



PROGRAMA DE ASIGNATURA

I. IDENTIFICACIÓN

Carrera o programa: Ingeniería en Tecnologías de Información

Unidad responsable: Escuela de Ingeniería

Nombre de la asignatura: Ingeniería de Datos y Big Data

Código: ECIN-00762

Semestre en la malla¹: 7

Créditos SCT – Chile: 5

Ciclo de Formación	Básico		Profesional	X
Tipo de Asignatura	Obligatoria	X	Electiva	

Clasificación de área de conocimiento²

Área: Ingeniería y Tecnología **Subárea:** Ingeniería Informática

Requisitos:

Prerrequisitos:	Requisitos para:
- Desarrollo de Soluciones Web/Móvil - Ciencia de Datos	

II. ORGANIZACIÓN SEMESTRAL

Horas Dedicación Semanal (Cronológicas)	Docencia Directa	4.5		Trabajo Autónomo		3.5	Total	8
Detalle Horas Directas	Cátedra	Ayudantía	Laboratorio	Taller		Terreno	Exp. Clínica	Supervisión
	3		1.5					

III. APOORTE AL PERFIL DE EGRESO

La asignatura contribuye al dominio 1 del perfil de egreso, “Conocimiento científico y disciplinario”. Además, contribuye al dominio 2 “Habilidades y Actitudes Personales y Profesionales”. También contribuye al dominio 3 Habilidades Interpersonales”. Por último, contribuye al dominio 4 “Habilidades para la Práctica de la Ingeniería”. Al finalizar la asignatura las y los estudiantes serán capaces de modelar un problema de datos y aplicar el diseño, implementación y programación de tecnologías de big data para resolver una problemática del mundo real.

IV. COMPETENCIAS

La carrera declara las siguientes habilidades:

- 1.3 Aplicar conocimientos, métodos y herramientas de la especialidad para resolver problemas complejos en Ingeniería de Software, Plataformas y Gestión de Tecnologías.
- 2.1 Identificar, formular, modelar y resolver problemas complejos de ingeniería considerando las interacciones y la dinámica de las variables.
- 3.1 Liderar y trabajar en equipos de trabajo multidisciplinarios.

¹ Este campo sólo se completa en caso de carreras con programas semestrales.

² Clasificación del curso de acuerdo a la OCDE



3.2 Comunicar de manera comprensible información técnica en español, de forma oral, escrita y gráfica, a nivel avanzado.

4.4 Diseñar soluciones TIC para la industria intensiva en procesamiento de la información. Estas soluciones toman en cuenta las arquitecturas TI junto con sus modelos de servicios y modelos operativos, los sistemas de software y las plataformas de cómputo y comunicaciones junto con sus servicios asociados.

4.5 Implementar soluciones TIC. Estas soluciones consideran las arquitecturas TI junto con sus modelos de servicios y modelos operativos, los sistemas de software y las plataformas de cómputo y comunicaciones junto con sus servicios asociados.

V. RESULTADOS DE APRENDIZAJE

1. Diseñar modelos de almacenamiento de datos eficientes y escalables para gestionar grandes volúmenes de información en el contexto de proyectos de Big Data.
2. Implementar procesos de extracción, transformación y carga (ETL) de datos utilizando herramientas y técnicas modernas, considerando la diversidad de fuentes y formatos de datos en entornos de Big Data.
3. Analizar conjuntos de datos masivos para identificar patrones, tendencias y oportunidades de negocio, utilizando herramientas de análisis de datos y visualización de datos.
4. Desarrollar algoritmos de procesamiento distribuido para realizar análisis y cálculos en entornos de Big Data, aprovechando las capacidades de paralelismo y escalabilidad.
5. Evaluar tecnologías y soluciones de almacenamiento, procesamiento y análisis de datos que sean apropiadas para diferentes escenarios y requerimientos en proyectos de Big Data.
6. Aplicar prácticas de seguridad y privacidad de datos en el diseño y desarrollo de soluciones de ingeniería de datos y Big Data, cumpliendo con las regulaciones y estándares relevantes.

VI. ÁREAS TEMÁTICAS

1. Big Data y su Ecosistema
 - 1.1. Definición y características de Big Data
 - 1.2. Componentes del ecosistema de Big Data (Hadoop, Spark, etc.)
 - 1.3. Casos de uso y aplicaciones de Big Data
2. Almacenamiento de Datos en Big Data
 - 2.1. Bases de datos NoSQL
 - 2.2. Almacenamiento distribuido
 - 2.3. Almacenamiento en la nube (cloud storage)
3. Procesamiento de Datos en Big Data
 - 3.1. Procesamiento distribuido
 - 3.2. MapReduce y programación en Spark
 - 3.3. Flujo de datos en tiempo real
4. Extracción y Transformación de Datos (ETL)
 - 4.1. Técnicas y herramientas ETL
 - 4.2. Limpieza y preparación de datos
 - 4.3. Integración de datos de múltiples fuentes
5. Análisis y Visualización de Datos en Big Data
 - 5.1. Minería de datos



- 5.2. Visualización de datos masivos
- 5.3. Análisis predictivo y prescriptivo
- 6. Machine Learning en Big Data
 - 6.1. Fundamentos de machine learning
 - 6.2. Entrenamiento de modelos en grandes conjuntos de datos
 - 6.3. Aplicaciones de machine learning en Big Data
- 7. Arquitecturas y Diseño de Soluciones de Big Data
 - 7.1. Arquitecturas de procesamiento en lotes y en tiempo real
 - 7.2. Diseño de pipelines de datos
 - 7.3. Escalabilidad y rendimiento
- 8. Seguridad y Privacidad en Big Data
 - 8.1. Retos de seguridad en Big Data
 - 8.2. Políticas y prácticas de privacidad
 - 8.3. Cumplimiento normativo
- 9. Ética y Responsabilidad en el Uso de Datos
 - 9.1. Consideraciones éticas en la recopilación y uso de datos
 - 9.2. Implicaciones sociales y legales
- 10. Tendencias Emergentes en Big Data
 - 10.1. Internet de las cosas (IoT) y Big Data
 - 10.2. Big Data en la inteligencia artificial
 - 10.3. Avances tecnológicos en Big Data

VII. ORIENTACIONES METODOLÓGICAS

- 1. La metodología a desarrollar en esta asignatura es principalmente práctica, por lo que se debe fomentar la interacción entre los estudiantes a través de trabajos prácticos colaborativos.
- 2. Se sugiere abordar la teoría mediante metodologías activas como la clase invertida, el aprendizaje basado en problemas (ABP), entre otras, creando oportunidades para presentaciones orales individuales y/o grupales, lo que favorece el aprendizaje contextualizado.
- 3. Las experiencias en la cátedra y el taller deben llevarse a cabo utilizando software moderno aplicable a la asignatura.
- 4. Se recomienda que los estudiantes realicen presentaciones periódicas sobre el trabajo realizado en el taller.

VIII. ORIENTACIONES Y CRITERIOS PARA EVALUACIÓN

- 1. Se recomienda la aplicación de una evaluación diagnóstica al inicio de la asignatura.
- 2. En la cátedra, se sugiere implementar estrategias de evaluación sumativa, con un peso del 70% en la nota final de la asignatura.
- 3. En el taller, se sugiere implementar estrategias de evaluación sumativa, con un peso del 30% en la nota final de la asignatura.
- 4. Las actividades pueden ser individuales o grupales.
- 5. La asistencia mínima requerida para las actividades de cátedra es del 70%.



6. Se requerirá un logro del 60% de los objetivos para aprobar las actividades de evaluación.
7. Se recomienda llevar a cabo evaluaciones de carácter formativo con retroalimentación de carácter personal.

IX. RECURSOS BIBLIOGRÁFICOS

Bibliografía Mínima

Reis, J., & Housley, M. (2022). *Fundamentals of Data Engineering*. O'Reilly Media. ISBN: 9781098108250