



Programa de Asignatura

2.5.29 Ciencia de Datos

I. IDENTIFICACIÓN				
Carrera o Programa: Ingeniería Civil Industrial				
Unidad responsable: Escuela de Ingeniería				
Nombre de la asignatura: Ciencia de Datos				
Código: ECIN- XXXX				
Semestre en la malla⁵⁸ : 5				
Créditos SCT – Chile: 5				
Ciclo de Formación	Básico		Profesional	X
Tipo de Asignatura	Obligatoria	X	Electiva	
Clasificación de área de Conocimiento⁵⁹				
Área: Ingeniería y Tecnología			Subárea: Otras Ingenierías y Tecnologías	
Requisitos				
Pre - Requisitos:			Requisito para:	
<ul style="list-style-type: none"> ▪ Programación ▪ Estadística Aplicada 			<ul style="list-style-type: none"> ▪ Tecnologías de Información 	

II. ORGANIZACIÓN SEMESTRAL							
Horas Dedicación Semanal (Cronológicas)		Docencia Directa	4,5	Trabajo Autónomo	3,5	Total	8,0
Detalle Horas Directas	Cátedra	Ayudantía	Laboratorio	Taller	Terreno	Exp. Clínica	Supervisión
	3,0	1,5	-	-	-	-	-

⁵⁸ Este campo sólo se completa en caso de carreras con programas semestrales.

⁵⁹ Clasificación del curso de acuerdo a la OCDE



III. APORTE AL PERFIL DE EGRESO

Al final del curso el estudiante podrá aplicar técnicas de ciencia de datos y de sus disciplinas asociadas (estadística y aprendizaje automático) para modelar y resolver problemas aplicados en diversos contextos.

IV. HABILIDADES PERFIL DE EGRESO (Relación)

De acuerdo con las habilidades descritas en el perfil de egreso, Habilidades para las Prácticas de la Ingeniería, la habilidad a desarrollar es:

- 4.3 Concebir sistemas para gestionar las operaciones, la calidad y confiabilidad y la cadena de abastecimiento, generando valor a las organizaciones, orientado por el uso eficiente del capital.
- 4.4 Diseñar sistemas para gestionar las operaciones, la cadena de abastecimiento, la calidad y confiabilidad, orientado por el uso eficiente del capital humano y recursos.
- 4.5 Conducir procesos de implementación de mejoras de los sistemas de gestión de operaciones, la cadena de abastecimiento, la calidad y confiabilidad, orientado por el uso eficiente del capital humano y recursos.

Competencias Genéricas:

CG6: Trabajo en equipo

IV. RESULTADOS DE APRENDIZAJE

1. Comprender los conceptos fundamentales de ciencia de datos y su relación con otras disciplinas como estadística y aprendizaje automático.
2. Implementar métodos de obtención de datos en distintos dominios de aplicación.
3. Diseñar modelos para el análisis de datos en diversos contextos de aplicación.
4. Aplicar modelos de análisis de datos teniendo en consideración el dominio que está siendo analizado.
5. Interpretar los resultados obtenidos mediante modelos de análisis de datos.



6. Comunicar los análisis realizados en un trabajo de análisis de datos en distintos dominios de aplicación.

V. ÁREAS TEMÁTICAS

1. Unidad I: Introducción

- 1.1. Definición de Ciencia de Datos
- 1.2. Metodologías
- 1.3. Relación con Estadística y Machine Learning
- 1.4. Aplicaciones

2. Unidad II: Obtención de datos

- 2.1. Data scraping
- 2.2. Repaso álgebra lineal
- 2.3. Grafos y procesamiento de redes
- 2.4. Procesamiento de lenguaje natural

3. Unidad III: Modelamiento estadístico y aprendizaje automático

- 3.1. Regresión y clasificación lineales
- 3.2. Modelamiento no lineal, cross-validation y regularización
- 3.3. Repaso de probabilidades y estadísticas
- 3.4. Pruebas de hipótesis y diseño de experimentos

4. Unidad IV: Técnicas de modelamiento avanzado

- 4.1. Aprendizaje supervisado
 - 4.1.1. Árboles de decisión
 - 4.1.2. Redes neuronales
- 4.2. Aprendizaje no supervisado
 - 4.2.1. Clustering
 - 4.2.2. Reducción de dimensionalidad

VI. ORIENTACIONES METODOLÓGICAS

1. Las metodologías sugeridas en la asignatura son clases de método combinado, es decir, clases expositiva con alternancia de trabajos de grupo de corta duración para responder preguntas.



2. Se utilizará la metodología activa de análisis de casos para desarrollar experiencias necesarias que permitan incorporar elementos prácticos/ teóricos con técnicas de ciencia de datos y de sus disciplinas asociadas.
3. Se realizarán mini talleres complementarios que se relacionan directamente con la materia, estos son desarrollados en equipos para fortalecer la comunicación, expresión oral y escrita, y trabajo en equipo.
4. Se realizarán ayudantías prácticas con ejercicios de programación e implementación.
5. Se entregarán talleres grupales para trabajo autónomo.

VII. ORIENTACIONES Y CRITERIOS PARA LA EVALUACIÓN

1. Se evaluará el conocimiento conceptual y metodológico mediante pruebas presenciales sumativas.
2. Se considera además la evaluación de talleres grupales autónomo.
3. Se realizarán evaluaciones de carácter formativo (procesual) lo que permitirá observar el proceso y progreso que están teniendo los y las estudiantes en relación con las unidades temáticas y sus respectivos aprendizajes esperados. De esta forma, el docente podrá introducir correcciones, añadir alternativas y reforzar ciertos aspectos para ayudar al estudiantado a mejorar su aprendizaje.
4. Asistencia y aprobación de la asignatura según reglamento de docencia de pregrado.

VIII. RECURSOS BIBLIOGRÁFICOS

Bibliografía mínima

- Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython, William McKinney, 2012, O'Reilly Media.
- Machine Learning for Hackers: Case Studies and Algorithms to Get You Started, Drew Conway & John White, 2012, O'Reilly Media.



Bibliografía complementaria

- An Introduction to Statistical Learning with Applications in R, Gareth James, Daniela Witten, Trevor Hastie & Robert Tibshirani, 7 ma edición, 2017, Springer, disponible en <http://www-bcf.usc.edu/~gareth/ISL/>
- Data Mining: Practical Machine Learning Tools and Techniques, Ian Witten, Eibe Frank & Mark Hall, 3 ra edición, 2011, Elsevier.
- Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking, Foster Provost & Tom Fawcett, 2013, O'Reilly Media.

Software:

Python y Knime