

Abordando el desbalance de datos y la equidad en modelos de clasificación de Machine Learning para COVID-19

Andreea Madalina Oprescu Popescu

Departamento de Tecnología Electrónica, Universidad de Sevilla. Instituto de Ingeniería Informática de la Universidad de Sevilla.

Facultad de Informática

Sala de Grados - miércoles 11 de junio de 2025 - 15:00

Entrada libre hasta completar el aforo

Resumen:

En el contexto del avance de la Inteligencia Artificial, ¿cómo podemos aplicar de manera efectiva el poder predictivo del Machine Learning a problemas sanitarios complejos y de gran relevancia, como la clasificación de la mortalidad en pacientes con COVID-19? En esta charla, exploraremos no solo una aplicación concreta del aprendizaje automático en el ámbito clínico, sino que también abordaremos uno de sus mayores desafíos: el desbalance de datos. Analizaremos diversas estrategias para mitigar este problema, entendiendo que corregir la asimetría entre clases, aunque necesario, no garantiza por sí solo un comportamiento justo del modelo. Exploraremos también cómo los modelos pueden aprender y perpetuar sesgos intrínsecos en los datos, llevando a predicciones injustas o menos precisas para ciertos grupos de pacientes. Utilizando herramientas como Aequitas, auditaremos los modelos para desvelar y cuantificar disparidades de su rendimiento entre diferentes subgrupos.

Sobre Andreea Madalina Oprescu Popescu:

Profesora en la U. Sevilla. Doctora en Ingeniería Informática por la Universidad de Sevilla (2023) y miembro del grupo Tecnología Electrónica e Informática Industrial. Su actividad investigadora se centra en la inteligencia artificial responsable y confiable aplicada a la salud, con un interés específico en cómo poder mejorar la vida de las personas y prestar apoyo a equipos clínicos. Colabora en investigación con instituciones como el Hospital Universitario Virgen del Rocío y la Fundación Progreso y Salud. Su investigación ha dado como resultado 8 publicaciones en revistas de alto impacto, además de varias ponencias en congresos nacionales e internacionales. Específicamente, ha investigado acerca del uso de métodos IA y aprendizaje automático para la detección temprana y la predicción de complicaciones obstétricas. Asimismo, ha propuesto una metodología responsable de recolección de datos para la IA en salud, subrayando aspectos éticos, de calidad de datos y la relevancia de estudios cualitativos y prospectivos.