

PROGRAMA DE CURSO

Código	Nombre			
MDS7204	Aprendizaje de Máquinas Avanzado			
Nombre en Inglés				
Advanced Machine Learning				
SCT	Unidades Docentes	Horas de Cátedra	Horas Docencia Auxiliar	Horas de Trabajo Personal
6	10	3		7
Requisitos			Carácter del Curso	
MDS7104 (o equivalente) y MDS7101 (o equivalente)			Electivo Magíster en Ciencia de Datos	
Resultados de Aprendizaje				
<p>Este curso enseña técnicas avanzadas de aprendizaje de máquinas, relacionadas con series de tiempo, modelos generativos, filtrado, modelos bayesianos no paramétricos e inferencia aproximada. El enfoque teórico del curso es basado probabilidades, optimización y estadística, mientras que el aspecto práctico considera programación en Python.</p> <p>Al final del curso, el alumno deberá conocer el estado del arte en aprendizaje de máquinas probabilístico y deberá ser capaz de aplicar las herramientas aprendidas a un problema real de la industria o su propia investigación.</p>				

Metodología Docente	Evaluación General
<p>La metodología del curso comprenderá los siguientes ítems:</p> <ul style="list-style-type: none"> • Asignación de artículos y temas de estudio por parte del instructor • Clases expositivas, donde se presentará la teoría estudiada e implementaciones de los mismos métodos • Discusión de los temas expuestos por parte de los alumnos, cuerpo docente e invitados • Realización de tareas al final de cada unidad del curso 	<p>La evaluación será en aspectos tanto teóricos como prácticos durante el desarrollo del curso e incluirá:</p> <ul style="list-style-type: none"> • Asistencia y participación en las clases expositivas • Participación en discusión de las técnicas y discusiones grupales en base a material de grupo de lectura • Tareas prácticas al final de cada unidad • Proyecto final del curso: aplicación práctica, incorporación a las temáticas de investigación, o bien desarrollo de nueva teoría.

Unidades Temáticas

Número	Nombre de la Unidad	Duración en Semanas
1	Series de Tiempo	3
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
<ul style="list-style-type: none"> - Modelos autoregresivos - Filtros lineales - Estimación espectral - Filtro de partículas 	Al final de la unidad se espera que el estudiante pueda usar distintas herramientas para modelar series de tiempo mediante el uso de observaciones ruidosas	[BJR] [smc]

Número	Nombre de la Unidad	Duración en Semanas
2	Bayesian nonparametrics	3
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
<ul style="list-style-type: none"> - Procesos gaussianos --- Diseño de covarianzas --- aproximaciones sparse --- representación espectral --- Multioutput GPs - Proceso de Dirichlet --- Construcción stick-breaking --- Proceso del restaurant chino --- Proceso del buffet indio 	Al final de la unidad se espera que el estudiante comprenda la naturaleza de los modelos no paramétricos, sus ventajas, costos de implementación y ajuste.	[gp] [bnp]

Número	Nombre de la Unidad	Duración en Semanas
3	Inferencia variacional	3
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
<ul style="list-style-type: none"> - Motivación 1: Latent Dirichlet allocation - Motivación 2: Bayesian mixture of Gaussians - Motivación 3: Autoencoders variacionales - Inferencia Variacional --- Laplace & Expectation Maximisation --- Evidence lower bound --- La familia de campo --- Coordinate-ascent VI --- Stochastic/Blackbox VI 	Al final de la unidad se espera que el estudiante comprenda cuando aproximar un problema de inferencia y qué técnica de usar. Además, el estudiante debe entender la relación entre compresión y representación	[murphy] [vi] [mackay] [bishop]

Número	Nombre de la Unidad	Duración en Semanas
4	Tendencias en Aprendizaje de Máquinas	3
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
<ul style="list-style-type: none"> - Aprendizaje Profundo Bayesiano - Transporte Óptimo - Flujos normalizadores 	<p>Al final de la unidad se espera que el estudiante comprenda y domine las tendencias de la investigación en el área que puedan servir para subsecuente desarrollo de su propia investigación. En particular, el estudiante podrá entender la conexión entre el área de aprendizaje de máquinas y otras como probabilidades, estadística y optimización</p>	<p>[BDL] [OT] [NF]</p>

Bibliografía General

[PyTorch] <https://www.udemy.com/practical-deep-learning-with-pytorch/>

[Pyro] <http://pyro.ai/examples/>

[bnp. 1] <https://papers.nips.cc/paper/1956-the-infinite-hidden-markov-model.pdf>

[bnp. 2] <http://www.tamarabroderick.com/tutorials.html>

[bnp. 3] <http://stat.columbia.edu/~porbanz/talks/nipstutorial.pdf>

[bnp. 4] <https://www.stats.ox.ac.uk/~teh/npbayes.html>

[bnp. 5] S. Gershman and D. Blei. A tutorial on Bayesian nonparametric models

[vi. 1] <https://arxiv.org/pdf/1601.00670.pdf>

[NF] <https://arxiv.org/pdf/1505.05770v6.pdf>

[smc. 1] A. Doucet, N. de Freitas and N. Gordon. Sequential Monte Carlo Methods in Practice. Springer, 2001.

[ot.1] <https://arxiv.org/abs/1803.00567>

[ot.2] <https://papers.nips.cc/paper/4927-sinkhorn-distances-lightspeed-computation-of-optimal-transport>

[BDL.1] <http://mlg.eng.cam.ac.uk/yarin/thesis/thesis.pdf>

[BDL.2] Y. Bengio, Deep Architectures for AI.

[murphy] K. Murphy. Machine Learning: A Probabilistic Perspective. MIT Press, 2012.

[mackay] D. Mackay. Information Theory, Inference, and Learning Algorithms. Cambridge University Press. 2003.

[bishop] C. Bishop. Pattern Recognition and Machine Learning. Springer, 2006.

[gp.1] C. Rasmussen and C. Williams, Gaussian Processes for Machine Learning. MIT Press 2006.

[gp.2] <http://www.jmlr.org/papers/v6/quinonero-candela05a.html>

[gp.3] http://www2.aueb.gr/users/mtitsias/papers/aistats_varGP.pdf

[BJR] G. Box, G. Jenkins and G. Reinsel. Times Series Analysis: Forecasting and Control. Wiley, 2008.



FACULTAD DE CIENCIAS
FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE

Vigencia desde:	Primavera 2022
Elaborado por:	Felipe Tobar
Revisado por:	Comité Académico MDS