



## Becas colaboración curso 2021/2022

Fecha: 28 Mayo 2021

### Vicerrectorado de Investigación, Innovación y Transferencia

Subcomisión de I+D+i

Propuesta del departamento *MATEMÁTICA APLICADA*

**Núm Proyecto: 2021/26/00006**

#### Responsable

Conejero Casares, José Alberto

#### E-mail

aconejero@upv.es

#### Ext.

19664

#### Título proyecto

Análisis de calidad de datos en el ámbito biomédico

#### Valoración proyecto

3

#### Descripción proyecto

La creciente disponibilidad de bases de datos biomédicas masivas permite aumentar la capacidad de desarrollar modelos predictivos en cuanto a la evolución de ciertas enfermedades y de los pacientes. A la hora de considerar estos problemas es conveniente hacer un análisis comparativo de las distintas fuentes de procedencia de la información con el fin de determinar si estas son comparables o no, y en su caso si se pueden agregar con fines investigadores. Estas medidas de calidad están basadas en la Teoría de Información y son un paso preliminar antes de establecer los modelos mencionados anteriormente.

#### Actividades a realizar por el alumno

El alumno aplicará la metodología descrita en [1] y aplicada recientemente en [2] a varios conjuntos de datos con el fin de poder realizar un estudio de carácter investigador. La temática será a convenir con el alumno. Para ello se cuenta con la colaboración del Prof. Juan Miguel García Gómez y del personal técnico del Biomedical Data Science Lab (ITACA-UPV).  
Más información en:

[1] C. Sáez, et al.. EHRtemporalVariability: delineating temporal data-set shifts in electronic health records. *GigaScience*, 9(8), giaa079 (2020).

[2] C. Sáez, N. Romero, J.A. Conejero, and J.M. García-Gómez. Potential limitations in COVID-19 machine learning due to data source variability: a case study in the nCov2019 dataset. 15:28, 360-364 (2021).  
doi:10.1093/jamia/ocaa258

#### Horario

A convenir