**Título de las prácticas (Title of the internship):**

Utilizing Deep Learning Models for Protein Folding Prediction: An Exploration of Explainability, Biases, Ethics, and Trustworthiness

**Descripción de las funciones del alumno (Description of the student´s tasks)**

Protein folding is a critical aspect of understanding molecular biology, with applications regarding from drug discovery to disease mechanisms. Recent advances in deep learning, exemplified by AlphaFold, have shown remarkable promise in predicting protein structures. This research project aims to explore the utilization of deep learning models, like AlphaFold, for protein folding prediction, while addressing key challenges related to explainability, biases, ethics, and trustworthiness. The research will investigate the explainability of deep learning predictions in the context of protein folding to ensure transparency and comprehensibility. It will also undertake a critical examination of potential biases in the training data and model outputs, considering their implications for protein structure predictions. Ethical considerations regarding the use of AI models for protein folding will be address. The study will further explore and propose strategies to determine the trustworthiness of deep learning models used in protein folding prediction, ensuring their responsible and reliable application in the field.

**Requisitos (Prerequisites):** *(indicar titulación y curso) (give Grade and academic year); otros requisitos adicionales (idiomas, informática, otros conocimientos, etc) (other aditional prerequitistes (languages, informatics, other knowledge, etc)*

- Candidates should hold a bachelor's degree in biotechnology, biology, biochemistry or related field.
- It is recommended to complete the subjects of machine learning, and knowledge representation and acquisition.
- Proficiency in English is a requirement, as the analysis of scientific literature and communication predominantly occurs in English.

**Proyecto formativo (Training Project)**

Module EXTERNAL PRACTICES. The fundamental goal of the external practices is to guide the student in applying his previously acquired knowledge to real tasks in a group work environment the realistically represents the work conditions the students will encounter in their future roles. In this way, the student will be able to get familiar with a working environment (work schedule, responsibility, attitude, organization, etc…), and with the adequate working methodology in professional reality, contrasting and applying the acquired academic knowledge.

This training initiative provides students with an exceptional chance to integrate their knowledge in molecular biology, proteomics, and artificial intelligence. It empowers them with the ability to critically assess and make meaningful contributions as researchers in contemporary scientific domains.

-Supervision and Guidance: The mentor or advisor will provide ongoing guidance and support to the student throughout the project, offering consistent feedback, addressing inquiries, and cultivating a conducive learning environment.

-Project Duration and Scheduling: This program includes specific start and end dates and implements a flexible timetable that adapts to the student's availability.

-Learning objectives: The study is significant as it bridges the gap between advanced AI techniques and the complex field of protein folding. The results will not only advance our understanding of molecular biology but also contribute to responsible AI adoption in life sciences.

## Actividades a desarrollar en la práctica académica (Activities that will be performed in the academic internship):

1. Literature Review: Review existing literature on deep learning in protein folding, explainability in AI, biases in data, ethics in AI, and trustworthiness of AI models.

2. Data collection and Preprocessing: collect pertinent protein structure data and datasets, and prepare them through preprocessing for subsequent model training and evaluation.

3. Model selection: Select and utilize deep learning models for protein folding prediction, including models similar to AlphaFold.

4. Explainability Analysis: Employ interpretability techniques to understand and elucidate model decisions in protein folding predictions.

5. Bias Assessment: Examine potential biases in the training data and model predictions and propose strategies to mitigate them.

6. Ethical Considerations: Investigate the ethical implications of using deep learning models in molecular biology, including issues related to privacy, consent, and social impact.

7. Trustworthiness Enhancement: Develop measures to enhance the trustworthiness of the models, such as robustness, fairness, and accountability.

8. Evaluation and Validation: Assess the performance of the models, their explainability, and ethical implications through rigorous validation and testing.

9. Documentation and Reporting: Record the methodologies employed, document the findings and insights derived from the analysis, and compile a comprehensive report summarizing the project's results and contributions.

| | |
|---|---|
| **Nº de plazas**:<br><br>(Nr. of places) | **1** |
| **¿El alumno tendrá trato habitual con menores?**<br><br>**(Has the student dealings with underage persons?)** | **No** |
| **Fecha de inicio:**<br><br>**(Starting date)** | **08 / 01 / 23** |
| **Fecha de fin:**<br><br>**(End date)** | **31/05/2023** |
| **Horas semanales**:<br><br>**(Weekly hours)** | **20** |
| **Horario jornada laboral**:<br><br>**(Working hours)** | **Flexible** |
| **Importe Ayuda/Bolsa de estudio**:<br><br>**(Amount of fellowship / remuneration)** | **0 €/mes** |
| **Tutor académico**:<br>(Academic tutor (UPM))<br>Email: | **Emilio Serrano Fernández**<br>emilio.serrano@upm.es |

| | |
|---|---|
| **Departamento tutor académico**:<br><br>**(Dept. of academic tutor)** | **Inteligencia Artificial** |
| **Tutor empresa**:<br><br>**(External tutor)** | Paloma Tejera Nevado |
| **Email tutor empresa**:<br><br>**(Email external tutor)** | paloma.tejera@upm.es |
| **Departamento tutor empresa**:<br><br>**(Dept. of external tutor)** | Medical Data Analytics Laboratory (MEDAL) |
| **Ubicación de la estancia de las practicas**<br><br>**(Location of the internship)** | Parque Científico y Tecnológico de la UPM, Crta. M40, Km. 38,<br><br>28223 Pozuelo de Alarcón, Madrid |
| **ENTIDAD COLABORADORA**:<br><br>**(Collaborating Entity)** | CTB (CTB-UPM) Center for Biomedical Technology |
| *A cumplimentar por Oficina Prácticas ETSIAAB:*<br>**Créditos a reconocer (Nº ECTS):** | |

**Enviar por email a: OFICINA DE PRÁCTICAS ACADEMICAS EXTERNAS – ETSIAAB**
secretaria.pei.etsiaab@upm.es – Secretarias: Visitación Pérez / Susana Pardo - Tfno: 913363686)