UNIVERSIDAD DE GRANADA

Propuesta TFGB. Curso: 2020-2021

DEPARTAMENTO: CCIA

CÓDIGO DEL TFG: CCIA-09

1. DATOS DEL TFG OFERTADO:

Título: Modelos computacionales de Ciencia de Datos aplicados a problemas biológicos

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

La extracción de conocimiento sobre diferente problemas en el área de biología según modelos de Machine Learning y Data Mining. Actualmente la dificultad a la hora de modelar nuevos escenarios de trabajo está creciendo. Por un lado, se encuentra el gran volumen de información que se es capaz de extraer de los distintos dominios de aplicación, lo que se conoce como Big Data. Por otro, la representación de esta información necesita ser preprocesada con metodologías robustas para mejorar la calidad de los datos, hasta alcanzar el Smart Data. Por ello se habla de la Ciencia de Datos como el proceso completo necesario para la extracción, interpretación, modificación, generación de modelos, y visualización de resultados El manejo y comprensión de las herramientas de IA disponibles para realizar esta metodología es de especial interés en el área de la biología, la química, o la medicina.

Se pretende familiarizar al alumno/a con el proceso completo de Ciencia de Datos, mediante el uso e implementación de técnicas a partir de su programación en lenguaje Python, R o similar, y su uso posterior en algunos problemas relacionados con la biología. Para ello se propone el siguiente plan de trabajo:
- Estudiar, comprender y resumir el funcionamiento de los algoritmos básicos más conocidos para preprocesamiento de información: selección de características y filtrado de ruido.
- Seleccionar métodos robustos de generación de modelos para clasificación a partir de bibliotecas estándar.
- Familiarizarse con los problemas relativos a su campo de trabajo, en particular basados en genómica. Ha de utilizar y buscar en los repositorios públicos disponibles conjuntos de datos relacionados con la biología.
- Realizar un análisis comparativo de los distintos algoritmos para determinar cuales son los más interesantes.
- Implementar en Python o lenguaje de programación equivalente los algoritmos que considere más interesantes.
- Intentar mejorar los resultados obtenidos por los modelos de clasificación estándar.
- Realizar un informe de visualización para representar los resultados obtenidos.

Tabla de actividades y dedicación estimada:

Planteamiento y desarrollo del trabajo	200
Elaboración de la memoria	80
Preparación y ejecución de la exposición	20
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: Fernández Hilario **Nombre**: Alberto

Teléfono: 958240079 ***e-mail:*** alberto@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: **Nombre:**

e-mail institucional: