

PROGRAMA DE CURSO

Código	Nombre			
CC6204	Deep Learning			
Nombre en Inglés				
Deep Learning				
SCT	Unidades Docentes	Horas de Cátedra	Horas Docencia Auxiliar	Horas de Trabajo Personal
6	10	3	0	7
Requisitos			Carácter del Curso	
CC3001			Electivo (pre y posgrado)	
Resultados de Aprendizaje				
<p>El objetivo del curso es introducir a los estudiantes al área de aprendizaje basado en redes neuronales profundas, comúnmente conocida como Deep Learning. Las técnicas utilizadas en esta área han resultado fundamentales en los últimos avances en Inteligencia Artificial en particular en tareas como procesamiento de lenguaje natural y visión computacional. Durante el curso el estudiante será expuesto a la teoría detrás de los modelos de Deep Learning, comprenderá su funcionamiento, los usos posibles, y será capaz de construir y entrenar modelos que permitan solucionar problemas reales.</p> <p>Desde el punto de vista práctico, el curso pretende que los estudiantes sean capaces de implementar soluciones a problemas ingenieriles complejos que requieran conceptos de Inteligencia Artificial y Deep Learning, con foco en visión computacional y procesamiento de lenguaje natural. Desde el punto de vista teórico el curso busca que los estudiantes sean capaces de leer y comprender artículos científicos recientes del área y puedan así estar al día en los últimos avances en Deep Learning. Si bien la parte teórica es sumamente importante, el curso tendrá un claro sesgo práctico y la evaluación será en su gran parte basada en tareas de programación.</p> <p>En específico, al finalizar el curso los estudiantes serán capaces de:</p> <ul style="list-style-type: none"> – Concebir la solución a un problema real que se beneficie de las técnicas de Deep Learning, entendiendo claramente los pro y contras de una posible solución basada en esta tecnología, y las necesidades de datos, software, hardware y tiempo de computación. – Diseñar una red neuronal profunda para un problema en específico considerando las distintas arquitecturas de redes posibles, las especificidades del problema en cuestión y los datos disponibles para entrenamiento. – Implementar un modelo de red neuronal profunda usando librerías de software especializadas, y compararlo a otros modelos implementados según métricas objetivas de eficiencia y efectividad. – Operar y optimizar el funcionamiento de modelos de redes neuronales profundas, entendiendo cómo modificar (hiper)parámetros, cuándo se debe re-entrenar, cuándo se debe conseguir más datos, etc. – Saber cuándo y cómo se debe iterar sobre todos los puntos anteriores para mejorar la solución a un problema en cuestión, y, de ser necesario, consultar el estado del arte del área para implementar nuevas arquitecturas u optimizaciones. 				

Metodología Docente	Evaluación General
<p>La metodología consiste en clases teóricas, revisión y presentación de artículos del estado del arte, trabajos prácticos y un proyecto. Asimismo, se considera un examen de medio semestre y uno final.</p> <p>El curso considera una parte teórica importante, pero tiene una tendencia altamente práctica a través de las tareas y proyectos.</p>	<p>Práctica</p> <ul style="list-style-type: none"> ● 2 Tareas ● 1 Proyecto (o 2 tareas adicionales) <p>Teórica</p> <ul style="list-style-type: none"> ● 1 control (mid-term) ● 1 examen <p>Participación</p> <ul style="list-style-type: none"> ● Revisión y presentación de papers. <p>Nota Final: 40% Práctica, 40% Teórica y 20% Participación.</p>

Unidades Temáticas

Número	Nombre de la Unidad	Duración en Semanas
1	Fundamentos	4
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
<p>I introducción a redes neuronales modernas. Capítulos: I.2, I.3, II.6 [1]</p> <ul style="list-style-type: none"> – Perceptrón, perceptrón multi-capas, funciones de activación, no linealidad. – Redes neuronales, cómo computan, qué no pueden computar, representación tensorial. – Álgebra tensorial y cálculo tensorial. – Funciones de error/pérdida y entrenamiento por descenso de gradiente. – Grafos de computación y el algoritmo de BackPropagation. 	<p>El alumno es capaz de:</p> <ul style="list-style-type: none"> - Entender en detalle los fundamentos de redes neuronales profundas. - Implementar redes neuronales sin la necesidad de paquetes especializados, incluyendo algoritmos de optimización avanzados. - Analizar las bondades y deficiencias de un clasificador basado en redes neuronales profundas. - Operar y optimizar el funcionamiento de un modelo basado en redes neuronales profundas. - Organizar un proyecto simple de ingeniería basado en redes neuronales profundas considerando las necesidades de datos, software, hardware y tiempo de computación. 	<p>Cap. I.2-3, II.6-8, II.11 [1]</p>

II Algoritmos de aprendizaje, regularización y optimización.

Capítulos: II.7, II.8 [1]

- Descenso de gradiente estocástico.
- Inicialización de parámetros, normalización, normalización de paquetes.
- Aprendizaje adaptativo
- Dropout
- Penalización de parámetros

III Aspectos prácticos de entrenamiento y aprendizaje.

Capítulo: II.11 [1]

- Conceptos clásicos de aprendizaje de máquina.
- Métricas de eficiencia, baselines, overfitting, underfitting.
- Búsqueda y selección de hiperparámetros.
- Técnicas de debugging
- Uso de GPUs en el entrenamiento
- Organizando una solución basada en Machine Learning y Deep Learning.

Número	Nombre de la Unidad	Duración en Semanas
2	Redes Convolucionales y Aplicaciones	5
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
<ul style="list-style-type: none"> - Redes Neuronales convolucionales - Relación de CNN con el modelo biológico - Estudio de arquitecturas CNN del estado del arte. - Caffe y TensorFlow - Visualización de modelos CNN. - Búsqueda por Similitud usando Deep Features. - Hashing de vectores de características usando Deep Learning. - <i>Class Activation Mapping</i>. - Modelos de Deep Learning para Detección de Objetos (imágenes) - Modelos de Deep Learning para Segmentación de imágenes 	<ul style="list-style-type: none"> - Entender y aplicar redes neuronales en problemas de visión por computadora. - Entender y aplicar redes convolucionales. - Entender y desarrollar aplicaciones de búsqueda por similitud basada en deep-learning. - Utilizar herramientas como Caffe y TensorFlow para desarrollar modelos de deep-learning. - Entender y desarrollar aplicaciones de detección y etiquetado de imágenes utilizando modelos convolucionales. - Entender y desarrollar aplicaciones de segmentación de imágenes utilizando modelos de deep-learning. - Entender los conceptos de convolución y deconvolución en CNN. 	<p>[1] Cap. II.9, II.12.</p> <p>[2] Cap 4, paper científicos.</p>

Número	Nombre de la Unidad	Duración en Semanas
3	Redes Recurrentes y Aplicaciones	3
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
<ul style="list-style-type: none"> - Redes recurrentes - Backpropagation en el tiempo - Redes recurrentes bidireccionales. - Dependencias temporales a largo plazo. - Modelos con memoria externa explícita. - Aplicaciones en procesamiento de lenguaje natural y otras aplicaciones basadas en secuencias. 	<p>El alumno es capaz de:</p> <ul style="list-style-type: none"> - Entender en detalle los fundamentos de redes neuronales recurrentes y sus variaciones con dependencias temporales y memoria explícita. - Implementar modelos basados en redes neuronales recurrentes usando librerías tipo TensorFlow, Theano, o Pytorch. - Analizar el funcionamiento de redes neuronales recurrentes. 	<p>Cap II.10, II.12 [1]</p>

	<ul style="list-style-type: none"> - Operar y optimizar el funcionamiento de un modelo basado en redes neuronales recurrentes. - Aplicar redes neuronales recurrentes al procesamiento de lenguaje natural y otras aplicaciones basadas en secuencias. 	
--	--	--

Número	Nombre de la Unidad	Duración en Semanas
4	Tópicos Avanzados	3
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
<p>Se elegirán algunos de los siguientes contenidos:</p> <ul style="list-style-type: none"> - Introducción a los Modelos Generativos. - Autoencoders - Autoencoder Variacionales - Generative Adversarial Networks - Neural Turing Machine (NeuralTM). - Computación Neuronal Derivable (DNC). - CapsNet. 	<p>El estudiante será capaz de enfrentarse exitosamente a literatura científica del área, y entender los conceptos del estado del arte.</p> <p>Algunos de los resultados específicos serán:</p> <ul style="list-style-type: none"> - Entender y aplicar autoencoders variacionales (VAEs). - Entender y aplicar modelos generativos usando Generative Adversarial Nets (GANs). - Identificar el contexto de aplicación de cada uno de los diferentes modelos deep analizados en el curso. - Conocer avances en el áreas (e.g. capsNet, NeuralTM, DNC). 	<p>Cap. III.14, III.20 [1] papers científicos</p>

Bibliografía

Referencias Bibliográficas:

- **[1] Deep Learning.** Ian Goodfellow , Yoshua Bengio , Aaron Courville . 2016. ISBN : 9780262035613. <http://www.deeplearningbook.org/>
- **[2] Deep Learning: A practitioner's Approach.** Josh Patterson, Adam Gibson. O'Reilly, 2017.

Papers Científicos:

- Artículos de Conferencias tipo CVPR, ICCV, ICML, ECCV, NIPS, ICLR, EMNLP.
- Artículos de Revistas tipo IJCV, PAMI.

Librerías de código:

- Caffe, <http://caffe.berkeleyvision.org/>
- TensorFlow, <https://www.tensorflow.org/>
- Pytorch, <http://pytorch.org/>
- Theano, <http://deeplearning.net/software/theano/>
- Keras, <https://keras.io/>

Vigencia desde:	Otoño 2018
Elaborado por:	Jorge Pérez y José Manuel Saavedra