



EdV
Para estudiantes de Educación Básica y Media.
UNIVERSIDAD DE CHILE

PROGRAMA

Programación Científica para Proyectos de Inteligencia Artificial (IA)

FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS



Nombre del curso	Programación Científica para Proyectos de Inteligencia Artificial (IA)
Facultad Asociada	Facultad de Ciencias Físicas y Matemáticas
Nivele Educativos	I.º, II.º, III.º y IV.º medio

¿Qué aprenderemos en este curso?

Este curso tiene como propósito que las y los estudiantes desarrollen una comprensión integral y práctica del ciclo de vida de un proyecto de Inteligencia Artificial, fomentando la intuición y el pensamiento algorítmico para la resolución de problemas de ingeniería.

El programa está diseñado para que, al finalizar, los estudiantes sean capaces de manipular y visualizar datos, construir y evaluar modelos de aprendizaje supervisado y no supervisado. Además, aprenderán a optimizar su rendimiento mediante técnicas de selección de características y ajuste de hiperparámetros. De esta manera, estarán preparados para enfrentar los desafíos de una carrera en ciencia y tecnología con una perspectiva analítica y aplicada.

Objetivos de Aprendizaje

Explicar los fundamentos del lenguaje de programación Python en el contexto de la Ciencia de Datos.

Aplicar conceptos esenciales de estadística y técnicas de visualización para analizar y comunicar información relevante.

Analizar los desafíos de diseño, implementación y ejecución de un proyecto de Ciencia de Datos.

Seleccionar modelos de aprendizaje supervisado y no supervisado de acuerdo con las características de una tarea de Machine Learning.

Evaluar el rendimiento de modelos de Machine Learning mediante métricas adecuadas y argumentar la elección de dichas métricas.

Justificar el uso de métodos numéricos y técnicas de aproximación para problemas que no poseen una solución analítica cerrada.

Implementar técnicas de selección de características y optimización de hiperparámetros orientadas a mejorar el desempeño de modelos de Ciencia de Datos.

Día 1	Sesión 1	Introducción al curso. Operaciones y estructuras de datos fundamentales en el lenguaje de programación Python.
	Sesión 2	Ciclos y funciones en Python.
Día 2	Sesión 1	Estadística. Medidas de tendencia central, dispersión y relación entre variables. Diferencia entre correlación y causalidad. Definición de valores atípicos.
	Sesión 2	Estadística. Ley de los Grandes Números y Teorema Central del Límite. Diferencias entre el enfoque frecuentista y bayesiano.
Día 3	Sesión 1	Herramientas elementales para cargar y manipular datos en Python con Pandas. Formatos principales para guardar datos y sus diferencias.
	Sesión 2	Visualización de información. Gráficos (barras, histogramas, líneas, dispersión, etc.) y sus objetivos. Uso de colores e ilusiones ópticas.
Día 4	Sesión 1	Sesión práctica. Análisis exploratorio de datos.
	Sesión 2	Ingeniería de características. Creación, eliminación y modificación de variables categóricas y numéricas en datos tabulares.
Día 5	Sesión 1	Aprendizaje supervisado y no supervisado en Machine Learning. Dilema sesgo-varianza, subajuste y sobreajuste de modelos.
	Sesión 2	Estimación de parámetros en problemas de predicción con regresión lineal. Aplicación de regularización en modelos de Machine Learning.
Día 6	Sesión 1	Técnicas de aproximación de soluciones: intuición detrás del método del gradiente descendente y gradiente descendente estocástico. Aplicación a predicciones con modelos no lineales.
	Sesión 2	Modelo de regresión logística y análisis discriminante lineal (LDA) para aprendizaje supervisado.
Día 7	Sesión 1	Más modelos de aprendizaje supervisado: k-Nearest Neighbors (k-NN), Support Vector Machines (SVM), Naïve Bayes. Intuición detrás de las variantes de SVM y Naïve Bayes.
	Sesión 2	Modelos de aprendizaje supervisado basados en árboles: Decision Tree, Random Forest, XGBoost, LightGBM, CatBoost.
Día 8	Sesión 1	Métodos de evaluación para modelos de aprendizaje supervisado. Revisión de métricas, técnica de cross-validation y problema de Data Leakage.
	Sesión 2	Introducción al aprendizaje no supervisado. Modelos particionales (K-Means, DBSCAN) y jerárquicos (aglomerativo, divisivo).
Día 9	Sesión 1	Problema de maldición de la dimensionalidad. Reducción de dimensiones en los datos. Algoritmos PCA, UMAP y t-SNE.
	Sesión 2	Métodos de evaluación para modelos de aprendizaje no supervisado.

		Métricas supervisadas, no supervisadas y relativas.
Día 10	Sesión 1	Técnicas de selección de características y optimización de hiperparámetros.
	Sesión 2	Cierre del curso. Pipelines y Column Transformers en Scikit-Learn.

○