

PROGRAMA DE CURSO

Código	Nombre			
EL 7006	Redes Neuronales y Teoría de Información para el Aprendizaje			
Nombre en Inglés				
Neural Networks and Information Theoretic Learning				
SCT	Unidades Docentes	Horas de Cátedra	Horas Docencia Auxiliar	Horas de Trabajo Personal
6	10	3	2	5
Requisitos			Carácter del Curso	
EL4006 Inteligencia Computacional (AD para alumnos de doctorado)			Electivo de Línea de Especialización Electivo del Magister en Ciencias de la Ingeniería, mención Eléctrica Electivo del Doctorado en Ingeniería Eléctrica	
Resultado de Aprendizaje del Curso				
Al final del curso se espera que el estudiante demuestre que: <ul style="list-style-type: none"> • Aplica y evalúa técnicas avanzadas de redes neuronales artificiales, teoría de información y métodos de kernels en problemas de reconocimiento de patrones, análisis de datos, y procesamiento temporal de señales. 				

Metodología Docente	Evaluación General
<p>La metodología de trabajo será activo-participativa, en donde se desarrollarán:</p> <ul style="list-style-type: none"> • Clases expositivas. • Tareas. • Proyectos. <p>Si el curso se dicta en forma tutorial entonces en vez de clases expositivas se harán reuniones periódicas con los alumnos para monitorear el grado de avance de los proyectos.</p>	<p>La evaluación permitirá que los estudiantes demuestren los resultados de aprendizaje alcanzadas en los distintos momentos del proceso de enseñanza, siendo éstos:</p> <ul style="list-style-type: none"> • Controles. • Tareas. • Proyectos <p>El examen dará cuenta del resultado de aprendizaje del curso. Si el curso se dicta en forma tutorial, entonces no hay controles, y el examen se reemplaza por la presentación final de proyecto.</p>

Unidades Temáticas

Número	Nombre de la Unidad	Duración en Semanas	
1	Procesamiento Temporal con Redes Neuronales	5 Semanas	
Contenidos		Resultados de Aprendizaje de la Unidad	Referencias a la Bibliografía
1. Redes neuronales prealimentadas. Redes TDNN de retardo temporal. Filtro de memoria Gamma. Redes con línea de retardos. 2. Redes neuronales recurrentes y redes neuronales convolucionales. Teorema de aproximación universal. Redes de Jordan y Elman. Retropropagación a través del tiempo. Redes FIR. Aprendizaje recurrente en tiempo real (<i>Real time recurrent learning</i>). Modelo LSTM: Long Short Term Memory. 3. Redes neuronales auto-organizativas recurrentes. Mapas de Kohonen y Gas Neural temporales. Memorias gamma. 4. Análisis de series de tiempo. Metodología Box-Jenkins para modelos ARIMA. Métodos de reconstrucción de espacio de estado. Teorema de Takens. Estimación de parámetros de embedding: Falsos vecinos más cercanos, información mutua promedio. Algoritmos de Basis Pursuit y Matching Pursuit.		Al final de la unidad se espera que el estudiante demuestre que: <ol style="list-style-type: none"> 1. Comprenda los paradigmas de redes neuronales artificiales prealimentadas y recurrentes. 2. Simule redes neuronales artificiales. 3. Aplique redes neuronales al procesamiento de datos temporales y secuenciales. 4. Evalúe el desempeño de redes neuronales artificiales aplicadas a series de tiempo. 	[3] Cap. 6, 9-10 [4] Cap. 15 [7] Cap. 3,9

Número	Nombre de la Unidad	Duración en Semanas	
2	Teoría de Información para el Aprendizaje	6 Semanas	
Contenidos		Resultados de Aprendizaje de la Unidad	Referencias a la Bibliografía
1. Conceptos básicos de teoría de información: Entropía, entropía conjunta, condicional y relativa. Divergencia. Información mutua. Estimadores de entropía e información mutua. 2. Conceptos de teoría de información para el aprendizaje. Principios de optimización. Criterio unificado de aprendizaje. Entropía de Renyi. Propiedad de		Al final de la unidad se espera que el estudiante: <ol style="list-style-type: none"> 1. Comprenda los fundamentos de teoría de información para el aprendizaje 2. Simule algoritmos de aprendizaje basados en teoría de información 3. Utilice teoría de información para resolver problemas de aprendizaje tales como 	[1] Cap. 1-3,5-8, 10-11 [4] Cap. 10 [8] Cap. 1-2

<p>convolución de Gaussianas. Estimación no-paramétrica. Ventanas de Parzen. Potencial de información. Información mutua cuadrática. Divergencia de Cauchy-Schwarz y Euclideana.</p> <p>3. Criterio mínima entropía del error (MEE). Entrenamiento de clasificadores con MEE. Medidas de divergencia e información cuadrática mutua. Selección de características. Clustering.</p> <p>4. Principio de información relevante. Aplicación a cuantización vectorial, clustering y curvas principales.</p> <p>5. Correntropía. Propiedades. Coeficiente de correntropía. Test de no-linealidad. Relación con potencial de información y EHKR. Densidad espectral de correntropía. Ejemplos de análisis de series de tiempo.</p> <p>6. Análisis y aprendizaje de redes neuronales con métodos de teoría de información. Plano de información. Principio de cuello de botella de información.</p>	<p>análisis de series de tiempo, clasificación, clustering, cuantización vectorial y selección de características</p> <p>4. Evalué el desempeño de algoritmos de teoría de información y compare sus resultados con métodos alternativos</p>	
---	--	--

Número	Nombre de la Unidad	Duración en Semanas
3	Teoría de Información y Métodos de Kernels	4 Semanas
Contenidos	Resultados de Aprendizaje de la Unidad	Referencias a la Bibliografía
<p>1. Definición de kernels. Mapa de kernel reproductor. Espacio de Hilbert con kernel reproductor (EHKR). Mercer kernel. Propiedades de kernels. Regularización. Teorema de representación. Kernels invariantes a la traslación y kernels periódicos.</p> <p>2. Relación entre EHKR y teoría de información para el aprendizaje. Estimación no-paramétrica de funciones de densidad de probabilidad usando métodos de kernels. Relación entre EHKR, potencial de información cruzado</p>	<p>Al final de la unidad se espera que el estudiante:</p> <ol style="list-style-type: none"> 1. Comprenda métodos de aprendizaje y filtraje basados en kernels 2. Simule métodos de aprendizaje y filtraje basados en kernels 3. Utilice métodos de kernels para el filtraje adaptivo de señales y procesamiento de datos. 4. Contraste el desempeño de métodos de kernels con métodos alternativos 	<p>[2] Cap. 1-2 [5] Cap. 2 y 4</p>

<p>y funciones de costo usadas en teoría de la información.</p> <p>3. Filtros lineales adaptivos: LMS (<i>least mean square</i>), RLS (<i>recursive least squares</i>), RLS extendido. Uso de EHKR para diseñar filtros adaptivos no-lineales. Algoritmo kernel-LMS. Selección de parámetros. Criterio de novedad. Kernel LMS normalizado. Kernel ADALINE. Algoritmo kernel-RLS. Aplicación a series de tiempo.</p>		
---	--	--

Bibliografía	
<p><u>Bibliografía Básica</u></p> <p>[1] PRINCIPE, J.C., <i>Information Theoretic Learning: Renyi's Entropy and Kernel Perspectives</i> Springer-Verlag, 2010.</p> <p>[2] LIU, W., PRINCIPE, J.C., HAYKIN, S., <i>Kernel Adaptive Filtering: A Comprehensive Introduction</i>. Wiley, 2010.</p> <p>[3] GOODFELLOW, I., BENGIO, Y., COURVILLE, A., <i>Deep Learning</i>, MIT Press, 2016.</p> <p>[4] HAYKIN, S., <i>Neural Networks and Learning Machines</i>, 3rd edition, Pearson, 2009.</p> <p><u>Bibliografía Complementaria</u></p> <p>[5] SCHÖLKOPF, B., SMOLA, A.J., <i>Learning with Kernels</i>, MIT Press, 2002.</p> <p>[6] THEODORIDIS, S., KOUTROUMBAS, K. <i>Pattern Recognition</i>. Elsevier, Segunda Edición, 2003.</p> <p>[7] KANTZ, H., SCHREIBER, T. <i>Nonlinear Time Series Analysis</i>, 2dn edition, Wiley, 2005.</p> <p>[8] COVER, T., THOMAS, J. <i>Elements of Information Theory</i>. Wiley, 1991</p>	

Vigencia desde:	24 de Julio 2011, revisado 4 de Enero de 2019
Elaborado por:	Pablo Estévez
Revisado por:	Área de Desarrollo Docente Martin Adams, Javier Ruiz del Solar, Doris Saéz, Claudio Pérez Comité Docencia DIE