ÉCOLE DOCTORALE GALILÉE
Université Sorbonne Paris Nord

Contrat doctoral – ED Galilée

<u>Titre du sujet :</u> Explainable Prediction Models using Analogical Reasoning: Application to medical datasets

- Unité de recherche : Laboratoire de recherche en informatique pour la santé–LIMICS-UMRS 1142
- Discipline : Informatique
- Direction de thèse : Karima Sedki, MCF HdR, co-encadrants : Fadi Badra, MCF & Bounhas Myriam, MCF, HdR
- Contact : **karima.sedki@univ-paris13.fr**
- Domaine de recherche : Intelligence artificielle, méthodes d'explicabilité, données médicales
- Mots clés : Modèles de prédiction, raisonnement analogique, explicabilité, données médicales

---

Medical data is available in a huge amount and is increasingly growing thanks to the large use of electronic health records, wearables and mobile Apps, etc. This data is rich in detail but complex as is composed of **structured data** which is organized and standardized and generally stored in databases with predefined formats (e.g., patient demographics, vital signs, laboratory results), and **unstructured data** which is not organized into predefined formats (e.g., clinical notes and discharge summaries). **MIMIC** [1] is an example of a well-known medical data set, it contains a large amount of information such as admissions, movement of patients between one unit to another, and prescriptions. This data is an important resource for applying artificial intelligence (AI) in healthcare research.
Developing advanced AI techniques to learn predictive models and explain them is a very active topic of research [2,3,4,5,6]. In the medical setting, such approaches aim to predict patient outcomes (mortality, risk of developing diseases), assist in diagnosis, support treatment planning, personalize treatments, and explain decisions. However, many clinicians still hesitate to adopt AI recommendations due to the lack of interpretability, justifiable reasoning, and adaptability. Thus, **more advanced AI approaches are needed to improve the transparency of predictive models trained on medical data, to reflect how clinicians' reason, and to better explain decisions**. **Reasoning by analogy** (AR) [3,4] is an advanced approach that mimics human cognition by analyzing similarities and differences between new and previously solved problems. Recent research works demonstrate that AR is promising to enhance explainability of machine learning models, as both the inference process and the knowledge it uses (such as a case base and similarity/dissimilarity measures) can be inherently interpretable. Central to this approach is the concept of **analogical proportions**, which formally represent relationships between items as *"A is to B as C is to D"* or *A:B::C:D*. This denotes that the difference between A and B is analogous to the difference between C and D. Analogical inference takes advantage of these proportions to predict unknown feature. A simple example of analogical inference in the context of medical setting involves three patients with different symptoms and treatments as follows: Case A: Patient with infection A was treated with antibiotic X and recovered. Case B: Patient with infection B was treated with antibiotic Y and recovered. Case C: Patient with infection C was treated with antibiotic X but did not recover. Now given a new case D: Patient with infection D is similar to infection C in some traits and similar to infection B in others. Because D shares traits with both C (resistant to X) and B (responsive to Y), it's reasonable to hypothesize that D might respond to Y, like B.

The aim of this PhD thesis is to **develop an analogy-based framework to learn explainable models with application to the medical datasets**. The key objectives are to address the following challenges:

1) **Reduce the impact of high dimensionality:** Medical datasets is described with an important number of features, many of them may be irrelevant. Identifying which features are clinically relevant is an important issue to address to avoid misleading comparisons. The problem concerns also the number of cases in the data base (patient records). So, developing competence models of the case base could be of great help for medical decision support. For example, similar patients

have similar diagnostics, but is it possible to identify a small set of key patients that a new patient ought to be compared to in order to perform prediction by analogy.

2) **Integrate clinical context, causal Information and expert knowledge:** Raw numerical features often fail in capturing the clinical reasoning or causal relationships. As a result, analogies based only on such data can miss the underlying justifications or symptoms of predicted diagnosis or treatment. In addition, study the role of expert knowledge can have to tune the learning process.
3) **Develop analogy-based systems to provide explanations**: Presenting relevant explanations in an understandable manner is fundamental for the acceptability and effective use of developed tools in the medical domain. Among the major problems that limit existing approaches to explanation are: (i) the overwhelming number of explanations, which compromises their interpretability and usability in practice, and (ii) the level of explanations, which is often very basic and low-level, such as merely listing the values of attributes that influence certain predictions
Analogical learners can answer questions like "Why this treatment?", "Why not this medicine?", and "Which treatment should be recommended for a given patient?" Therefore, the necessity to develop analogy-based systems that provide explanations for decision-making becomes crucial.

A research idea to explore for addressing these challenges during this Phd thesis is to use with **analogical reasoning more context-aware methods like Argumentation mining** [7]. It allows automatically identifying and extracting arguments and counter arguments from texts (particularly form unstructured data or from experts knowledge which constitutes also a precious piece of information in the medical setting). An example of argument is in the form ***Premise → Claim***, where Claim is a conclusion (e.g., Sepsis is suspected for the patient), and Premise concerns reasons that support or oppose the claim (e.g., the patient has a persistent fever above 38.5°C and hypotension). Combined with analogical reasoning, argumentation mining can make predictive models more efficient and more explainable. It adds logical, human-like reasoning structures, prioritize relevant clinical evidence, captures relationships, and produces transparent, structured justifications for predictions.

Data: The analogy-based framework will be validated on mimic data and another real datasets in the domain of neuro-cognitive disorders in older adults. The Cognum dataset [8] is a small dataset (164 patients, 129 attributes) containing the answers of a set of questionnaires, cognitive and manual dexterity tests carried out on a digital tablet. Each patient is described by a set of answers to the tests, and labeled with its real diagnostic. The NACC UDS dataset (https://naccdata.org) is a larger dataset that contains longitudinal neurocognitive and clinical phenotypic data about more than 20,000 patients with or without dementia. Data is collected annually for each patient, so the dataset can be used to detect early signs or predict the evolution of the disease for a patient by leveraging the data collected for similar patients in their follow-up visits.

Bibliographie

[1] Johnson et al. Mimic-iv (version 0.4). PhysioNet, 2020.
[2] R. Tsopra, JB. Lamy, K. Sedki. Using preference learning for detecting inconsistencies in clinical practice guidelines: Methods and application to antibiotherapy. Artif. Intell. Medicine 89: 24-33 (2018)
[3] H. Prade and G. Richard. Analogical proportions and multiple valued logics. In L.C. van der Gaag, editor, Proc. 12th Eur. Conf. of Symb. And Quantit. Appr. To Reasoning with uncertainty (ECSQUARU'13)
[4] M. Bounhas et H. Prade. Analogy-based classifiers: An improved algorithm exploiting competent data pairs, International Journal of Approximate Reasoning, vol. 158, p.108923, 2023.
[5] F. Badra, E. Marquer, MJ. Lesot, M. Couceiro, and D. Leake. EnergyCompress: A General Case Base Learning Strategy, to appear in Proc. of the Int Joint Conf. on Artificial Intelligence, IJCAI'25.
[6] Guidotti et al. A Survey of Methods for Explaining Black Box Models. ACM Comput. Surv. 51, 5, Article 93, 42 pages, 2019.
[7] T. Mayer, S. Marro, E. Cabrio, and S.Villata. Enhancing evidence-based medicine with natural language argumentative analysis of clinical trials. Artificial Intelligence in Medicine, 118:102098, 2021.
[8] N. Um Din, F. Maronnat, Q. Le Boterff, S. Otmane, F. Badra, P. Lindberg, J. Belmin. Contribution of a Tablet-Based Tests to the Diagnosis of Neurocognitive Disorders in Older Adults: A Feasibility Study. 34th Medical Informatics Europe Conference (MIE 2024), Athens, Greece. pp.1163–1164, 2024.