

Programa de Asignatura

1. Identificación Asignatura

Nombre:	Ingeniería de Datos			Código:	IN1032
Carrera:	Ingeniería Civil Industrial	Unidad Académica:	Departamento Ciencias Naturales y Tecnología		
Ciclo Formativo:	Licenciatura	Línea formativa:	Especialidad		
Semestre	VII	Tipo de actividad:	Obligatoria		
N° SCT:	6	Horas Cronológicas Semanales			
		Presenciales:	4.5	Trabajo Autónomo:	4.5
Pre-requisitos	IN1028 – Introducción a las TICA				

2. Propósito formativo

El propósito de esta asignatura es que el estudiante comprenda, aplique y evalúe conocimientos y habilidades de la **Ingeniería en Ciencia de Datos** para el análisis de datos, visualización y aprendizaje automático, abarcando el ciclo completo de trabajo con datos: desde su recepción y limpieza hasta su análisis exploratorio, modelado y presentación de resultados. A través del desarrollo de competencias en técnicas de preprocesamiento, identificación de patrones y construcción de modelos predictivos, el estudiante aprenderá a transformar datos en información valiosa para la toma de decisiones en distintos contextos profesionales.

El aprendizaje se desarrollará a través de una combinación de clases teóricas y prácticas, integrando herramientas de programación y software especializado en análisis de datos, visualización y aprendizaje automático. Se trabajará con conjuntos de datos reales, fomentando el pensamiento crítico y la resolución de problemas mediante estudios de caso y proyectos aplicados. Además, se promoverá el trabajo colaborativo y multidisciplinario, incentivando la comunicación efectiva de hallazgos y resultados.

Esta asignatura es clave en la formación del futuro profesional, ya que proporciona habilidades esenciales para la gestión y análisis de datos en sectores como la industria, la investigación y la toma de decisiones estratégicas. La capacidad de limpiar, analizar, modelar y visualizar datos permitirá a los egresados generar conocimientos accionables, optimizar procesos y desarrollar soluciones basadas en evidencia.

En cuanto a su conexión curricular, esta asignatura complementa conocimientos adquiridos en los cursos de Herramientas Informáticas, Programación I, Programación II, Introducción a las TICA y Probabilidades y Estadísticas Aplicadas, proporcionando una perspectiva aplicada y orientada a la resolución de problemas con datos. Además, sirve como base para cursos más avanzados en Big Data, Ciencia de Datos e Inteligencia Artificial, fortaleciendo la formación en métodos analíticos y predictivos.

3. Contribución al perfil de egreso

Esta asignatura contribuye a los siguientes desempeños o resultados de aprendizaje globales declarados en el Perfil de Egreso de la carrera:

Demuestra dominio de las ciencias básicas, de energía y de las ciencias de la ingeniería y la capacidad de realizar análisis de tipo económico, políticos y sociales.

Concibe soluciones a los problemas que surgen en las organizaciones.

Diseña, selecciona y adapta desarrollos tecnológicos y científicos propios de la ingeniería industrial a los desafíos de las organizaciones.

4. Resultados de aprendizaje específicos

Resultado de Aprendizaje Específico	Criterios de evaluación	Evidencia
1. Comprende la importancia de los datos, su preprocesamiento, análisis y visualización en la toma de decisiones y el desarrollo de soluciones basadas en datos.	<p>1.1. Define correctamente los principales pasos del preprocesamiento de datos (limpieza, transformación, reducción de dimensionalidad)</p> <p>1.2 Identifica problemas comunes en los datos crudos y propone soluciones teóricas.</p> <p>1.3 Explica el propósito del análisis exploratorio de datos (EDA) y sus principales técnicas.</p> <p>1.4 Distingue entre distintos métodos de análisis de datos y juzga cuándo aplicarlos.</p> <p>1.5 Selecciona los tipos de gráficos más adecuados para distintos tipos de datos.</p>	<p>Prueba escrita:</p> <p>Preguntas de análisis de caso donde el estudiante deberá identificar problemas en un conjunto de datos ficticios y proponer soluciones teóricas.</p>
2. Aplica técnicas de preprocesamiento de datos, incluyendo limpieza, transformación y reducción de dimensionalidad, con el fin de optimizar la calidad y utilidad de los datos para su posterior análisis y modelado.	<p>2.1. Limpieza de datos: Identifica y corrige errores en los datos, como valores faltantes, duplicados o inconsistencias. Aplica técnicas adecuadas para manejar valores atípicos y datos ruidosos.</p> <p>2.2. Transformación de datos: Implementa y justifica el uso de técnicas de transformación de datos en función de la estructura y características de los datos.</p>	<p>Trabajo práctico.</p>
3. Desarrolla visualizaciones exploratorias y de resultados de análisis de datos de forma efectiva utilizando herramientas de visualización, permitiendo la comprensión clara y accesible de los resultados.	<p>3.1. Las visualizaciones presentadas son claras, comprensibles y están orientadas a los objetivos del análisis.</p> <p>3.2. Utiliza las herramientas de visualización en Python como Matplotlib y Seaborn de manera apropiada para los tipos de datos y resultados a presentar.</p> <p>3.3. Los resultados son presentados de forma coherente y los hallazgos son comunicados de manera efectiva al público objetivo.</p>	<p>Presentación con visualizaciones de los resultados obtenidos en un análisis de datos. El estudiante debe explicar cómo cada visualización contribuye a la interpretación de los datos y toma de decisiones.</p>
4. Desarrolla y evalúa modelos básicos de aprendizaje automático para la solución de un problema práctico.	<p>4.1. Comprende los conceptos teóricos fundamentales del aprendizaje automático.</p> <p>4.2. Justifica y aplica correctamente un método de aprendizaje automático a un problema práctico.</p> <p>4.3. Evalúa el desempeño de un modelo con métricas de error, exactitud y precisión.</p> <p>4.4. Presenta resultados de manera clara.</p>	<p>Prueba escrita.</p> <p>Trabajo práctico y presentación con póster.</p>

5. Unidades de Aprendizaje

Unidad 1: Análisis exploratorio de datos

0. Introducción

1.1. Tipos de datos

1.2. Preprocesamiento de datos

1.3. Estadística descriptiva

1.4. Métodos de visualización

1.5. Tipos de análisis exploratorios

1.6. Python para el análisis de datos

Unidad 2: Introducción al aprendizaje automático

2.1. Aprendizaje automático supervisado – Regresión y clasificación

2.2. Aprendizaje automático no supervisado – Agrupamiento

2.3. Reducción de dimensionalidad

2.4. Aprendizaje por refuerzo

Unidad 3: Aprendizaje supervisado

3.1. Regresión lineal

3.2. Regresión logística

3.3. Modelos lineales generalizados (opcional)

3.3. Máquinas de vectores de soporte

3.3. Árboles de decisión

Unidad 4: Selección de modelos

4.1. Sesgo y varianza. Overfitting y underfitting.

4.2. Set de datos para entrenamiento, validación y testeo.

4.3. Métodos de medición de error en regresión y clasificación.

4.4. Regularización y selección de variables.

Unidad 5: Aprendizaje no supervisado

5.1. K-Means clustering

5.2. PCA (Principal component analysis)

Unidad 6: Introducción a las redes neuronales (Opcional)

6. Recursos de Aprendizaje

Bibliografía

B1. Downey, A. (2012). Think Python. " O'Reilly Media, Inc. " .

B2. Han, J., Pei, J., & Kamber, M. (2011). Data mining: concepts and techniques. Elsevier.

B3. GÈron, A. (2019). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly Media.

B4. McKinney, W. (2012). Python for data analysis: Data wrangling with Pandas, NumPy, and IPython. " O'Reilly Media, Inc. " .

B5. Raschka, S., & Mirjalili, V. (2019). Python Machine Learning: Machine

Recursos materiales e infraestructura:

Acceso a U-Campus.

Computadores debidamente equipados para utilizar lenguajes de alto nivel (por ej.: Python, Tableau, Power BI, otros).

7. Comportamiento y ética académica:

Se espera que los estudiantes actúen en sus diversas actividades académicas y estudiantiles en concordancia con los principios de comportamiento ético y honestidad académica propios de todo espacio universitario y que están estipulados en el *Reglamento de Estudiantes de la Universidad de Aysén*, especialmente aquellos dispuestos en los artículos 23°, 24° y 26°.

Todo acto contrario a la honestidad académica realizado durante el desarrollo, presentación o entrega de una actividad académica del curso sujeta a evaluación, será sancionado con la suspensión inmediata de la actividad y con la aplicación de la nota mínima (1.0).

Planificación del curso

8. Responsables

Académico (s) Responsable (s) y equipo docente	Felipe Ochoa		
Contacto	felipe.ochoa@docentes.uaysen.cl		
Año	2025	Periodo Académico	Primer semestre
Horario clases	Cátedra: Lu 14.30 – 16.00 h*. Mi 14.30 – 16.00 h*. Sesiones Prácticas: Mi 16.15 – 17.45 h*.	Horario de atención estudiantes	Contactar al académico previamente mediante correo electrónico. Ju 14.30 – 16.00 h
Sala / Campus			

9. Metodología de Trabajo:

La asignatura contiene:			
Actividades de vinculación con el medio		Actividades relacionadas con proyectos de investigación	
En el curso se contemplan cuatro tipos de actividades docentes, las cuales se asocian a requerimientos de sala y al nivel de intervención del profesor/ayudante:			
Actividad docente	Descripción	Intervención del profesor/ayudante	Requerimiento de sala
Exposición conceptual	El profesor introduce los fundamentos preliminares de programación y conceptos esenciales para el desarrollo de habilidades prácticas, con una participación activa en el aula tradicional o entornos virtuales.	Alta	Sala de clases UCampus Online UCampus Offline
Programación expositiva	Mediante la resolución de problemas específicos de ingeniería en ciencia de datos, el profesor guía a los estudiantes a profundizar en el entendimiento de conceptos clave, empleando ejemplos prácticos en un entorno de sala de clases o plataformas online.	Alta	Sala de clases UCampus Online UCampus Offline

Programación tutorial	Esta modalidad incorpora pausas estratégicas en la exposición, permitiendo a los estudiantes completar tareas específicas bajo la supervisión del docente, favoreciendo el aprendizaje interactivo en laboratorios de computación o con uso de equipos personales	Media	Laboratorio de computación Computador persona
Actividad práctica / Programación autónoma	Fomenta la independencia de los estudiantes al enfrentar y resolver desafíos de ingeniería de datos, en grupos o individualmente, con mínima intervención docente, ideal para el trabajo en laboratorios de computación o mediante dispositivos personales.	Baja	Laboratorio de computación Computador persona

10. Evaluaciones:

<u>Evaluaciones y ponderaciones</u>			
Evaluación	Descripción	Ponderación	Fecha
Evaluación 1	Se evaluarán los contenidos de la unidad 1 por escrito a través de preguntas y estudios de caso.	25%	09-04-2025
Evaluación 2	Se evaluarán los contenidos de las unidades 2 y 3 mediante la entrega de un trabajo práctico.	25%	12-05-2025
Evaluación 3	Se evaluarán los contenidos de la unidad 4 mediante una prueba escrita.	25%	16-06-2025
Tareas prácticas	Tareas prácticas serán entregadas en formato digital, en los que se evaluará tanto los contenidos como su presentación y formato en función de lo señalado en clases.	25%	Todo el semestre

b) Examen
El curso tendrá un trabajo de aplicación práctico final en el cual se evaluará la aplicación de los contenidos expuestos y trabajados durante el semestre.

c) Ponderación nota final de la asignatura
Nota presentación: 70%
Examen final: 30% (Exposición de proyecto práctico)

Todos los estudiantes deberán presentar el proyecto práctico final.

d) Requisitos de aprobación de la asignatura (calificaciones y asistencia)
La nota final exigida para aprobar la asignatura es 4,0 o mayor.
Asistencia: 65%

e) Disposiciones reglamentarias de calificaciones y aprobación
Todas las calificaciones, incluidos los promedios ponderados, se expresarán en cifras con un decimal. La centésima igual o mayor a cinco se aproximará a la décima superior y la menor a cinco se desestimará.

En casos oportunamente justificados con el jefe de carrera, y con una suficiente antelación, el/la estudiante que no haya asistido a una o más evaluaciones tendrá derecho a rendir una evaluación recuperativa que integre los contenidos a evaluar en fecha establecida por el profesor. Dicha evaluación tendrá una ponderación equivalente a aquella no rendida y deberá cubrir los mismos objetivos de evaluación. Las inasistencias no justificadas a evaluaciones harán que ésta sea calificada automáticamente con la nota mínima (1,0).

11. Otros aspectos asociados al funcionamiento del curso:

Toda la coordinación del curso (comunicaciones, actualización de notas, material, etc.) será realizada a través de UCampus.

Es deber del estudiante mantenerse informado de las noticias, avisos y material entregado por el profesor a través de estos medios, se sugiere instalar en su dispositivo móvil la aplicación de UCampus.

Los correos electrónicos serán respondidos en horario laboral (lunes a viernes de 9:00 a 18:00hs), no se responderán correos fuera de ese horario.

Durante el desarrollo de las actividades lectivas, los teléfonos celulares deberán estar en silencio y guardados, a menos que el/la profesor/a específicamente requiera de estos equipos para la realización de su clase o durante algunos casos excepcionales conversados previamente con el/la docente a cargo.

Las actividades lectivas se dictarán de forma presencial, salvo excepciones sujeto a contingencias presentes durante el transcurso de la asignatura.

12. Planificación de las actividades de enseñanza- aprendizaje y de evaluación

Semana / Sesión	Fecha	Resultado(s) de Aprendizaje	Tema (Unidades de aprendizaje) y actividades	Recursos utilizados o lecturas	Actividad(es) de Trabajo Autónomo
Semana 1	10/03	1	Introducción estructura del curso, planificación.		Revisión programa del curso
Semana 1	12/03	1	Unidad 1: Análisis exploratorio de datos 0. Introducción curso 1.1. Tipos de datos		
Semana 2	17/03	1	Python y matemáticas para aprendizaje automático		
Semana 2	19/03	2	1.2. Preprocesamiento de datos. Práctico: Python en Visual Studio Code		
Semana 3	24/03	2	1.3. Estadística descriptiva		Revisión de apuntes y de literatura sugerida. Consultas a profesor
Semana 3	26/03	2	1.3. Estadística descriptiva Práctico: Estructuras de datos en Python. Uso de dataframes.		
Semana 4	31/03	3	1.4. Métodos de visualización		Revisión de apuntes y de literatura sugerida. Consultas a profesor

Semana 4	02/04	3	1.4. Métodos de visualización Práctico: Visualización de datos en Python. Matplotlib y seaborn		Revisión de apuntes y de literatura sugerida. Consultas a profesor
Semana 5	07/04	1	1.5. Tipos de análisis exploratorios		Revisión de apuntes y de literatura sugerida. Consultas a profesor
Semana 5	09/04	1,2,3	Evaluación 1 Práctico: Presentación análisis exploratorio		
Semana 6	14/04	4	Unidad 2: Introducción al aprendizaje automático 2.1. Aprendizaje automático supervisado – Regresión y clasificación 2.2. Aprendizaje automático no supervisado – Agrupamiento		Revisión de apuntes y de literatura sugerida. Consultas a profesor
Semana 6	16/04	2,4	2.3. Reducción de dimensionalidad 2.4. Aprendizaje por refuerzo Práctico: Visualización de datos en Python		Revisión de apuntes y de literatura sugerida. Consultas a profesor
Semana 7	21/04	4	Unidad 3: Aprendizaje supervisado 3.1. Regresión lineal		
Semana 7	23/04	3,4	3.1. Regresión lineal Práctico: Regresión lineal		
Semana 8	28/04	4	3.2. Regresión logística		
Semana 8	30/04	3,4	3.2. Regresión logística Práctico: Regresión logística		
Semana 9	05/05	4	3.3. Máquinas de vectores de soporte (SVM)		
Semana 9	07/05	3,4	3.3. Máquinas de vectores de soporte (SVM) Práctico: SVM		
Semana 10	12/05		Entrega Evaluación 2		
Semana 10	14/05	3,4	3.3. Árboles de decisión Práctico: Árboles de decisión y random forest		
Semana 11	19/05	4	Unidad 4: Selección de modelos 4.1. Sesgo y varianza. Overfitting y underfitting.		Revisión de apuntes y de literatura sugerida. Consultas a profesor
Semana 11	21/05		FERIADO		
Semana 12	26/05	1,4	4.1. Sesgo y varianza. Overfitting y underfitting.		Revisión de apuntes y de literatura sugerida. Consultas a profesor
Semana 12	28/05	1,3,4	4.1. Sesgo y varianza. Overfitting y underfitting. Práctico: Visualización de sesgo y varianza		Revisión de apuntes y de literatura sugerida. Consultas a profesor
Semana 13	02/06	4	4.2. Set de datos para entrenamiento, validación y testeo.		

Semana 13	04/06	4	4.3. Métodos de medición de error en regresión y clasificación. Práctico: Uso de librería scikit-learn		
Semana 14	09/06	4	4.4. Regularización y selección de variables.		
Semana 14	11/06	4	4.4. Regularización y selección de variables. Práctico: Algoritmos de selección de variables		
Semana 15	16/06		Evaluación 3		
Semana 15	18/06	2,4	Unidad 5: Aprendizaje no supervisado 5.1. K-Means clustering 5.2. PCA (Principal component analysis) Práctico: K-Means		
Semana 16	23/06		Presentación proyecto final		
Semana 17	30/06		Prueba recuperativa		
Semana 18	07/06		Examen		

*Este programa y/o planificación de actividades podrá sufrir modificaciones, las que, de ser el caso, serán informadas de manera oportuna