SCIENCES & BIOENGINEERING SCIENCES

The Research Group

Artificial Intelligence Lab

has the honor to invite you to the public defense of the PhD thesis of

Alexandre Renaux

to obtain the degree of Doctor of Sciences

Joint PhD with Université libre de Bruxelles

Title of the PhD thesis:

Bridging the Knowledge Gap in Multigenic Pathogenicity Predictions Leveraging integrated networks and machine learning to explain predictions in oligogenic diseases

Promotors:
Prof. dr. Ann Nowé (VUB)
Prof. dr. Tom Lenaerts (ULB)

The defense will take place on

Tuesday, October 10, 2023 at 16h30 in the ULB Solvay Room, NO Building - Floor 5

Members of the jury

Prof. dr. Gianluca Bontempi (ULB, chair)
Prof. dr. Bart Bogaerts (VUB, secretary)
Dr. Catharina Olsen (VUB)
Dr. Guillaume Smits (ULB)
Prof. dr. Catia Pesquita (University of Lisbon, Portugal)
Prof. dr. Kris Laukens (UAntwerpen)

Curriculum vitae

Alexandre Renaux obtained his MSc in Bioinformatics at the University of Rouen, France, in 2014. After working as a bioinformatics engineer at the CEA in France and at the EMBL-EBI in the UK, he embarked on his PhD research at the crossroads of human genetics, machine learning, and knowledge representation.
During his PhD, he co-authored four peer-reviewed articles, two of which as the first author, and gave three talks at national and international conferences. He also developed and publicly released a genetic analysis web platform, a knowledge graph, and an open-source software package. Additionally, he co-supervised two MSc students in their master's theses.

Abstract of the PhD research

Recent years have seen significant progress in medical genetics, with enhanced access to sequenced genomic data and the evolution of computational methods for analysing genetic variation. These advances have improved our understanding of Mendelian genetic models, characterised by ‘one gene - one phenotype' relationships. However, the shift towards the study of oligogenic diseases, where a few genes jointly contribute to disease manifestation, remains a substantial challenge.

With the rising reports of clinical oligogenic cases, resources and machine learning tools have emerged to leverage this data. Nevertheless, despite their high accuracy, these predictors can be viewed as 'black-boxes' due to their limited interpretability. This complexity poses challenges for medical professionals in validating these predictions and restricts their understanding of potential underlying disease mechanisms. Our research aimed to tackle these limitations using structured background knowledge to provide additional context to predictions.

Our first key contribution aimed to provide better interpretability for existing black-box predictors for oligogenic disease diagnosis. This objective was achieved by implementing post-hoc interpretability techniques and by placing predicted pathogenic gene networks within background biological knowledge. These tools have been made accessible through a user-friendly web platform, enabling geneticists to submit, filter, predict, analyse, and interpret individual patient variant data.

To expand our understanding of these genetic interactions, we constructed a biomedical knowledge graph that integrates data from known oligogenic diseases with multi-scale biological networks. This novel resource facilitates the formulation of complex queries and inference tasks, underscoring the importance of heterogenous biological information in exploring the intricate and multi-factorial origins of oligogenic diseases. Building upon our knowledge graph, we devised a white-box machine learning approach that can provide both predictions and meaningful explanations. This interpretable model leverages path semantics between gene pairs and explain its prediction by highlighting relevant entities and relationships in the graph. This new method enables geneticists to validate predictions and hypothesise about the causal mechanisms underlying oligogenic diseases.

In conclusion, this PhD research highlights how the incorporation of background knowledge can enrich the interpretation of pathogenic gene interactions, ultimately advancing our understanding of oligogenic diseases.