

PROGRAMA DE CURSO

Código	Nombre			
MA5203	Aprendizaje de Máquinas Probabilístico			
Nombre en Inglés				
Probabilistic Machine Learning				
SCT	Unidades Docentes	Horas de Cátedra	Horas Docencia Auxiliar	Horas de Trabajo Personal
	10	3	2	5
Requisitos			Carácter del Curso	
MA2001 Cálculo en Varias Variables MA3401 Probabilidades (alumnos DIM) MA3403 Probabilidades y Estadística (otros departamentos)			Electivo para especialidad y Magister	
Resultados de Aprendizaje				
Este curso enseña los fundamentos del enfoque probabilístico para aprendizaje de máquinas, su importancia en relación a otras ramas que requieren análisis de datos y su relación con métodos determinísticos. Al final del curso, el alumno deberá conocer las propiedades de distintos métodos de aprendizaje probabilístico, cómo máxima verosimilitud, máximo a posteriori, y estimación Bayesiana, tanto para modelos paramétricos como no-paramétricos. Adicionalmente, el alumno deberá aplicar las herramientas aprendidas a datos reales en problemas de regresión, predicción y clasificación.				

Metodología Docente	Evaluación General
La metodología del curso comprenderá los siguiente ítems: • Clases expositiva de 90 minutos • Demostraciones de las métodos aprendidos • Clases auxiliares donde se guiará a los alumnos a implementar las técnicas vistas en cátedra • Talleres de lectura.	La evaluación será en aspectos tanto teóricos como prácticos durante el desarrollo del curso e incluirá: • Participación en clase • Presentaciones y discusiones grupales en base a material de grupo de lectura • Tareas

Unidades Temáticas

Número	Nombre de la Unidad	Duración en Semanas
1	Introducción a Inferencia Bayesiana	1.5
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
a) Aprendizaje de máquinas: - Orígenes y conexión con otras disciplinas - Tipos de aprendizaje - ¿Cuándo usar aprendizaje de máquinas? b) Aprendizaje mediante inferencia - Repaso probabilidades - Densidad condicional - Teorema de Bayes - Prior, verosimilitud y posterior - Enfoque Bayesiano versus frecuentista	Al final de la unidad se espera que el estudiante: - Conozca los fundamentos del aprendizaje de máquinas y su importancia - Entienda la relación entre aprendizaje de máquinas y razonamiento probabilístico - Comprenda cómo extraer información y manejar incertidumbre de datos con un enfoque Bayesiano	Barber: [c1, c13] Murphy: [c1, c2] Mackay: [c28]

Número	Nombre de la Unidad	Duración en Semanas
2	Regresión Lineal	2.5
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
a) Métodos deterministas: - Mínimos cuadrados - Regularización de Tíjonov b) Métodos probabilísticos: - La distribución normal - Máxima verosimilitud - Máximo a posteriori - Regresión Bayesiana c) Predicción	Al final de la unidad se espera que el estudiante: - Formule un problema de regresión y lo resuelva de forma determinística o probabilística - Entienda las diferencias, virtudes y desventajas de ambos enfoques - Realice predicción en base a datos observados tomando en cuenta conceptos de incertidumbre y sobreajuste	Bishop: [c2, c3] Barber: [c8, c18] RW: [c2]

Número	Nombre de la Unidad	Duración en Semanas
3	Modelos no lineales: diseño, selección y combinación	3
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
a) Modelos lineales en los parámetros - selección de características b) Selección y combinación de modelos - ¿Máximizarse o integrar? - Ley de Parsimonia asociada al enfoque Bayesiano - Criterio de información Bayesiano c) Regresión logística - Aproximación de Laplace	Al final de la unidad se espera que el estudiante: - Entienda los desafíos en cuanto al diseño de los modelos no lineales - Diseñe, evalúe y combine distintas estructuras de modelos no lineales en base a datos observados - Aproxime la solución del problema de regresión cuando esta no tiene forma cerrada.	Bishop: [c3.4, c4] Barber: [c12, c18] Mackay: [c28] Murphy: [c8] RW: [c3.1-c3.2]

Número	Nombre de la Unidad	Duración en Semanas
4	Métodos basados en kernels	4
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
a) Diseño usando el truco del kernel - Espacios de Hilbert de kernel reproductivo - Máquinas de soporte vectorial para regresión y clasificación b) Procesos Gaussianos - Predicción con y sin ruido de observación - Inferencia - Comparación de mínimos locales - Diseño de funciones de covarianza - Costo computacional y aproximaciones	Al final de la unidad se espera que el estudiante: - Comprenda la diferencia entre modelos paramétricos y no-paramétricos para interpolación, clasificación, predicción. - Implemente algoritmos de estimación no paramétricos, tanto determinísticos como probabilísticos, mediante la generalización de datos reales observados. - Aplique dichas herramientas de estimación a series de tiempo y clasificación binaria. - Discrimine entre las soluciones encontradas en base a su verosimilitud	Murphy: [c14, c15] RW: [c2, c4, c5] SS: [c2, c7, c9]

Número	Nombre de la Unidad	Duración en Semanas
5	Inferencia aproximada	4
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
a) Motivación: - Clasificación usando procesos Gaussianos y ranking probabilístico b) Métodos de muestreo - Integración de Monte Carlo - Muestreo por rechazo e importancia - Cadenas de Markov Monte Carlo · Metrópolis-Hastings · Gibbs Sampling c) Método de propagación de esperanza	Al final de la unidad se espera que el estudiante: - Entienda la necesidad de inferencia aproximada en modelos que no admiten resolución en forma cerrada. - Formule el problema de aprendizaje estadístico como uno de extraer muestras de una distribución de probabilidad y lo resuelva usando métodos de Monte Carlo - Comprenda la diferencia entre: 1) solución aproximada a problema exacto y 2) solución exacta a problema aproximado	Bishop: [c10.7, c11.1] Barber: [c28.8, c27.7] Murphy: [c23, c24] RW: [c3]

Bibliografía General
[Murphy] K. Murphy. <i>Machine Learning: A Probabilistic Perspective</i> . MIT Press, 2012.
[Barber] D. Barber. <i>Bayesian Reasoning and Machine Learning</i> . Cambridge University Press 2012.
[Bishop] C. Bishop. <i>Pattern Recognition and Machine Learning</i> . Springer, 2006. □
[Mackay] D. Mackay. <i>Information Theory, Inference, and Learning Algorithms</i> . Cambridge University Press. 2003.
[RW] C. Rasmussen and C. Williams, <i>Gaussian Processes for Machine Learning</i> . MIT Press 2006.
[SS] B. Schölkopf and A. Smola. <i>Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond</i> . MIT Press 2002.

Vigencia desde:	Otoño 2016
Elaborado por:	Felipe Tobar
Revisado por:	Aris Daniilidis Jefe Docente y el CTD del DIM.