

PROGRAMA DE CURSO

Código	Nombre		
MA5309	Aprendizaje de Máquinas Avanzado		
Nombre en Inglés			
Advanced Machine Learning			
SCT	Horas de Cátedra	Horas Docencia Auxiliar	Horas de Trabajo Personal
6	2		8
Requisitos			Carácter del Curso
MA5203 (o equivalente) y autorización			Electivo de magister y doctorado
Resultados de Aprendizaje			
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. 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.			

Metodología Docente	Evaluación General
La metodología del curso comprenderá los siguientes ítems: • 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	La evaluación será en aspectos tanto teóricos como prácticos durante el desarrollo del curso e incluirá: • 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	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
2	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 - Gaussian processes --- Covariance design --- Sparse approximations --- Spectral representation --- Multioutput GPs		[bnp]

Número	Nombre de la Unidad	Duración en Semanas
3	Approximate inference I: Monte Carlo	2
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
- 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
4	Approximate inference II: Variational Inference	2
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
- Problem statement - Laplace & Expectation Maximisation		[murphy] [vi]

- Evidence lower bound - The mean-field variational family - Coordinate ascent VI - Example: Bayesian mixture of Gaussians - VI with exponential families - Stochastic/Blackbox VI		[mackay] [bishop]
---	--	----------------------

Número	Nombre de la Unidad	Duración en Semanas
5	Computational Optimal Transport	2
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
- Theoretical foundations - Algorithmics - Wasserstein PCA - Entropic regularisation: Sinkhorn - Empirical Barycentres - Applications		[ot]

Número	Nombre de la Unidad	Duración en Semanas
6	Trends on Machine Learning Research	5
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
- The information bottleneck method - Determinantal point processes - Dynamic time warping - Random matrices - Bayesian deep learning - Bayesian optimisation		[IB] [DPP] [DTW] [BDL] [BO]

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_intromontecarlolearning.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>
- [BO.0] <https://arxiv.org/pdf/1012.2599.pdf>
- [BO.1] <https://arxiv.org/abs/1807.02811>
- [BO.2] <http://pyro.ai/examples/bo.html>
- [DTW.1] <https://link.springer.com/content/pdf/10.1007/s10115-004-0154-9.pdf>
- [DTW.2] <https://www.aai.org/Papers/Workshops/1994/WS-94-03/WS94-03-031.pdf>
- [RM.1] <https://papers.nips.cc/paper/6857-nonlinear-random-matrix-theory-for-deep-learning.pdf>
- [RM.2] <https://arxiv.org/pdf/1702.05419.pdf>
- [RM.3] <https://arxiv.org/pdf/1805.11917.pdf>
- [IB.1] https://iopscience.iop.org/article/10.1088/1742-5468/ab3985/pdf?casa_token=rWo9eHGP5cgAAAAA:58IFs9HTS92sQKxqk262NZfg4RiYQ2zy7cCdFtJX_upoFFXcRnO
- [IB.2] <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7133169>
- [IB.3] <https://arxiv.org/pdf/2004.14941.pdf>
- [IB.4] <https://arxiv.org/abs/1703.00810>
- [IB.5] <https://arxiv.org/abs/physics/0004057>
- [IB.6] <https://arxiv.org/abs/1503.02406>
- [DPP] <https://arxiv.org/abs/1207.6083>
- [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] <https://papers.nips.cc/paper/2540-gaussian-process-latent-variable-models-for-visualisation-of-high-dimen>
- [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.

Vigencia desde:	Primavera 2020
Elaborado por:	Felipe Tobar
Revisado por:	José Soto - Jefe Docente