



Brainteaser



LASIGE
driven by excellence



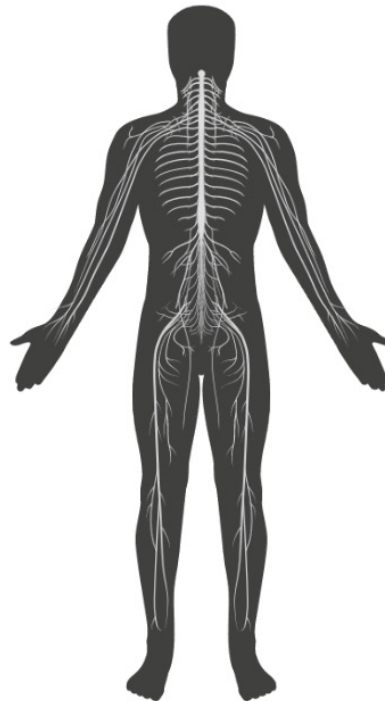
Explaining Artificial Intelligence Predictions of Disease Progression with Semantic Similarity

Susana Nunes, Rita T. Sousa, Filipa Serrano, Ruben Branco, Diogo F. Soares, Andreia S. Martins, Eleonora Auletta, Eduardo N. Castanho, Sara C. Madeira, Helena Aidos and Catia Pesquita

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal



Amyotrophic Lateral Sclerosis (ALS) and IDPP: CLEF 2022



ALS Progression leads to

- ➔ Non-invasive ventilation (NIV)
- ➔ Percutaneous endoscopic gastrostomy (PEG)
- ➔ Death

Why do we need
Explainable Artificial intelligence?

Understand
predictions

Trust
predictions

Machine Learning approaches are being
develop to predict disease progression.

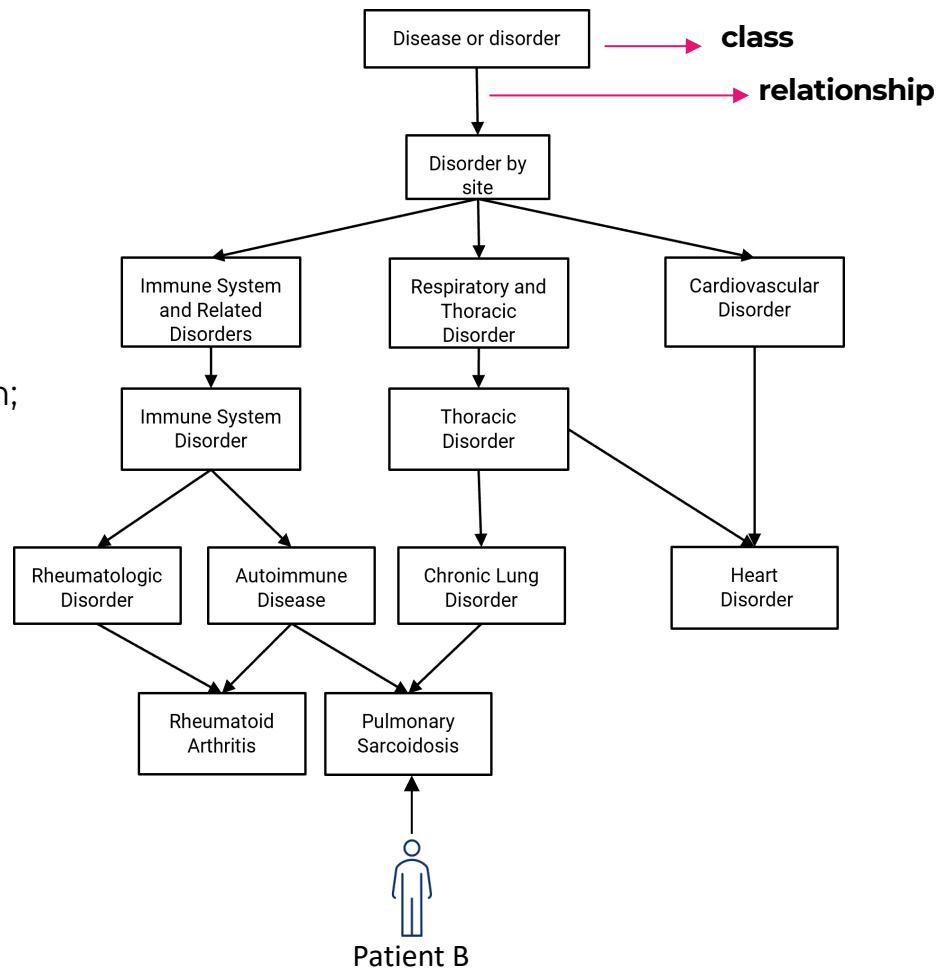


Introduction

Ontologies are a unique opportunity for explainability

Ontologies are semantic models that

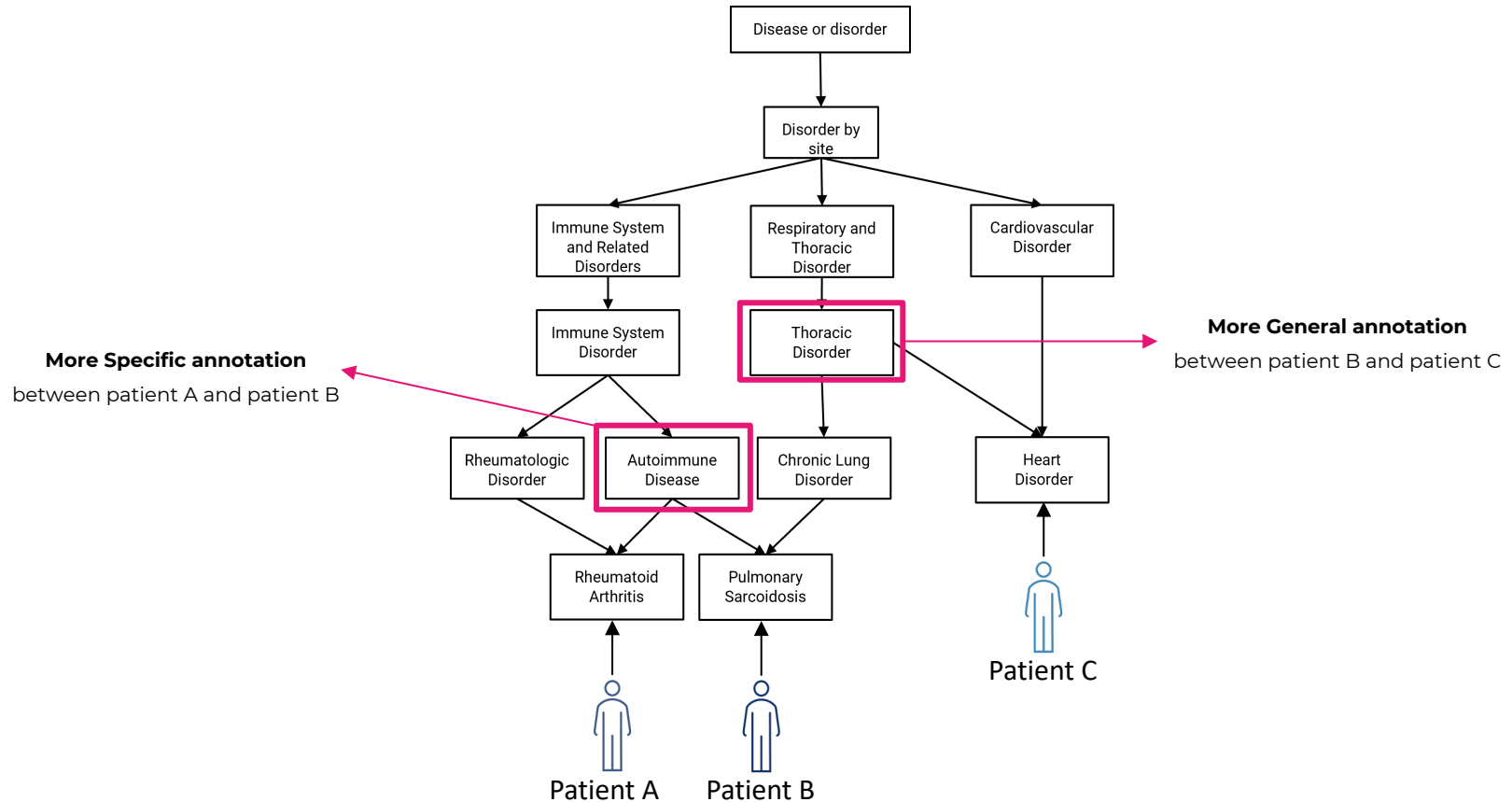
- encode the knowledge of a domain;
- can be used to describe the input data and outcomes.





Proposal

Semantic similarity can be used as an explanatory strategy



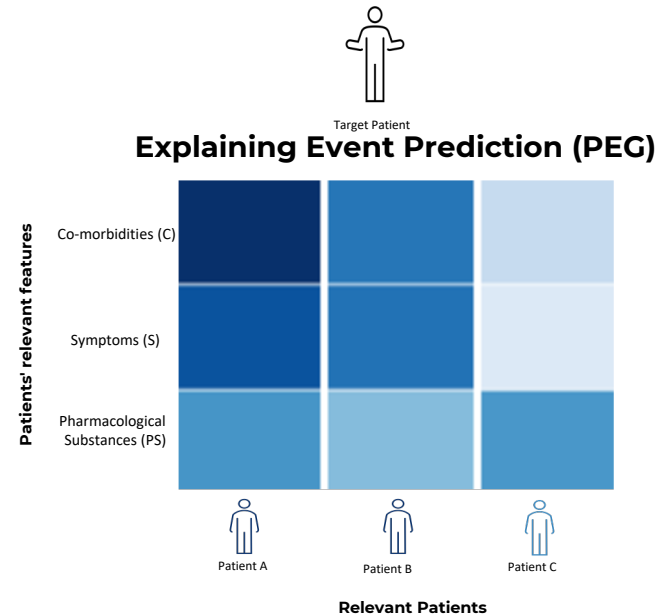
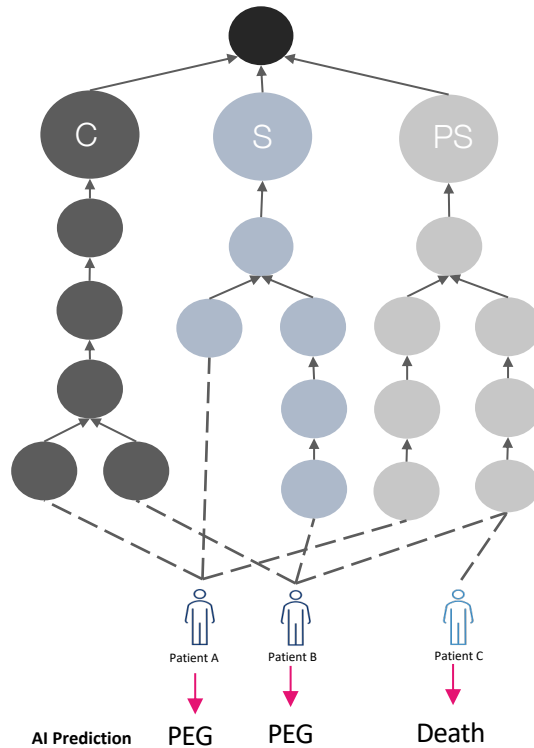
If both patients are annotated with the same ontology, they can be compared by comparing the classes with which they are annotated.

$$SS(P_B, P_A) = 0.8 > SS(P_B, P_C) = 0.3$$



Proposal

Semantic similarity-based explanations for patient-level predictions

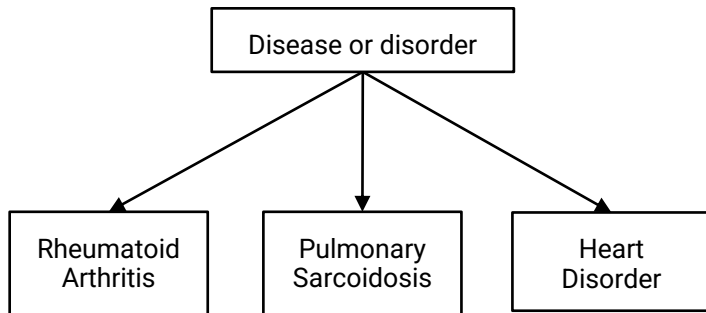


Explain the prediction for one patient by considering aspect-oriented semantic similarity with other relevant patients.

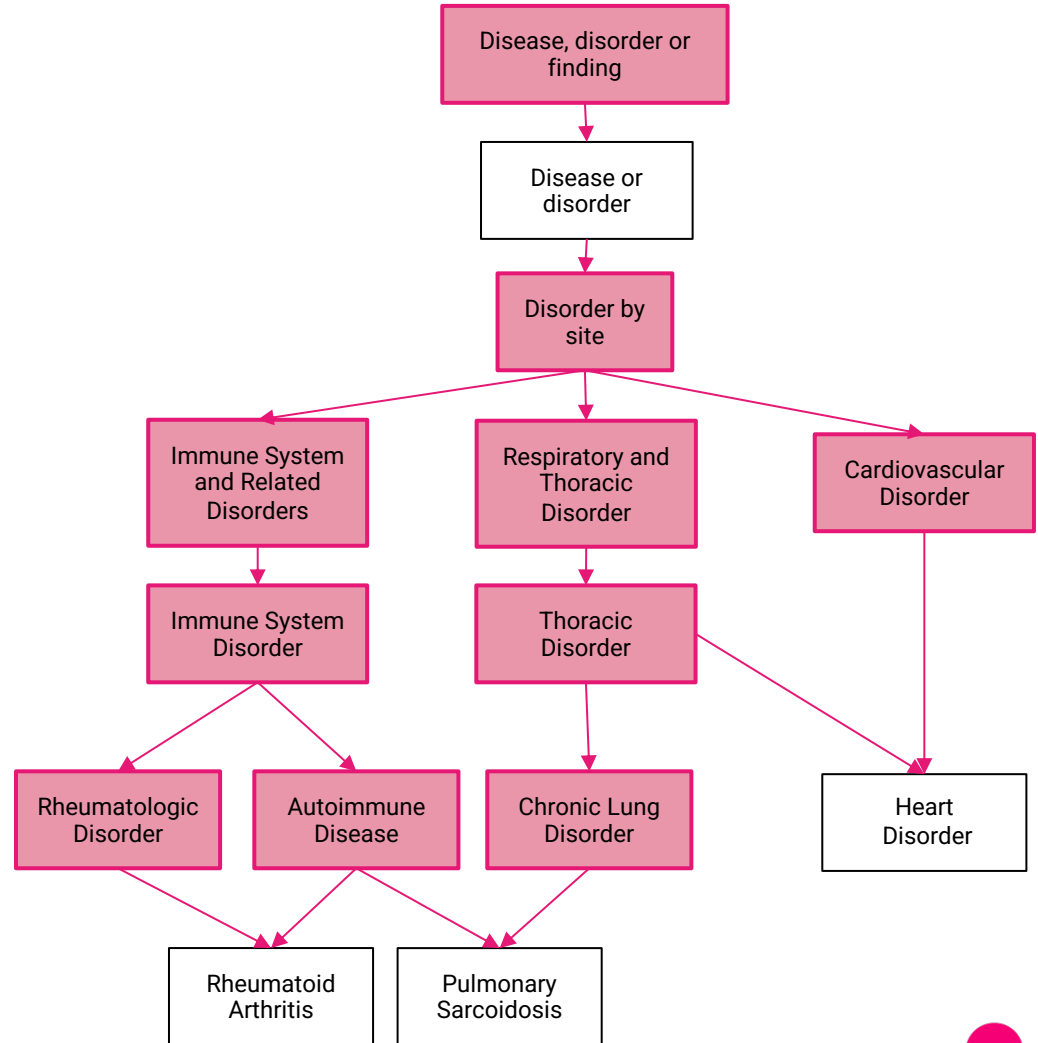


BO Enrichment Impact

Brainteaser Ontology



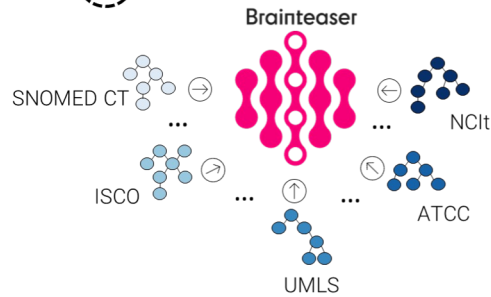
Enriched Ontology





Methodology

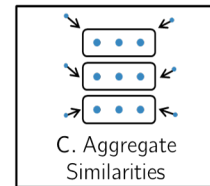
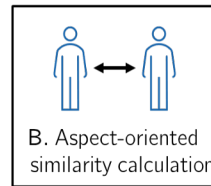
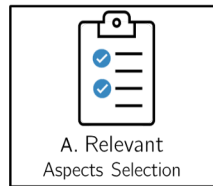
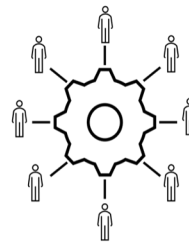
1 Ontology Integration



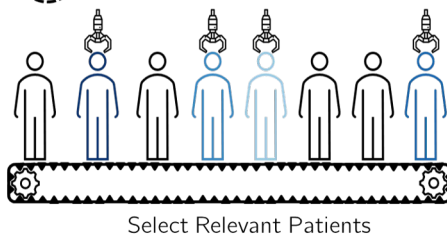
2 Semantic Annotation



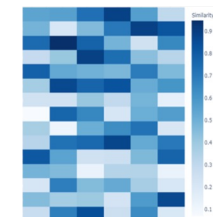
3 Similarity Calculation



4 Patient Selection



5 Explanations Visualization



Build Heatmap



Ontology Integration

Main ontologies and number of classes used to enrich the Brainteaser Ontology

Brainteaser Ontology



362 classes

Ontology	Classes
National Cancer Institute thesaurus (NCIt)	173,001
Ontology of Genes and Genomes (OGG)	69,689
Uberon	20,849
Medical Action Ontology (MAxO)	15,086
Ontology for MIRNA Target (OMIT)	90,916
SNOMED Clinical Terms (SNOMED CT)	358,483
Anatomical Therapeutic Chemical Classification (ATCC)	6,567
International Standard Classification of Occupations (ISCO)	619
Experimental Factor Ontology (EFO)	35,470

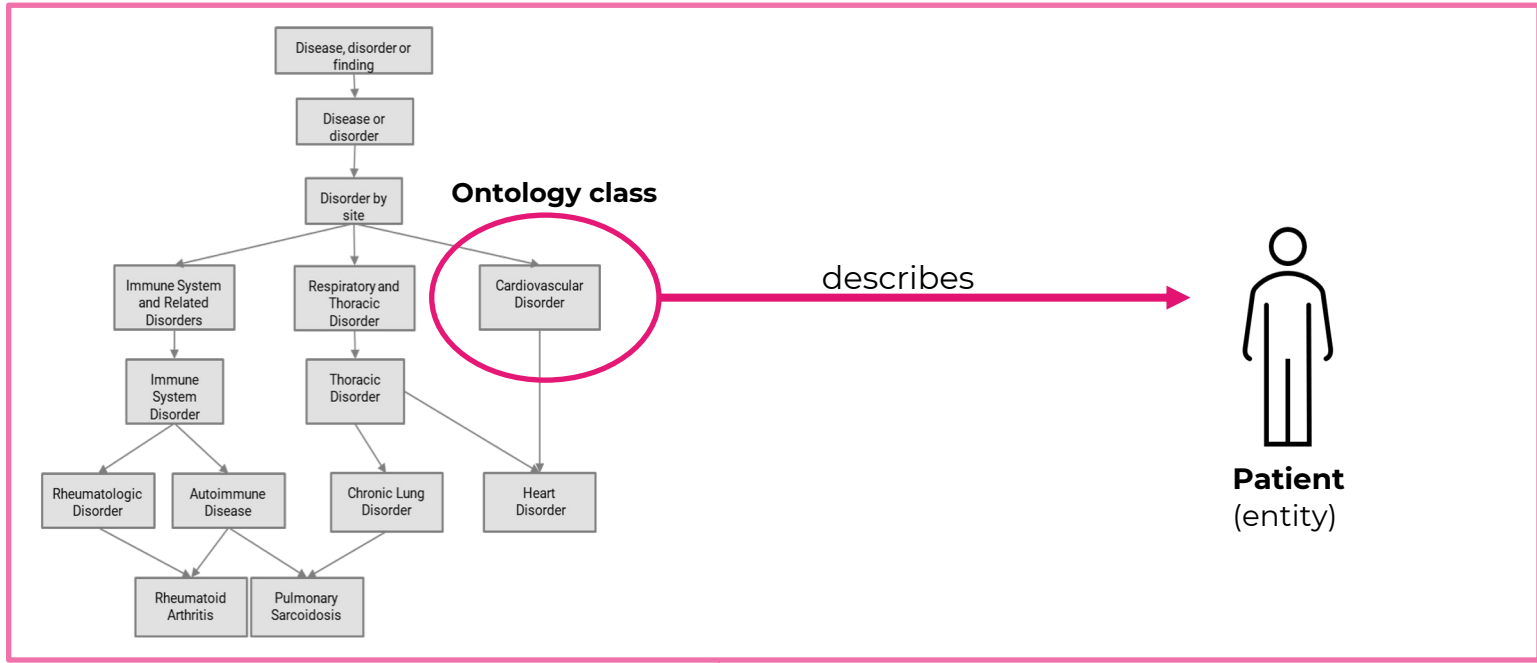
~ 770,680 classes






2

Patient Semantic Annotation



Annotation

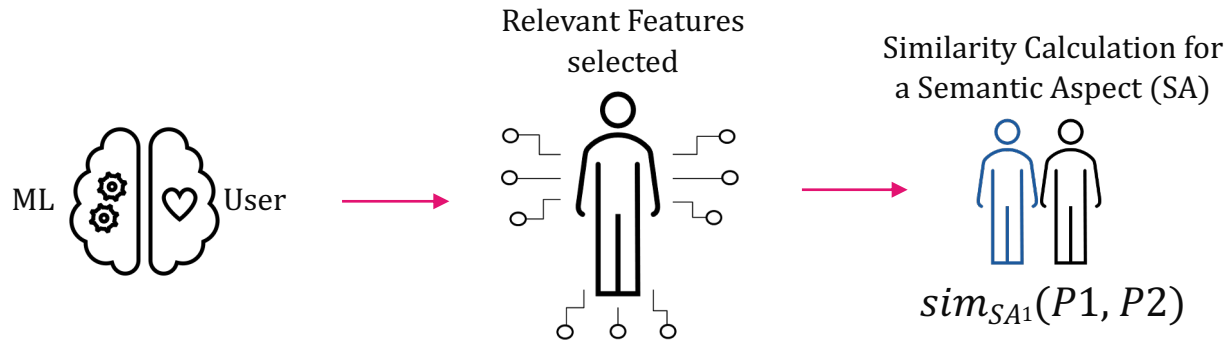
$a = \langle \text{patient}, \text{ontology class}, \text{time point} \rangle$

 $P = \{a_1, \dots, a_n\}$

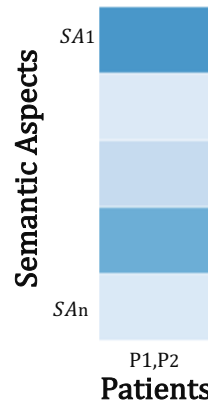


3

Semantic Similarities between Patients



Similarity Matrix

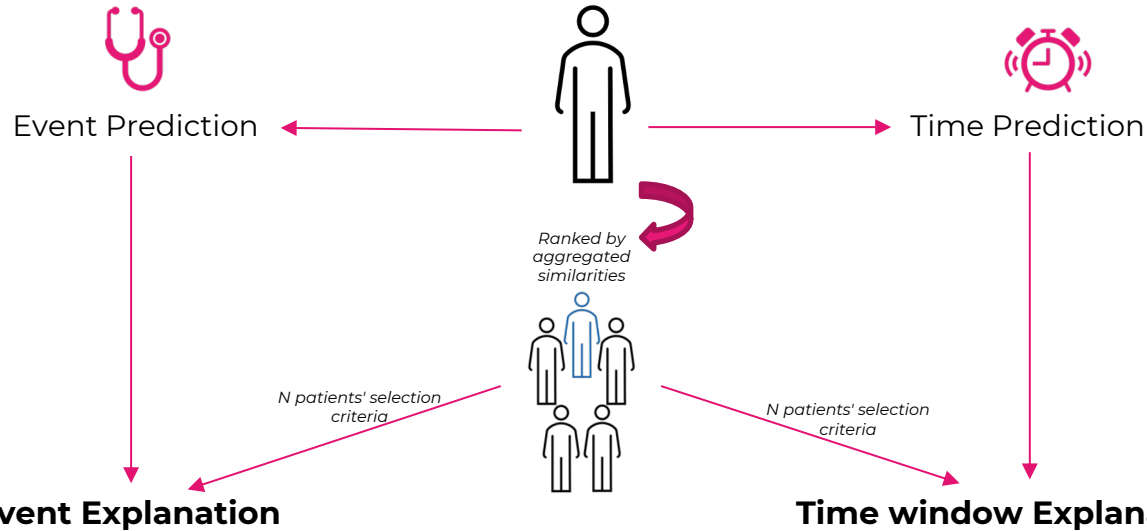


$$sim(P1, P2) = \{sim_{SA1}(P1, P2), \dots, sim_{SA_n}(P1, P2)\}$$



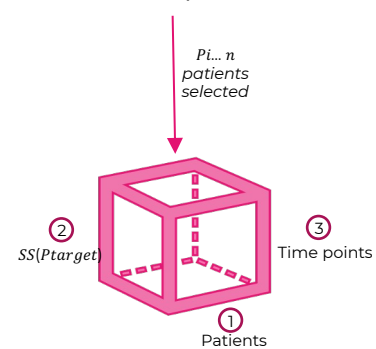
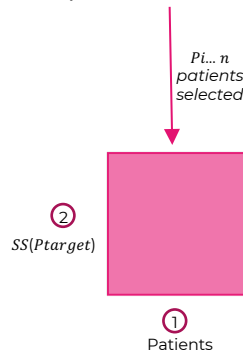
4

Building Semantic Similarity Explanations



1. The N **most similar** patients with **same outcome**;
2. The N **least similar** patients with **same outcome**;
3. The N **most similar** patients with **different outcome**.

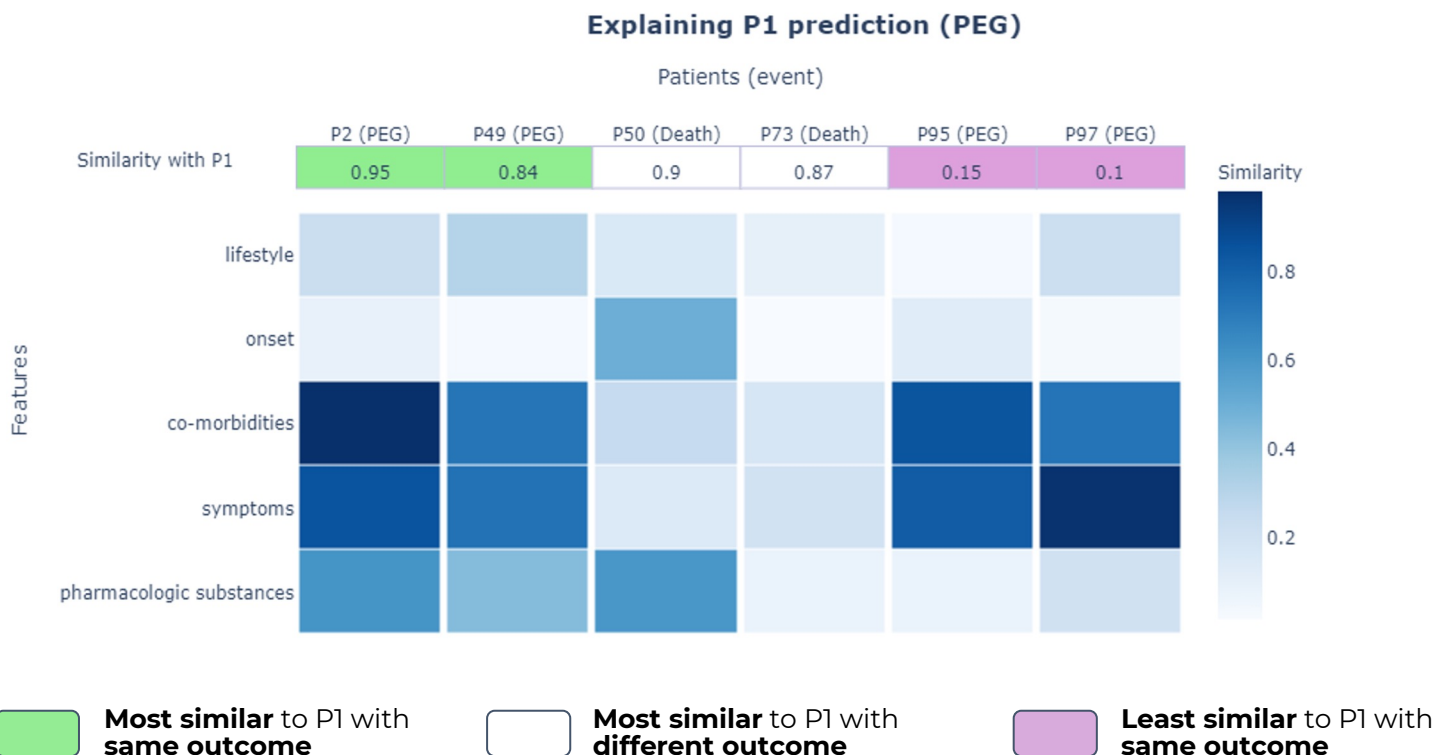
1. The N **most similar** patients in **same time window**;
2. The N **least similar** patients in **same time window**;
3. The N **most similar** patients in a **different time window**.





Semantic Similarity Explanations

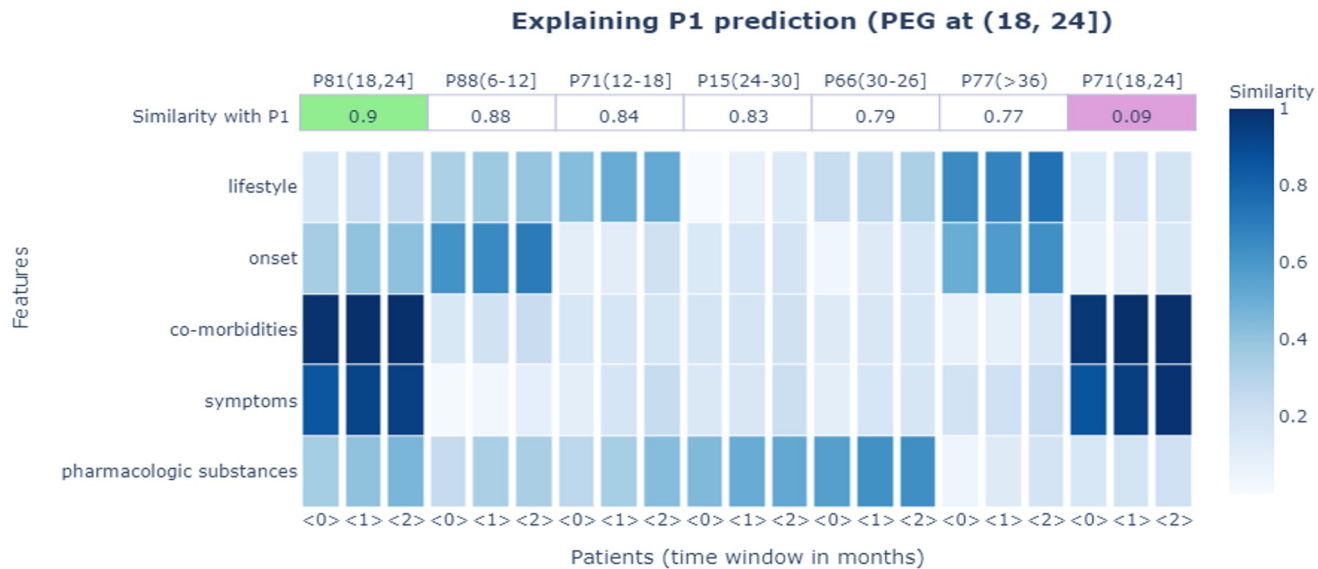
Explaining Event Prediction






Semantic Similarity Explanations

Explaining Time Window to the Event Prediction



 **Most similar** to P1 with **same outcome** and **time** to event

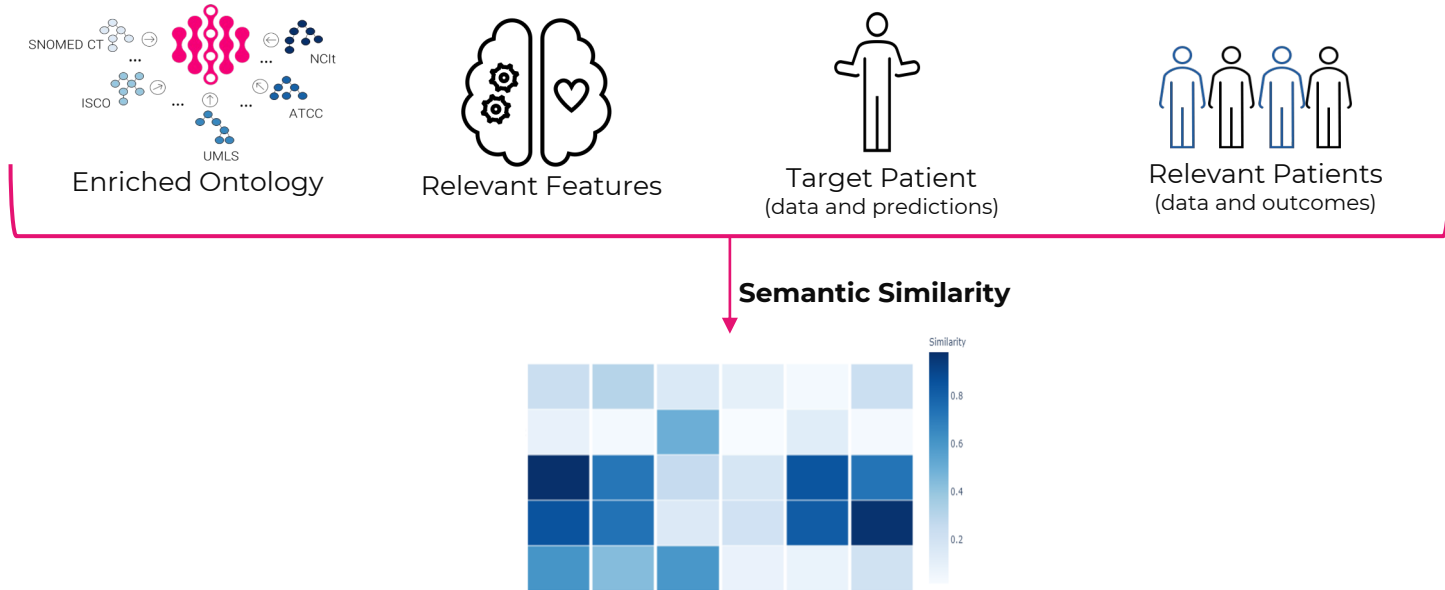
 **Most similar** to P1 with **same outcome** but **different time** to event

 **Least similar** to P1 with **same outcome** and **time** to event



Conclusions

Our Proposal



Future Work

- Integration with other similarity types.
(comparisons of ALSFRS-R scores).
- Explanations value measured in user studies.



Thank You



The authors are funded by the FCT through LASIGE Research Unit (ref. UIDB/00408/2020 and ref. UIDP/00408/2020), AlpALS (PTDC/CCI-CIF/4613/2020), PhD research scholarships to RTS (SFRH/BD/145377/2019), DFS (2020.05100.BD), ENC (2021.07810.BD); and by the BRAINTEASER project and KATY project that have received funding from the European Union's Horizon 2020 research and innovation programme under the grants agreements No 101017598 and No 101017453, respectively



scnunes@fc.ul.pt



[@S_Nunes](https://twitter.com/S_Nunes)