Biomedical Explainable AI

Contact
Prof. Jaume Bacardit, jaume.bacardit@newcastle.ac.uk

Research project
Biomedical data is growing at an unprecedented scale. Recent decades have seen the development of a very broad range of bio/digital-technologies able to generate data of human activity at all scales from the cellular to the whole-body level. Such data has little value unless appropriate analysis techniques are used. Artificial Intelligence (AI) methods have shown their value in analysing biomedical data of all kinds. However, a question often overlooked is understanding how AI models take decisions. This prevents the extraction of valuable novel insight from the data and a rigorous assessment of the reliability and fairness of AI model. Such assessment is crucial before their deployment to aid decision making in precision medicine initiatives [1-2].

Explainable AI [3-4] is a subfield of AI that focuses on how to extract knowledge and explain decisions of AI models. It has rapidly grown in recent years, with a broad range of strategies for understanding better how AI models operate and maximising the value we can obtain for the costly biomedical data. However, most Explainable AI methods are data-agnostic and do not exploit any of the wealth of domain knowledge that exists in biomedical data, still requiring substantial human effort in assessing the reliability of AI models from a biological and clinical perspective.

Building upon Bacardit's expertise in this area [5-10], this project will develop a new class of Explainable AI methods designed specifically to tackle biological and biomedical data, by incorporating the biological knowledge directly in the process of training and refining AI models. In this way models will be able to directly capture biomedically-relevant patterns in data, and it will accelerate the process of analysing biomedical data and enable their deployment in clinical settings. We have access to a broad range of biomedical data coming from clinical studies that will be used to develop and evaluate such technology, and a network of clinical experts that we can use to ensure that the developed technology is clinically meaningful.

Applicant skills/background
This project requires a student with strong computer science foundations. Existing experience in AI/machine learning will be highly valued. Biological knowledge will be a plus but is not crucial, as this will be provided through the project.

References
[1] Topol. *Nature medicine 25.1 (2019): 44-56.* **[2]** Goecks et al. *Cell 181.1 (2020): 92-101.* ***[3]*** *Arrieta et al. Information fusion 58 (2020): 82-115.* ***[4]*** *Tjoa & Guan. IEEE transactions on neural networks and learning systems 32.11 (2020): 4793-4813.* ***[5]*** *Widera et al. Osteoarthritis and Cartilage Open (2023): 100406.* ***[6]*** *van Helvoort et al. BMJ open 10.7 (2020): e035101.* ***[7]*** *Little et al. Psychological medicine 51.9 (2021): 1441-1450.* ***[8]*** *Danesh et al. Translational Vision Science & Technology 11.10 (2022): 10-10.* ***[9]*** *Swan et al. BMC genomics 16 (2015): 1-12.* ***[10]*** *Lazzarini et al. BioData mining 9.1 (2016): 1-23*