

DataReady4ML. Herramientas Deep Learning para la gestión del dato para la salud personalizada

Consortio: Ibermatica; i3B

Tecnología: Administración Digital & Slud; Inteligencia Artificial

Descripción general:

Algoritmia de estructuración de textos, de anonimización y generación de datos sintéticos y aseguramiento de la calidad del dato y ausencia de sesgo, que palie los problemas de explotación mediante Machine Learning de los datos almacenados en la HCE. Para ello propone. la investigación en la inteligencia artificial y de manera especial en las técnicas de aprendizaje profundo (Deep Learning), como la base tecnológica para lograrlo.

1. Trabajo en la estructuración automática del lenguaje natural en entornos clínicos.
 - Detectar y clasificar mediante DeepLearning los conceptos médicos en los informes médicos escritos en lenguaje natural y mapearlos con los códigos de las nomenclaturas de SNOMED y/o CIE-10.
 - Detectar relaciones entre sí en los informes médicos no estructurados, permitiendo obtener así información más rica y precisa que los conceptos por separado
 - Clasificación automática de la información
2. Técnicas para garantizar la anonimización del conjunto de datos extraídos y evaluación de riesgo y generación de datos sintéticos.
 - Investigar y desarrollar Técnicas y metodologías orientadas a la anonimización de información para su uso en Machine Learning.
 - Explorar e investigar métodos para la generación de datos sintéticos y sus potenciales usos en el ámbito de servicios de salud
3. Tratamiento en los datos para garantizar su calidad y equidad

Programa: Hazitek Competitivo (ZL-2022/00747; ZL-2023/00869)

Duración: 21 meses (2022 – 2023)

Presupuesto global proyecto: 185.399,22 €

Presupuesto Grupo Ayesa: 185.399,22 €

ACTUACIÓN COFINANCIADA POR EL DEPARTAMENTO DE DESARROLLO ECONÓMICO, SOSTENIBILIDAD Y MEDIO AMBIENTE DE GOBIERNO VASCO Y DE LA UNION EUROPEA A TRAVÉS DEL FONDO EUROPEO DE DESARROLLO REGIONAL (FEDER)



DataReady4ML. Herramientas Deep Learning para la gestión del dato para la salud personalizada

Consortio: Ibermática; i3B

Tecnología: Administración Digital & Slud; Inteligencia Artificial

Rol de Ayesa:

Ibermática ha definido un marco metodológico basado en modelos generativos interpretables, orientado a capturar patrones no aleatorios presentes en los datos, considerando además efectos asociados a datos ausentes, relaciones no lineales e incertidumbres. Este enfoque incluye técnicas que permiten ajustar los datos sintéticos generados a distintos niveles dentro del espectro privacidad–utilidad, garantizando un equilibrio adecuado entre protección del dato y valor analítico.

A nivel global, Ibermática lidera la gestión del proyecto, encargándose de la coordinación entre socios, el aseguramiento de la calidad de los resultados y la supervisión de todas las actividades técnicas y administrativas.

Participa en la definición de requisitos y en el diseño de la arquitectura de la solución, alineando el desarrollo tecnológico con las necesidades funcionales y operativas identificadas por el consorcio.

Desde el punto de vista técnico, desarrolla modelos avanzados de inteligencia artificial aplicados al procesamiento de texto clínico, incluyendo tareas de codificación médica, clasificación de documentos y extracción de entidades y relaciones mediante técnicas de deep learning.

De forma complementaria, aborda el uso seguro del dato, desarrollando técnicas de anonimización, generación de datos sintéticos y herramientas para el análisis de la privacidad. Asimismo, trabaja en la mejora de la calidad de los datos y de los modelos, tratando aspectos como la mitigación de sesgos, el ajuste automático de modelos, el tratamiento de grandes volúmenes de datos y la detección de cambios o derivas en los mismos.

Finalmente, Ibermática se encarga de la validación de las soluciones desarrolladas, tanto a nivel individual como integrado en escenarios realistas, y define la estrategia de explotación y puesta en mercado, liderando también las actividades de difusión de resultados.

