UNO IS&T Research Seminar Series - **Towards Explainable Machine Learning: A Case Study for Liver Transplantation**
Dr. Pedro Cabalar
https://www.dc.fi.udc.es/~cabalar/

Dr. Pedro Cabalar is a full professor at the Computer Science and Information Technology Department, University of A Coruña, Spain. His research interests focus on Artificial Intelligence and Knowledge Representation and Reasoning with special emphasis on temporal, causal, epistemic, and qualitative spatial reasoning, mostly using non-classical logics and logic programming. Many of his contributions are related to logical foundations and extensions of Answer Set Programming. He is also interested in logic-based explanations of machine learning classifiers. Pedro Cabalar has collaborated with more than twenty coauthors from universities of nine different countries, publishing in relevant journals like AIJ, TPLP, AMAI, AI Communications, the Journal of Applied Non-Classical Logics or the Logic Journal of the IGPL, and reviewing for other journals like JAIR, J. of Logic and Computation, J. of Applied Logic or ACM TOCL. He is an editor of the Artificial Intelligence Journal. He has regularly published in the main relevant international conferences like IJCAI, ECAI, KR, AAAI, JELIA, ICLP and LPNMR, being a frequent PC member for these events and acting as senior/area PC for IJCAI, ECAI and KR. He co-chairs the 40th edition of ICLP in 2024 in Dallas, TX, USA and has also co-chaired and organized the 12th edition of LPNMR in 2013 in A Coruña. Pedro Cabalar is the general coordinator of the Master in Artificial Intelligence offered by the three public Galician Universities. Pedro Cabalar has been the principal investigator of five national research projects related to Knowledge Representation and Automated Reasoning with applications to the medical domain. He has acted as workgroup leader and Management Committee member of the European COST Action DigForASP, on Automated Reasoning applied to Digital Forensics.

Artificial Intelligence (AI), and more particularly Machine Learning (ML) techniques, are experiencing an unstoppable and widespread transfer to most aspects of society. The eventual (positive or negative) consequences of many AI applications are yet to be determined, and their potential risks constitute a serious concern. One relevant feature to make AI systems trustworthy is the capability of generating explanations that may not only describe the system's technical behavior but can also justify its decisions or results in a comprehensible and useful way for a human. In this talk, we analyze several challenges for explanation generation in the context of ML for the sensitive area of health systems. In particular, we take the case study of survival prediction for donor-patient matching in liver transplantation. We mostly consider symbolic ML models such as Decision Trees (DT), since their results are easier to present and justify in human terms. We show how, even for a DT, the extracted explanations can be sometimes counterintuitive due to a lack of relevance in the conditions, or to inappropriate reference to causation versus correlation. We also discuss several proposed solutions in terms of logical reasoning and logic programming.

Wednesday, October 10th, 4:00pm - 4:40pm
Room PKI 158

Contact: @Jorge Fandinno Garcia