UNIVERSIDAD DE GRANADA

biología Granada

Propuesta TFGB. Curso: 22/23

DEPARTAMENTO: CCIA

CÓDIGO DEL TFG: CCIA-05

1. DATOS DEL TFG OFERTADO:

Título: **Machine Learning para conjuntos de datos biológicos**

Resumen (Introducción, Objetivos y Plan de trabajo; máx. 2.460 caracteres con espacios y fuente Arial 9):

Introducción
La aplicación de técnicas de Machine Learning (aprendizaje automático) para el análisis de datos permite la obtención de modelos (basados en reglas, en árboles, en redes neuronales, etc.) que ayudan a la predicción (aprendizaje predictivo mediante clasificación o regresión) o que permiten entender la naturaleza de los datos en forma de información interpretable y útil (aprendizaje descriptivo mediante reglas de asociación o clustering). En ocasiones ambos escenarios se pueden dar a la vez, como en el caso de los modelos basados en reglas, en los que el modelo puede ser interpretado desde el punto de vista del usuario final, a la vez que permiten realizar una predicción acertada para un nuevo caso/objeto (normalmente desconocida y/o difícil de medir para ciertos problemas). El manejo y comprensión de éstas herramientas de la IA es de especial interés en el área de la biología, la química, la medicina, etc.…

Objetivos y plan de trabajo
El objetivo del presente TFG es familiarizar al alumno/a con dicho tipo de técnicas mediante su aplicación para resolver algunos problemas *relacionados con la biología,* ya sean de clasificación, regresión, asociación y/o clustering. Para ello se propone el siguiente plan de trabajo:

- El alumno/a debe estudiar y comprender el funcionamiento de los algoritmos básicos más conocidos que sean de aplicación a los datos biológicos con los que se trabaje.
- El alumno/a debe familiarizarse con herramientas como KEEL, Weka, Knime, librerías de R y/o Python, etc., para aplicar los algoritmos existentes *en los conjuntos de datos biológicos* anteriormente mencionados: Preparación de los datos (excel, etc.), carga de datos, ejecución, posibles comparativas entre algoritmos, etc.
- El alumno/a debe aplicar los algoritmos más prometedores para resolver uno o varios problemas del mundo real.
- El alumno/a debe analizar los modelos obtenidos para ver si se puede extraer conocimiento útil y/o tener cierta validación o justificación biológica.

Tabla de actividades y dedicación estimada:

Planteamiento y desarrollo del trabajo	200 horas
Elaboración de la memoria	75 horas
Preparación y ejecución de la exposición	25 horas
TOTAL (12 ECTS)	**300 horas**

2. MODALIDAD (*): Trabajo Experimental / de Investigación

(*) En el caso de trabajos experimentales, el tutor considera conveniente que el estudiante realice el taller "Prevención de riesgos y eliminación de residuos en el laboratorio" ☐

3. DATOS DEL TUTOR/A UGR (**):

Apellidos: Alcalá Fernández **Nombre**: Rafael

Teléfono: 958 240427 ***e-mail:*** alcala@decsai.ugr.es

(**) En el caso de trabajos desarrollados en Empresas u otras Instituciones ajenas a la Universidad de Granada, por favor completar la siguiente información:

TUTOR/A DE LA EMPRESA O INSTITUCIÓN:

Apellidos: **Nombre:**

Empresa / Institución:

Teléfono: ***e-mail:***

4. DATOS DEL ESTUDIANTE (***):

(***) Si ha sido acordado por el estudiante y profesor/a, por favor completar la siguiente información sobre el estudiante:

Apellidos: Baños Sánchez **Nombre:** Rosa

e-mail institucional: banossanchez@correo.ugr.es