares:

- Clases teóricas: El profesor expondrá cada tema del programa, que se complementará con lecturas previas. Se realizarán trabajos individuales y grupales, para facilitar el intercambio de opiniones en cada clase. Se espera la participación de los alumnos, basada en la teoría expuesta en la clase y en las experiencias obtenidas de los trabajos y de los prácticos.
- Clases prácticas: Se realizarán trabajos individuales y grupales y el análisis de casos (asignados a los alumnos previamente o en clase - según su extensión -), para facilitar el intercambio de opiniones en cada clase. Se espera la participación de los alumnos, basada en la teoría expuesta en la clase y en las experiencias obtenidas de los trabajos y de los casos.
- Coloquios: Al principio de la cursada se asignan artículos o capítulos de libros para que cada alumno individualmente exponga el tema, tal como se realiza en coloquios de posgrado, buscando que los alumnos indaguen la bibliografía, descubran técnicas nuevas y trasmitan ese conocimiento a sus alumnos con claridad expositiva.

6.

FORMACIÓN PRÁCTICA

Casos prácticos: Durante la cursada se expondrán casos prácticos que van desde la resolución de problemas concretos en la plataforma cloud de la Facultad, así como el diseño de plataformas de procesamiento de datos, que deberán ser analizados y resueltos de manera grupal.

7.

CRITERIO Y MODALIDAD DE EVALUACIÓN

La evaluación será formativa, en proceso, con el seguimiento de las actividades que realice el alumno y el monitoreo de la adquisición de las competencias establecidas. La evaluación continua se realiza a través de la participación en clase, en las prácticas y en los coloquios. Asimismo, se cuenta con tres instancias de evaluación sumativa: dos exámenes parciales y el examen final de la materia.

Criterios de evaluación

- Claridad y precisión conceptual.
- Coherencia en la exposición de las ideas.
- Establecimiento de relaciones pertinentes.
- Utilización de vocabulario específico.
- Inclusión de elementos teóricos adecuados en los análisis realizados
- Utilización de herramientas teóricas pertinentes en la fundamentación
- Integralidad del abordaje realizado frente a las cuestiones planteadas.
- Coherencia entre la argumentación y la cuestión planteada.
- Identificación clara de las herramientas conceptuales implicadas.
- Identificación de los procedimientos que corresponde utilizar
- Aplicación adecuada de los procedimientos requeridos.

Criterios de calificación

- Escala numérica para los coloquios.
- Escala numérica para las evaluaciones parciales y el examen final.

Instancias de evaluación sumativa

Exámenes parciales: La materia tiene dos exámenes parciales, con una instancia de recuperación. El examen parcial consiste en una evaluación escrita y presencial, donde el alumno deberá demostrar los conocimientos teóricos adquiridos y deberá resolver algunos ejercicios similares a los resueltos en la parte práctica de la materia. Los recuperatorios se realizarán a los alumnos que no hayan aprobado el parcial y los que estuvieron ausentes, en condiciones idénticas al mismo.

FIRMA DE LA CURSADA

Para aprobar la cursada es necesario tener los niveles de asistencia a clases que fija la Universidad, tener aprobados dos parciales o sus recuperatorios y aprobar el Coloquio individual.

Examen final: El examen final consiste en una evaluación oral y escrita, presencial e individual, donde el alumno deberá demostrar conocimientos teóricos y prácticos. El examen final se diferencia en que abarca todos los temas del programa, incluyendo aquellos temas que no fueron evaluados en el examen parcial. Además los ejercicios prácticos tendrán un carácter integrador y superador con respecto a los realizados en clase. Finalmente, en las preguntas teóricas se pretende que el alumno demuestre un conocimiento profundo de los temas, relacionando conceptos entre sí.