UNIVERSIDAD DE GRANADA

Propuesta TFGB. Curso: 2021-2022

DEPARTAMENTO: DECSAI

CÓDIGO DEL TFG: CCIA-04

1. DATOS DEL TFG OFERTADO:

Título: Desarrollo de un flujo de trabajo en bioinformática para problemas biomédicos

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

El campo de la Biología es cada días más cuantitativo. Por ello, el análisis de datos y la extracción de conocimiento es una tarea fundamental en este área de trabajo. En este contexto, se define 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. Es importante destacar asimismo que el valor de la información se encuentra en los datos en crudo, y éstos necesitan de un proceso de captura, transformación y preparación para estar listos de cara a realizar el aprendizaje de los modelos. Esta fase, que se ha relacionado tradicionalmente con el trabajo de la bioinformática, es de vital conocimiento para un egresado en Biología.

Este TFG pretende familiarizar al alumno/a con el proceso completo de Bioinformática y 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:
- Conocer y utilizar el proceso de captura de datos a partir de ficheros en crudo tipo "fastq"
- 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:**