



SEMINARIO DE DEEP LEARNING

FECHA: Jueves, 20 de octubre, 2022

HORA: 11:00 HORAS

LUGAR: Aula Luis Rodríguez Marín (Room 2.32 ETSI Industriales).



INTRODUCCIÓN AL DEEP LEARNING DESDE LA MATEMÁTICA APLICADA

Jorge Muñoz (Investigador Deep Learning)

Resumen.-

Se comenzará con la presentación de los conceptos básicos de las redes de neuronas, que es una neurona artificial, y el mecanismo básico que se utiliza para que estas neuronas aprendan a base de ejemplos (la diferenciación automática, la función de error y el algoritmo de descenso del gradiente). Posteriormente se comentarán los principales problemas a la hora de entrenar estos modelos (desvanecimiento del gradiente, estabilidad numérica, colapso de pesos) y las soluciones que se han planteado (regularización, función relu, conexiones residuales), también se explicará alguna alternativa al algoritmo de backpropagation usando ODEs. En la segunda parte se describirán las diferentes arquitecturas en las redes de neuronas, comenzando por las redes completamente conectadas y su relación con los espacios de Banach. A continuación se hablará de las redes convolucionales y su alternativa las capsule networks. Introduciremos los conceptos de autoencoder y embedding (y su relación con word2vec) y hablaremos de modelos basados en la teoría varacional (VAEs) y los modelos de flujo. Se describirán brevemente las redes de neuronas probabilísticas y los modelos con incertidumbre. Posteriormente se establecerá el denominado geometric deep learning basado en la teoría de grupos y sus simetrías, y se describirá la ventaja que ofrece sobre los modelos convolucionales. Se describirán los modelos recurrentes y como se originaron las neuronas recurrente con memoria (LSTM) y los modelos secuenciales que transforman secuencias de elementos en otras secuencias. Explicaremos la problemática de los modelos recurrentes por sus problemas de paralelización para ser ejecutados y como dieron lugar a los modelos de atención (Transformers) que son las bases de los modelos de procesamiento de lenguaje natural actuales (BERT, GPT-3). En la última parte se tratarán los modelos generativos, en particular de las GANs (su relación con la teoría de juegos, su problemática y alternativas como las WGAN) y de los diffusion models, explicando como es posible que las redes de neuronas sean capaces de generar imágenes tan precisas y detalladas a partir de descripciones textuales. Finalmente, se discutirá brevemente de las redes de neuronas de grafos, la relación del deep learning con el aprendizaje por refuerzo y con la computación cuántica.