

Seminari del Departament d'Estadística i Investigació Operativa

Universitat Politècnica de Catalunya

6 de novembre de 2026

Títol: Arquitectura Transformer i l'entrenament de LLMs.

Autors: Toni Lozano Bagén (UAB).

Resum: Els models de llengua grans (LLM) s'han imposat com a eina central en l'anàlisi i generació automàtica de text, i darrere de pràcticament tots hi ha la mateixa arquitectura: el transformer. En aquesta xerrada farem una passejada descriptiva per les seves peces fonamentals (el mecanisme d'atenció, les connexions residuals, la normalització) amb l'objectiu de fer comprensible el seu funcionament intern. A continuació, repassarem les tres fases principals que defineixen l'entrenament d'un LLM: el preentrenament no supervisat sobre grans corpus, el fine-tuning supervisat i l'aprenentatge per reforç a partir de la retroalimentació humana (RLHF).

Sobre l'autor: Toni Lozano és professor lector al Departament de Ciències de la Computació de la Universitat Autònoma de Barcelona (UAB). Es va doctorar en Matemàtiques a la UAB (2016), després d'obtenir, el 2011, la llicenciatura en Matemàtiques i l'enginyeria Informàtica. Des del 2016, treballa activament en l'àmbit de la ciència de dades, amb especialització en mètodes d'aprenentatge profund (Deep Learning). Al llarg dels darrers anys ha contribuït al desenvolupament de diversos projectes per a empreses i institucions, aplicant xarxes neuronals a l'anàlisi de grans conjunts de dades no estructurades.