

Código	Nombre			
MDS7104	Aprendizaje de Máquinas			
Nombre en Inglés				
Machine Learning				
CT	Unidades Docentes	Horas de Cátedra	Horas Docencia Auxiliar	Horas de Trabajo Personal
6	10	3	1,5	5,5
Requisitos			Carácter del Curso	
MA3401/MA3403/EL4003/(CC3301, CC3201)			Obligatorio Magíster en Ciencia de Datos	
Resultados de Aprendizaje				
<p>Este curso enseña los fundamentos teóricos del aprendizaje de máquinas, su importancia en relación a otras ramas que requieren análisis de datos y su implementación en problemas reales.</p> <p>Al final del curso, el alumno deberá conocer las propiedades de distintos métodos de aprendizaje, como mínimos cuadrados, máxima verosimilitud, máximo a posteriori, y estimación Bayesiana, tanto para modelos paramétricos como no-paramétricos. Finalmente, el alumno deberá implementar las herramientas aprendidas a datos reales en problemas de regresión, predicción, clasificación, selección de modelos y reducción de dimensionalidad.</p>				

Metodología Docente	Evaluación General
<p>La metodología del curso comprenderá los siguientes ítems:</p> <ul style="list-style-type: none"> • Clases expositivas • Demostraciones de los métodos aprendidos • Clases auxiliares donde se guiará a los alumnos a implementar las técnicas vistas en cátedra • Realización de un proyecto individual 	<p>La evaluación será en aspectos tanto teóricos como prácticos durante el desarrollo del curso e incluirá:</p> <ul style="list-style-type: none"> • Participación en clase • Presentaciones y discusiones grupales en base a material de grupo de lectura • Tareas • Proyecto

Unidades Temáticas

Número	Nombre de la Unidad	Duración en Semanas
1	Inteligencia artificial vía aprendizaje de máquinas	1.5
Contenidos		Referencias a la Bibliografía
a) Introducción		Ver notas de referencia en cada contenido
Resultados de Aprendizajes de la Unidad		
- Conocer los fundamentos del aprendizaje de		

<ul style="list-style-type: none"> • Inteligencia Artificial, fundamentos filosóficos, agentes [IAMA, MA] • Sistemas expertos versus automatización [IAMA] • Definición de aprendizaje [LFD, ISL] • Ejemplo: Ajuste de curvas polinomiales • Aplicaciones reales y motivación <p>b) Definición de Aprendizaje de Máquinas [IAMA cap18]</p> <ul style="list-style-type: none"> • Taxonomía: Aprendizaje supervisado, no-supervisado, semi-supervisado, reforzado [LFD] • Sobreajuste / subajuste [LFD, ISL] • Entrenamiento, validación y test [LFD, ISL] • Navaja de Occam [LFD] 	<p>máquinas en relación a la computación clásica y a la inteligencia artificial</p> <p>- Entender la relación entre aprendizaje de máquinas y razonamiento probabilístico</p> <p>- Comprender cómo extraer información y manejar incertidumbre de datos con un enfoque probabilístico</p>	
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--

Número	Nombre de la Unidad	Duración en Semanas
2	Regresión: inferencia exacta y aproximada	3
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
<p>a) Regresión lineal</p> <ul style="list-style-type: none"> • Modelo lineal [ISL] • Mínimos cuadrados y regularización • Aprendizaje mediante optimización: gradiente descendente estocástico • Repaso de optimización y álgebra lineal • Regresión lineal multivariada • Regresión no lineal: polinomial, escalones, bases, <i>splines</i> [ISL] • Modelos lineales generalizados [MLPP, ISL] • Repaso de programación científica en Python 	<p>- Habilidad para formular un problema de regresión y lo resolverlo de forma determinística o probabilística</p> <p>- Resolver modelos intratables mediante métodos de Monte Carlo</p> <p>- Realizar predicción en base a datos observados tomando en cuenta conceptos de incertidumbre y sobreajuste</p>	<p>Ver notas de referencia en cada contenido</p>

<p>b) Aprendizaje mediante inferencia probabilística [BDA]</p> <ul style="list-style-type: none"> • Tres pasos: modelo, distribución posterior, evaluación. • Repaso de probabilidades • Modelos generativos: Gaussian discriminant analysis y <i>naïve Bayes</i> • Inferencia bayesiana • Probabilidad como medida de incertidumbre • Ejemplo: un modelo de un parámetro • Regularización mediante elección de distribución <i>a priori</i> 		
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--	--

Número	Nombre de la Unidad	Duración en Semanas
3	Selección y promedio de modelos [MSMA]	1
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
<ul style="list-style-type: none"> • Selección: Akaike IC, Bayesian IC, Negative log-predictive distribution [BDA, ESL (ch7)] • Promedio de modelos • Selección/ingeniería de características 	<p>- Evaluar, seleccionar y combinar distintas estructuras de modelos en base a índices de desempeño, incluso cuando la cantidad de estructuras es infinita</p>	<p>Ver notas de referencia en cada contenido</p>

Número	Nombre de la Unidad	Duración en Semanas
4	Clasificación	2.5
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
<ul style="list-style-type: none"> • Regresión logística • Análisis de discriminante lineal • Clasificación lineal: binaria y multiclase • K-vecinos más cercanos [ISL] • <i>Naive Bayes</i> [ISL] 	<p>-Representar probabilísticamente el problema de clasificación y relacionarlo con enfoques deterministas.</p>	<p>Ver notas de referencia en cada contenido</p>

<ul style="list-style-type: none"> • Máquinas de soporte vectorial [LWK] • Random Forests • Medidas de desempeño: Error cuadrático medio, entropía cruzada, validación cruzada, curvas ROC, curvas de precisión-recall, F-score, matrices de confusión, etc. 	<p>-Conocer la teoría de aprendizaje estadístico y su relación con las máquinas de soporte vectorial</p> <p>-Implementar máquinas de soporte vectorial</p>	
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------	--

Número	Nombre de la Unidad	Duración en Semanas
5	Redes neuronales [ISL, MLPR, DL]	2
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
<ul style="list-style-type: none"> • Definición, perceptrón, adaline • Backpropagation, ADAM • Tipos de redes (CNN, RNN, FFNN). • Funciones de activación • Autoencoders y redes generativas adversariales • Métricas de evaluación • Aprendizaje profundo: convolución, submuestreo, dropout, normalización por batches. • Ejemplos de redes convolutivas y residuales. 	<p>-Conocer las estructuras estándar de redes neuronales y técnicas de entrenamiento</p> <p>-Implementar estructuras actuales de redes neuronales usando toolboxes disponibles</p> <p>-Nociones sobre el estado del arte en redes bayesianas</p>	Ver notas de referencia en cada contenido

Número	Nombre de la Unidad	Duración en Semanas
7	Reducción de dimensionalidad	1
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
<ul style="list-style-type: none"> • Análisis de componentes principales: lineal, probabilístico y de kernel • LDA/ANOVA/ICA • Algoritmo de visualización t-SNE 	-Conocer la importancia de las técnicas de reducción de dimensionalidad para preprocesamiento de datos, visualización e interpretación.	Ver notas de referencia en cada contenido

	-Dominar distintas técnicas de reducción de dimensionalidad y saber cuándo usarlas	
--	------------------------------------------------------------------------------------	--

Número	Nombre de la Unidad	Duración en Semanas
8	Clustering	2
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
<ul style="list-style-type: none"> • K-medias • DBSCAN: <i>Density-based spatial clustering of applications with noise</i> • <i>Kernel density estimation</i> [MLPP] • Factor analysis • Mezcla de gaussianas: El algoritmo <i>expectation maximisation</i> [MLPP] • Métricas de clustering 	<ul style="list-style-type: none"> -Entender la diferencia entre métodos paramétricos y no-paramétricos -interpretar <i>clustering</i> como el diseño de un modelo generativo -ocupar un enfoque variacional para resolver modelos intratables (inferencia aproximada) 	Ver notas de referencia en cada contenido

Bibliografía General
<p>[IAMA] S. Russell and P. Norvig, <i>Artificial Intelligence: A Modern Approach</i>, 1994. Prentice Hall.</p> <p>[MA] P. Domingues, <i>The Master Algorithm</i>, 2015. Basic Books.</p> <p>[LFD] Y. Abu-Mostafa, M. Magdon-Ismael and H. Lin, <i>Learning from data</i>, 2012. AMLBook</p> <p>[ISL]: G. James, D. Witten, T. Hastie and R. Tibshirani, <i>Introduction to statistical learning</i>, 2013. Springer.</p> <p>[BDA]: A. Gelman, J. Carlin, H. Stern, D. Dunson, A. Vehtari, and D. Rubin, <i>Bayesian Data Analysis</i>, 2013. CRC Press.</p> <p>[ITILA]: D. MacKay, <i>Information theory, inference and learning algorithms</i>, 2003. Cambridge.</p> <p>[MSMA]: G. Claeskens and N. L. Hjort, <i>Model selection and model average</i>, 2008. Cambridge.</p> <p>[PRML]: C. Bishop, <i>Pattern recognition and machine learning</i>, 2006. Springer.</p> <p>[GP4ML]: C. Rasmussen and C. Williams, <i>Gaussian processes for machine learning</i>, 2006. MIT.</p> <p>[MLPP]: K. Murphy, <i>Machine learning: A probabilistic perspective</i>, 2012. MIT.</p> <p>[ESL]: T. Hastie, R. Tibshirani, J. Friedman, <i>Elements of statistical learning</i>, 2009. Springer.</p> <p>[LWK]: B. Schölkopf and A. Smola, <i>Learning with kernels</i>, 2002. MIT.</p> <p>[UDL]: Y. Gal, <i>Uncertainty in Deep Learning</i>, 2015. http://mlg.eng.cam.ac.uk/yarin/thesis/thesis.pdf</p> <p>[VI]: D. M. Blei, A. Kucukelbir and J. D. McAuliffe, <i>Variational Inference: A Review for Statisticians</i>, 2015. https://arxiv.org/abs/1601.00670.</p>

Vigencia desde:	Diciembre 2018
Elaborado por:	Felipe Tobar
Revisado por:	Pablo Estévez