

RESPONSABLE(S) DE TUTORIZACIÓN			TRABAJO FIN DE GRADO		DETALLE DEL TFG					
Número	DPTO	RESPONSABLE DE TUTORIZACIÓN	RESPONSABLE DE COTUTORIZACIÓN si procede	TIPOLOGÍA	TÍTULO	ESTUDIANTE	Descripción, resumen de contenidos y actividades a desarrollar en el ámbito de la Informática	Descripción, resumen de contenidos y actividades a desarrollar en el ámbito de las Matemáticas	Materias del Grado relacionadas	HARDWARE/SOFTWARE/BIBLIOGRAFIA
15	CCIA	Francisco Herrera Triguero		Complementario de profundización / Iniciación a la investigación	Problema del desvanecimiento del gradiente: Aplicación en Transformers	Juan Andrés Mauricio Martín	El problema del desvanecimiento del gradiente es un desafío fundamental en el entrenamiento de redes neuronales profundas, donde los gradientes de las funciones de pérdida se vuelven extremadamente pequeños, dificultando el ajuste efectivo de los pesos y, en consecuencia, el aprendizaje. Este Trabajo de Fin de Grado (TFG) se centra en el análisis matemático de este fenómeno, explorando sus causas, efectos y posibles soluciones, así como su relación con las arquitecturas basadas en Transformers. Utilizando herramientas de análisis matemático y teoría de la optimización, se investigarán las condiciones bajo las cuales ocurre el desvanecimiento del gradiente y cómo diferentes arquitecturas de redes, incluidos los Transformers, y funciones de activación pueden mitigar su impacto. A través de simulaciones y análisis teórico, se evaluará la eficacia de estas técnicas en la mejora del	Se estudiará el concepto de espacios normados y su aplicación en el análisis de la convergencia del gradiente. Además de ello, se estudiarán distintos conceptos de Álgebra Lineal. Junto a ello, se estudiarán distintos teoremas de convergencia para estudiar rigurosamente los efectos de los distintos métodos de inicialización.	Aprendizaje Automático, Visión por Computador, Inteligencia Artificial, Análisis I y II, Inferencia Estadística, K5 Geometría I y II	- Briden, J., Gurga, T., Johnson, B., Cauligi, A., & Linares, R. (2024, January). Improving computational efficiency for powered descent guidance via transformer-based tight constraint prediction. SciTech GNC Conference. DOI: 10.2514/6.2024-1760 - Arora, S., Cohen, N., Golowich, N., & Hu, W. (2019). A convergence analysis of gradient descent for deep linear neural networks. arXiv preprint arXiv:1810.02281. - von Oswald, J., Niklasson, E., Randazzo, E., Sacramento, J., Mordvintsev, A., Zhmoginov, A., & Vladymyrov, M. (2023). Transformers learn in-context by gradient