

PROGRAMA DE CURSO

Código	Nombre			
MA5309	Aprendizaje de Máquinas Avanzado			
Nombre en Inglés				
Advanced Machine Learning				
SCT	Créditos	Horas de Cátedra	Horas Docencia Auxiliar	Horas de Trabajo Personal
6	6	2	0	8
Requisitos			Carácter del Curso	
MA5203 (o equivalente) y autorización			Electivo de magister y doctorado	
Resultados de Aprendizaje				
<p>Este curso enseña técnicas avanzadas de aprendizaje de máquinas, relacionadas con computación científica, modelos no paramétricos e inferencia aproximada. El enfoque teórico del curso es basado probabilidades y optimización, 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 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 dictadas por los alumnos, 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 un proyecto teórico o aplicado 	<p>La evaluación será en aspectos tanto teóricos como prácticos durante el desarrollo del curso e incluirá:</p> <ul style="list-style-type: none"> • Presentaciones: teoría e implementación. • Asistencia y participación en presentaciones de otros alumnos • Participación en discusión de las técnicas y discusiones grupales en base a material de grupo de lectura • Proyecto final del curso: aplicación práctica, incorporación a las temáticas de investigación del propio alumno, o bien desarrollo de nueva teoría.

Unidades Temáticas

Número	Nombre de la Unidad	Duración en Semanas
1	Advanced scientific computing	2
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
a) PyTorch -Numpy, Scipy, Pandas, Scikit-Learn - Fundamentals: tensors - Variables and gradients - Regression examples - CNN, RNNs and LSTMs b) Pyro - Models and distributions - Inference - examples		[PyTorch] [Pyro]

Número	Nombre de la Unidad	Duración en Semanas
2	Graphical models	1
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
- Structured representations - Directed and undirected graphs - Inference and marginalisation - Learning - Bayesian networks - Applications		[grph. mod.] [prob.]

Número	Nombre de la Unidad	Duración en Semanas
3	Bayesian nonparametrics	3
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
- Motivation - Dirichlet process - Stick-breaking construction - Chinese restaurant process - Indian buffet process - Infinite hidden Markov model - Hierarchical BNP		[bnp]

Número	Nombre de la Unidad	Duración en Semanas
4	Approximate inference I: Monte Carlo	2

Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
<ul style="list-style-type: none"> - Sampling techniques - MCMC: Metropolis and Gibbs - Transdimensional MCMC - Hybrid MC - Slice sampling - Particle filters 		[murphy] [mcmc] [mackay] [bishop]

Número	Nombre de la Unidad	Duración en Semanas
5	Approximate inference II: Variational Inference	2
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
<ul style="list-style-type: none"> - Problem statement - Laplace & Expectation Maximisation - Evidence lower bound - The mean-field variational family - Coordinate ascent VI - Example: Bayesian mixture of Gaussians - VI with exponential families - Stochastic VI - Black-box VI 		[murphy] [vi] [mackay] [bishop]

Número	Nombre de la Unidad	Duración en Semanas
6	Computational Optimal Transport	2
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
<ul style="list-style-type: none"> - Theoretical foundations - Algorithmics - Wasserstein PCA - Entropic regularisation: Sinkhorn - Empirical Barycentres - Applications 	To become familiar with the AI final de la unidad se espera que el estudiante comprenda las ventajas y desventajas de la formulación multicapa en redes neuronales y sus distintas arquitecturas	[ot]

Número	Nombre de la Unidad	Duración en Semanas
7	Gaussian Processes	4
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
<ul style="list-style-type: none"> - Constructions: Kolmogorov, RBFs and NNs - Sparse GPs: SOR, FITC, PITC, VI - Kernel design 	To understand the theory of modern implementations and approximations of GP models regarding computational	[gp]

- Multioutput GPs - GP latent variable model - Deep GPs - Bayesian spectral estimation	efficiency, multi-input/multi-output models, and their relationship to spectral estimation.	
---	---	--

Número	Nombre de la Unidad	Duración en Semanas
8	Bayesian Optimisation	2
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
- Motivation and problem statement - Acquisition functions - Applications		[bayesopt]

Número	Nombre de la Unidad	Duración en Semanas
9	Bayesian Deep Learning	1
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
- Bayesian neural networks - Stochastic regularisation - Applications		[bayesdl]

Número	Nombre de la Unidad	Duración en Semanas
10	Natural Language Processing	1
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
- Language modelling - Classification and clustering - Part-of-speech tagging - Sentiment analysis - Summarisation	Al final de la unidad se espera que el estudiante comprenda las ventajas y desventajas de la formulación multicapa en redes neuronales y sus distintas arquitecturas	[nlp]

Bibliografía General
[PyTorch] https://www.udemy.com/practical-deep-learning-with-pytorch/ [Pyro] http://pyro.ai/examples/
[grph. mod. 1] http://ai.stanford.edu/~paskin/gm-short-course/ [grph. mod. 2] https://www.cs.ubc.ca/~murphyk/Papers/intro_gm.pdf
[prob.] D. Bertsekas and J. Tsitsiklis (2002). Introduction to Probability. Athena Scientific, Belmont, Mass.
[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>

[mcmc. 1] http://vcla.stat.ucla.edu/old/MCMC/MCMC_tutorial.htm

[mcmc. 2] http://www.cs.ubc.ca/~arnaud/andrieu_defreitas_doucet_jordan_intromontecarlomachinel_earning.pdf

[mcmc. 3] <https://www.stat.fi/isi99/proceedings/arkisto/varasto/gree0167.pdf>

[mcmc. 4] <https://twiecki.github.io/blog/2015/11/10/mcmc-sampling/>

[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>

[bayesopt.0] <https://arxiv.org/pdf/1012.2599.pdf>

[bayesopt.1] <https://arxiv.org/abs/1807.02811>

[bayesopt.2] <http://pyro.ai/examples/bo.html>

[nlp.0] <https://cs.nyu.edu/courses/fall11/CSCI-GA.3033-001/>

[nlp.1] <https://www.amazon.com/Language-Processing-Prentice-Artificial-Intelligence/dp/0131873210>

[nlp.1] foundations-of-statistical-natural-language-processing

[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] <https://papers.nips.cc/paper/2540-gaussian-process-latent-variable-models-for-visualisation-of-high-dimensional-data.pdf>

[gp.4] 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.

[B] Y. Bengio, Deep Architectures for AI.

Vigencia desde:	Primavera 2018
Elaborado por:	Felipe Tobar
Revisado por:	