



## 1. DATOS BÁSICOS DEL TFG:

**Título:** Inteligencia Artificial Confiable en el Campo de la Bioquímica.

**Descripción general** (resumen y metodología):

Los sistemas de Inteligencia Artificial (IA) y Machine Learning son cada día más utilizados en numerosas áreas de la Ciencia, y la Biología no es una excepción. Sin embargo, dado el carácter sensible de los casos de uso donde se aplican, resulta imprescindible dotar a este tipo de soluciones de una transparencia tal que su uso sea confiable para los actores humanos. Estos desarrollos suelen contemplarse como "sistemas de ayuda a la decisión" y en ese sentido deben incorporar características que permitan realizar una correcta auditoría de estos. Sin embargo, resulta complejo encontrar cuáles son las soluciones actuales para dotar de interpretabilidad a las técnicas de IA debido a la explosión de opciones que existen. El manejo y comprensión de estas herramientas es de especial interés en el área de la bioquímica, la medicina, etc..... Para comprender las relaciones identificadas por los modelos obtenidos.

**Tipología:** Trabajo de investigación o desarrollo bioinformático

**Objetivos planteados:**

El objetivo del presente TFG es familiarizar al estudiante con dicho tipo de técnicas mediante su análisis en algunos problemas relacionados con la bioquímica. Para ello se propone el siguiente plan de trabajo:

- El estudiante debe acceder a la bibliografía principal usando los repositorios de Scopus y/o Pubmed, así como los estudios más relevantes de Arxiv y Bioarxiv.
- El estudiante debe estudiar, comprender y resumir el funcionamiento de los algoritmos básicos más conocidos para clasificación.
- El estudiante realizará algunas pruebas experimentales en Python o/y R para determinar el funcionamiento de las soluciones disponibles para dotar de explicabilidad a la IA en casos de uso de las biociencias.
- El estudiante debe interpretar desde el punto de vista bioquímico, las explicaciones generadas por las técnicas analizadas sobre alguna de las bases de datos estudiadas.

Finalmente, redactará una memoria que condense la información indicada en los objetivos del presente trabajo.

**Bibliografía básica:**

- Alejandro Barredo Arrieta, Natalia Díaz-Rodríguez, Javier Del Ser, Adrien Bennetot, Siham Tabik, Alberto Barbado, Salvador Garcia, Sergio Gil-Lopez, Daniel Molina, Richard Benjamins, Raja Chatila, Francisco Herrera, Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI, Information Fusion, Volume 58, 2020, Pages 82-115.
- Guidotti, R. et al. A survey of methods for explaining black box models. ACM Comput. Surv. 51, 93:1-93:42 (2018).
- S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in Advances in Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc., 2017, vol. 30, p. 4768-4777.

- Marco Ribeiro, Sameer Singh, and Carlos Guestrin. 2016. "Why Should I Trust You?": Explaining the Predictions of Any Classifier. In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations, pages 97-101, San Diego, California

**Recomendaciones y orientaciones para el estudiante:**

**Plazas:** 1

**2. DATOS DEL TUTOR/A:**

**Nombre y apellidos:** JESÚS ALCALÁ FERNÁNDEZ

**Ámbito de conocimiento/Departamento:** CIENCIAS DE LA COMPUTACIÓN E INTELIGENCIA ARTIFICIAL

**Correo electrónico:** jalcala@ugr.es

**3. COTUTOR/A DE LA UGR (en su caso):**

**Nombre y apellidos:**

**Ámbito de conocimiento/Departamento:**

**Correo electrónico:**

**4. COTUTOR/A EXTERNO/A (en su caso):**

**Nombre y apellidos:**

**Correo electrónico:**

**Nombre de la empresa o institución:**

**Dirección postal:**

**Puesto del tutor en la empresa o institución:**

**5. DATOS DEL ESTUDIANTE:**

**Nombre y apellidos:**

**Correo electrónico:**