



## Becas colaboración curso 2019/2020

Fecha: 07 Junio 2019

### Vicerrectorado de Investigación, Innovación y Transferencia

Subcomisión de I+D+i

Propuesta del departamento *MATEMATICA APLICADA*

**Núm Proyecto: 2019/26/00005**

#### Responsable

Conejero Casares, José Alberto

#### E-mail

aconejero@upv.es

#### Ext.

79664

#### Título proyecto

ANÁLISIS DE CALIDAD DE DATOS Y MÉTODOS DE DEEP LEARNING EN EL ÁMBITO BIOMÉDICO

#### Valoración proyecto

3,5

#### 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. A partir de los análisis de calidad se está en condiciones de implementar de manera óptima método de Deep Learning de cara a establecer métodos de clasificación y diagnóstico de enfermedades.

#### 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 iniciar 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. Saez et al. Applying probabilistic temporal and multisite data quality control methods to a public health mortality registry in Spain: a systematic approach to quality control of repositories. J Am Med Inform Assoc (2016). doi:10.1093/jamia/ocw010

[2] F.J. Pérez-Benito et al. Temporal variability analysis reveals biases in electronic health records due to hospital process reengineering interventions over six years. Preprint, 2018.

#### Horario

A convenir