

RESPONSABLE(S) DE TUTORIZACIÓN			TRABAJO FIN DE GRADO		DETALLE DEL TFG		Materias del Grado relacionadas	HARDWARE/SOFTWARE/BIBLIOGRAFIA		
Número	DPTO	RESPONSABLE DE TUTORIZACIÓN	RESPONSABLE DE TUTORIZACIÓN si procede	TIPOLOGÍA	TÍTULO	ESTUDIANTE				
20	CCIA / AM	Francisco Javier Meri de la Maza	Pablo Mesejo Santiago	Complementario de profundización / Iniciación a la investigación	Análisis teórico y empírico del deep double descent	Juan Antonio Ruiz Arévalo	<p>Descripción, resumen de contenidos y actividades a desarrollar en el ámbito de las Matemáticas</p> <p>Las técnicas de aprendizaje profundo han mostrado una gran efectividad en numerosos problemas, desde la visión por computador al procesamiento de lenguaje natural pasando por numerosas aplicaciones de gran complejidad (como pudiera ser la predicción de la estructura de proteínas). Sin embargo, este progreso a nivel empírico o experimental, no lleva aparejado un avance a nivel teórico sobre por qué estos modelos ofrecen el rendimiento que ofrecen. Dentro de este contexto, una de las contribuciones teóricas más interesantes la ofrece la reciente hipótesis del doble descenso de gradiente (deep double descent). Se trata del fenómeno según el cual, a medida que incrementamos la capacidad de un modelo, se observa que aumenta el error de test (moviéndonos en lo que sería el régimen clásico de aprendizaje automático, con el bias-variance trade-off) para, llegado un punto, descender enormemente y ofrecer</p>	<p>Descripción, resumen de contenidos y actividades a desarrollar en el ámbito de las Matemáticas</p> <p>Como se explica en la parte informática del TFG, el uso de redes neuronales profundas está obteniendo resultados de aproximación/predicción muy buenos usando modelos sobreparametrizados. Esto parece ir en contra de la creencia clásica de que hay un punto de equilibrio entre la tendencia y la variación, que se alcanza cuando hay una relación adecuada entre el número de datos y el número de parámetros del modelo. Se propone desarrollar con detalle algunas de las explicaciones conocidas que avalan esta creencia. Además, se propone modelizar matemáticamente algunas de las arquitecturas de redes profundas más utilizadas (perceptrón multicapa y convolucional) y tratar de obtener posibles explicaciones teóricas al alto rendimiento de los sistemas sobreparametrizados. Para ello, se desarrollará la teoría de aproximación no lineal en distintos espacios.</p>	<p>Analisis Matemático I y II, Analisis Funcional</p>	<p>[M1] R. DeVore, Non-linear Approximation, Acta Numerica (1998) pp. 51-150, Cambridge University Press.</p> <p>[1] Schaeffer, R., Khona, M., Robertson, Z., Boopathy, A., Pistunova, K., Rocks, J. W., ... &amp; Koyejo, O. (2023). Double descent demystified: Identifying, interpreting &amp; ablating the sources of a deep learning puzzle. arXiv preprint arXiv:2303.14151.</p> <p>[2] Nakkiran, P., Kaplan, G., Bansal, Y., Yang, T., Barak, B., &amp; Sutskever, I. (2021). Deep double descent: Where bigger models and more data hurt. Journal of Statistical Mechanics: Theory and Experiment, 2021(12), 124003.</p> <p>[3] Heckel, R., &amp; Yamaz,</p>