Universidad de Navarra
Facultad de Ciencias

Propuesta de Trabajo Fin de Máster
Año académico 2026-2027
MÁSTER EN CIENCIA DE DATOS PARA CIENCIAS EXPERIMENTALES

Proyecto Nº 12

Título: Unraveling the genetic makeup of rare diseases through AI and advance sequencing technologies

Departamento/ Laboratorio: Biologia Computacional (CIMA)

Director: Mikel Hernaez
Correo electrónico: mhernaez@unav.es
Codirector: Ana Patiño
Correo electrónico: apatigar@unav.es

Resumen:

This Master's thesis will build an AI-assisted pipeline (based on agentic models if posible) that integrates advanced sequencing (short-read and long-read whole-genome sequencing plus RNA-seq when available) to identify and prioritize causal variants in rare diseases. The central idea is to combine state-of-the-art variant discovery with a multi-modal learning model (such as LLMs) that scores candidate variants by jointly considering sequence context, regulatory impact, and patient phenotype.

Practically, the work targets three gaps that limit diagnostic yield: (1) structural and repeat variants missed by short reads, (2) splicing and expression dysregulation not captured by DNA alone, and (3) non-coding variant interpretation. The proposed pipeline will surface high-confidence, mechanistically plausible candidates and generate human-readable reports for clinicians.

OPTATIVAS RECOMENDADAS

1. Advance topics in Machine Learning
2. Deep Learning
3. Análisis de datos de alto rendimiento
4.