



Becas colaboración curso 2022/2023

Fecha: 01 Junio 2022

Vicerrectorado de Investigación

Subcomisión de I+D+i

Propuesta del departamento *SISTEMAS INFORMÁTICOS Y COMPUTACIÓN*

Núm Proyecto: 2022/32/00014

Responsable

Hurtado Oliver, Lluís Felip

E-mail

lhurtado@dsic.upv.es

Ext.

83500

Título proyecto

Aplicación de Técnicas de Procesamiento del Lenguaje Natural para la extracción de información en informes médicos

Valoración proyecto

4

Descripción proyecto

El grupo de investigación ELiRF tiene una dilatada experiencia en distintas aplicaciones del Procesamiento de Lenguaje Natural. Algunas de estas aplicaciones son la monitorización de redes sociales, el análisis de artículos de prensa y programas de debate televisivos, y el resumen automático. Actualmente ha empezado a trabajar en un campo de gran interés para la aplicación de técnicas de Procesamiento de Lenguaje Natural: la extracción de conocimiento a partir de informes médicos.

La mayor parte de la información clínica de los pacientes se encuentra en forma de texto libre en las historias clínicas electrónicas. Los avances en Inteligencia Artificial en el campo del Procesamiento del Lenguaje Natural habilitan la posibilidad de extraer conocimiento de forma automática mediante el procesamiento de dichas historias.

El análisis automático del texto libre disponible en los informes clínicos permite la extracción y etiquetado de múltiples elementos como el tipo, las características y la región anatómica de los hallazgos clínicos o las enfermedades diagnosticadas. Además, resultan de especial interés la estandarización de los términos utilizando vocabularios controlados del ámbito biomédico como Snomed-CT, UMLS o CIE.

El grupo de investigación dispone de gran cantidad de recursos hardware, software y corpus que darán soporte durante la realización del trabajo asociado a la beca.

Actividades a realizar por el alumno

El candidato o candidata se integrará en el grupo de investigación, que forma parte del instituto de investigación VRAIN, asistiendo a reuniones y colaborando con sus miembros en los distintos proyectos activos.

En particular trabajará en la tarea de extracción de información en informes médicos. Para ello se realizará un estudio de los términos médicos más frecuentes en los informes disponibles y se aprenderán representaciones densas de términos generales y específicos.

Con todo ello, el becario o becaria aprenderá los fundamentos teóricos (modelos de lenguaje, métricas de evaluación, etc.) y tecnológicos (word embeddings, deep learning, etc.) que serán necesarios para abordar las tareas encomendadas.

Localización de la actividad (Campus)

Campus de Vera

Horario

3 horas diarias, a acordar con el o la estudiante