PRINCIPAL ACADEMIC TUTOR
Emanuele Della Valle, DEIB, PoliMi

ACADEMIC TUTORS
Marco Agostino Deriu, DIMEAS, PoliTo
Michela Sperti, DIMEAS, PoliTo

EXTERNAL INSTITUTION
Dedalus Italia S.p,A.

EXTERNAL TUTORS
Davide Guerri, Dedalus Italia S.p.A.
Simone Paolucci, Dedalus Italia S.p.A.
Riccardo Banali, Dedalus Italia S.p.A.

TEAM MEMBERS

Andrea Cavallo, Computer Engineering, PoliTo

Davide Fassino, Mathematical Engineering, PoliTo

Karim Kassem, Biomedical Engineering, PoliMi

Andrea Mario Vergani, Computer Science and Engineering, PoliMi

YouForAll

Your digital twin for allowing a healthy society

Executive summary

The current healthcare system does not take full advantage of the huge potential of new technologies, which could revolutionise patient care and treatment. YouForAll (Your digital twin for allowing a healthy society) is a project from the Alta Scuola Politecnica XVII cycle that, in cooperation with Dedalus Italia, a company active in the world of clinical software, undertakes the challenge to move a first step towards this revolution. Indeed, the goal of the project is to realise an efficient system to support clinical decisions based on Artificial Intelligence (AI) techniques. The main innovations of this system with respect to existing ones are the high efficacy of the support offered and the clear explanations behind the provided suggestions, which can be easily understood by doctors.

As a matter of fact, the current approach to patient care still relies for the most part on clinical guidelines, sets of rules that define questions to be asked to patients based on their symptoms and actions to be undertaken based on their answers (e.g. what is the final diagnosis or which exams should be performed given the patient's health status). However, these guidelines fail to personalise the therapies and procedures to the specificities of each patient, leading to suboptimal treatments. Moreover, they need to be manually consulted by clinicians, resulting in poor healthcare efficiency.

In this context, YouForAll proposes NEAR (Neural imputed Explainable and Adaptive Risk score), an AI-based score to predict the risk of a clinical event. As proof of concept, NEAR is implemented to predict the risk of death and bleeding events for patients who have already suffered from cardiac disease, but the approach can be effortlessly extended to any other clinical condition for which a sufficient quantity of data from patients is available. Beyond its capability to model, in advance, the risk for a patient to incur in a given clinical condition, NEAR also provides easy-to-interpret explanations about its own predictions, thanks to which a clinician can immediately understand the clinical variables that contribute more to the score. Moreover, NEAR suggests actions to mitigate the risk or to have a more accurate prediction of the likelihood of the clinical event. Therefore, NEAR acts as a clinical decision support system for practitioners, who can integrate the suggestions with their professional experience to improve their diagnoses.

Key Words:

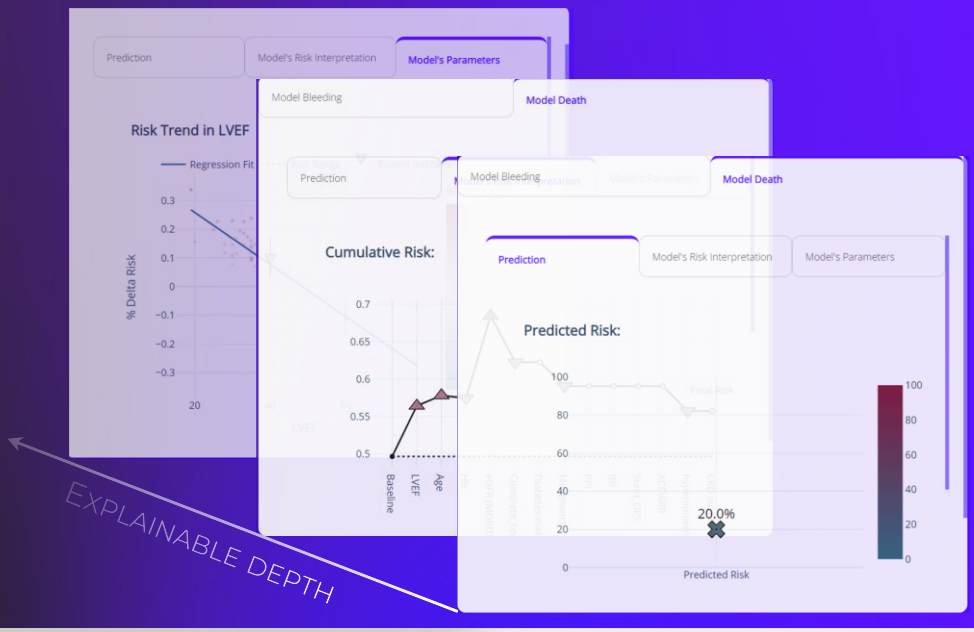
digital twin; healthcare; machine learning; clinical decision support system; explainability

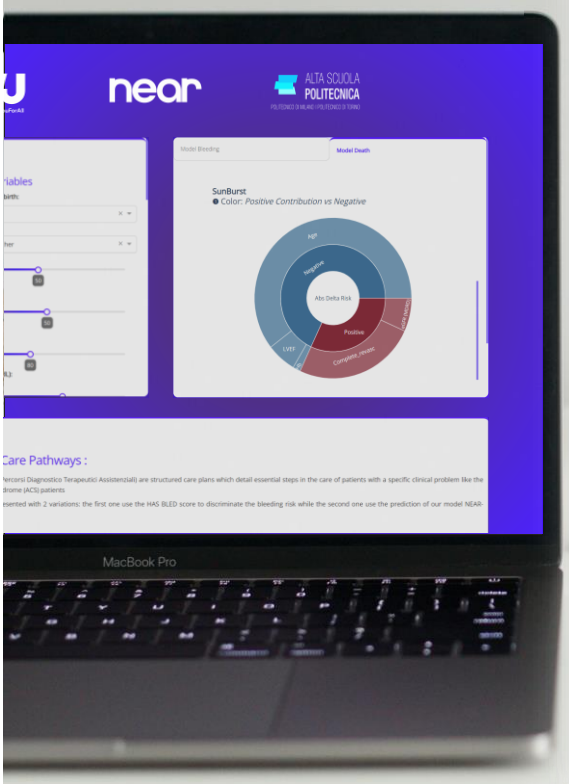
Your digital twin for allowing a healthy society

YouForAll

near

NEURAL IMPUTED EXPLAINABLE AND ADAPTIVE RISK SCORE





- **INTUITIVE WEB APP**
- **EXPLAINABLE SUGGESTIONS**
- **PERSONALIZED DIAGNOSES**
- **WORKFLOWS DIGITIZATION**

VISIT **NEAR-APP.HEROKUAPP.COM**

Project description written by the Principal Academic Tutor

The usage of Artificial Intelligence (AI) technology in health system will grow a lot in the next years. The main difficulty in this process stands in the fact that the clinicians would hardly trust the predictions by a black-box algorithm. Moreover, the present situation regarding healthcare digitization is even a step back: the clinical guidelines, which constitute the current decision support tool for clinicians, are not digitized in most of the situations; this translates into a low efficiency in the clinical decision-making process.

The YouForAll project fits in these problems. The team would be aware of the difficulties of the data collection in healthcare, and they would implement a tool that uses machine learning techniques for the classification of death and heart bleeding events starting from clinical, therapeutic and procedural features. The tool would be used as a clinical decision support system by the clinicians; for this reason, a great effort should be spent in the "explainability" of the model. A suitable user interface and simpler models would be integrated to help the clinicians to interpret the result of the AI predictions, giving the possibility to understand how each feature influences the final score for every specific patient.

In addition to the realization of such a data-driven system for the medical staff, the team would also concentrate on the digitization of the clinical guidelines, the current paradigm for clinical decision support, through the use of standard information technology languages in healthcare that would foster integration with health information systems and with electronic health records.

A great contribution for the realization of this project is due to Dedalus Italia S.p.A. The team had the occasion to learn how the software applications for healthcare are developed, as well as to understand the deep significance of interoperability, which has been kept in mind in the realization of the project, in order to promote health digitization.

Team description by skill

The YouForAll team is composed of four members:

- Karim Kassem, the team leader, a Biomedical Engineering student (specialisation in Technologies for electronics) from Politecnico di Milano;
- Andrea Cavallo, studying in the Master course of Computer Engineering at Politecnico di Torino, with a background in Electronics Engineering;
- Davide Fassino, currently a student of Mathematical Engineering at Politecnico di Torino, specialising in Mathematical models and numerical methods;
- Andrea Mario Vergani, Computer Engineering student from Politecnico di Milano, following a pathway in data science.

The reduced number of components of the team was initially considered only as a disadvantage. Indeed, for most of the activities, it was not possible to split the team in smaller subgroups. In the following months, we understood that it could also be a resource: all the members could contribute to each stage of the project, giving their own opinions and highlighting which aspects they considered critical and should have been modified.

Goal

Our primary goal is to develop a tool to help the clinicians to achieve more efficient and personalized diagnoses, which can be trusted and integrated with medical knowledge. We decline this main goal into the following sub-goals:

1. Digitization of the Integrated care pathways (the current decision support tools for clinical diagnoses). *The clinical pathways are nowadays static a priori rules provided by the regional healthcare systems to the clinicians. The first goal of our project consists in translating them into standard languages for an automatic integration with electronic health records, in order to improve the efficiency of diagnoses.*

2. Data-driven personalized diagnoses. *The second step consists in changing the paradigm of the current clinical pathways: we aim at shifting from a system of static rules to new rules driven by real data. The importance of clinical variables would then be targeted on the single patient, thus going in the direction of the personalization of diagnoses..*

3. Explainable suggestions. *We would like to open the black-box behind our data-driven clinical guidelines, in order for the clinicians to be able to interpret and "cooperate" with them. Indeed, our aim is not to substitute the doctors, but to implement a tool that can be integrated with the expert knowledge.*

4. Proof of concept for interoperability. *The key point to effectively use our tool in the real world is its interoperability with health information systems. For this reason, we need to keep an eye on the possibility of integrating our solution with the already existing clinical software tools and infrastructures, as well as with the electronic health records.*

Understanding the problem

In order to frame a complex problem like the one regarding healthcare digitization, the YouForAll team starts with an analysis of the involved stakeholders and their needs, for being able to translate the latter into system requirements.

In particular, in the context of our project, one of the main stakeholders is constituted by the clinicians and the medical staff, who ask for a higher efficiency and efficacy in the process of support for clinical decision-making; moreover, the practitioners' need is to have a decision support system with which they could "cooperate": basically, not simply a "black-box" predictor that aims at substituting the clinicians' role, but a system that allows them to integrate their expert knowldege for more accurate diagnoses and treatments.

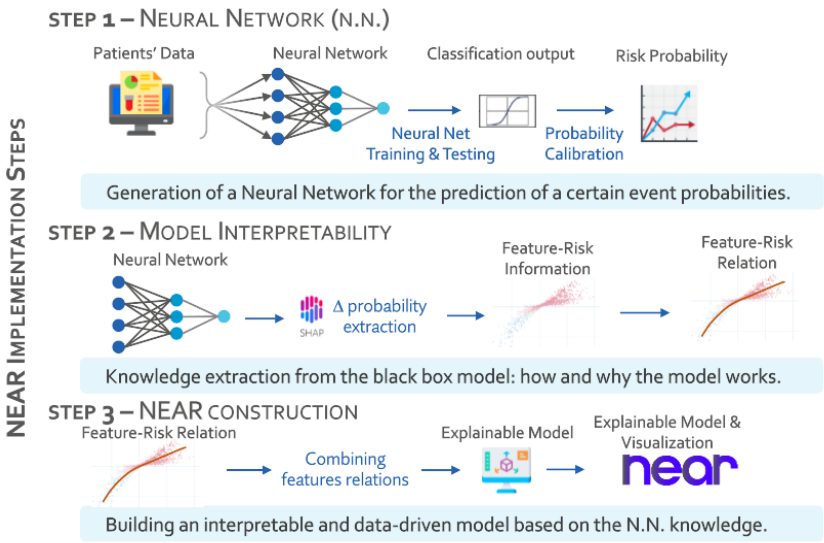
Of course, patients represent another class of important stakeholders, who feel a lack in the digitization of healthcare. They would welcome a massive shift to electronic health records, the exploitation of "big" medical data for advancements in the field, novel diagnosis and early detection methods, … On the other side, however, even in the case of a more digitized healthcare, patients would not like to lose the contact with the clinician, who is often seen as an expert, as a professional and, in some cases, as a trusted person.

The healthcare system and policy makers, instead, are trying to promote innovation in healthcare, but they are also aware that gradual change is necessary, especially in a field in which the main focus is the health of people. For instance, responsibility challenges, privacy concerns and property rights on data are currently limiting a full digitization of healthcare, as well as the introduction of artificial intelligence techniques and tools in the field of diagnosis, for instance.

Finally, information technology (IT) companies in the healthcare sector are currently working to fill the gap of digitization in the field. In order to foster integration and interoperability among different information systems, a crucial element in healthcare, they mainly work with IT clinical international standards.

From the stakeholders analysis, it is evident that one of the requirements for our project is to promote the use of real world clinical data, which are collected but whose potential is still not fully exploited. Of course, in order to help the clinicians, the suggestions deriving from data-driven techniques should be interpretable: explainability is one of the keys that we should achieve to support the clinical decision-making process.

Moreover, the care for interoperability and IT clinical standards is necessary, too, in order to promote a faster and effective healthcare digitization.



Main implementation steps behind the development of NEAR, our solution.

Exploring the opportunities

The issue of digitization in the healthcare sector and the introduction of new explainable and data-driven tools is extremely dynamic and prolific in terms of researches and new solutions. A new paradigm that is gaining momentum in the healthcare sector is that of the digital twin: a virtual model of a physical entity that can go from the whole human body to more specific body systems or functions or organs. A digital twin can be used to perform disease prediction and treatment simulation on specific individuals, detecting possible diseases in advance and suggesting the most appropriate treatments, behaving like a Clinical Decision Support System (CDSS). Nowadays CDSSs aim to improve medical decisions through the use of various types of health information. They can be divided into knowledge-based and non-knowledge-based systems.

The first category includes fixed and deterministic hand-crafted rules evaluated on the data retrieved by the system itself. Rules can be based on the literature, practice or patient evidence. Non-knowledge-based CDSSs, conversely, leverage artificial intelligence and machine learning techniques to automatically generate effective rules from a data source of past observations. The inferred AI rules must be interpretable and explainable. In many real world scenarios, it is fundamental to justify the decisions in order for the people to trust the algorithm and effectively use it. In August 2020, the National Institute of Standards and Technology published "Four Principles of Explainable Artificial Intelligence": such principles include concepts like explanation, meaningful, accuracy and knowledge limits.

Some knowledge extraction models are intrinsically explainable like the "Decision Trees" (which predict the value of a target variable by learning simple decision rules inferred from the data), while others are defined as "black-box models", since it is not straightforward to provide an explanation for their predictions without ancillary tools.

Generating a solution

NEAR, our solution, starts from a neural network, a popular model in AI that learns relevant patterns from data to classify an item based on its features (in our situation, classify a patient as likely subject to a clinical event, or not, based on her/his clinical information). Despite their high classification performances, neural networks have a relevant drawback: they are hard to interpret. Since the explainability of the score is one of the cornerstones of the project and it is necessary for the model to be effectively integrated into the healthcare ecosystem, a model explainer, SHAP (SHapley Additive exPlanations), is applied to the neural network. This tool computes the impact of each input feature on the final prediction. Then, the feature importance is translated into weights that allow for the definition of a simplified model for risk prediction. With respect to the plain neural network, this simplified approach presents two advantages: it also works if some input features are missing and it can be interpreted very easily. Indeed, each feature provides a separate contribution to the final risk score, therefore not all features are needed and it is immediate to verify which features have a stronger impact for a patient. Both these characteristics are fundamental for the deployment of the system in real scenarios and represent two of the main innovations introduced by YouForAll. To facilitate the integration of NEAR in clinical operations, a web application provides a user-friendly interface to visualise the risk score of a patient based on the inserted clinical data, underline the most relevant features, define the reliability of the prediction and suggest which exams or analyses could improve the accuracy of the risk. The efficacy of NEAR was tested both in terms of classification performances, by measuring the accuracy on a test dataset, and of usefulness, by presenting the web application to a group of doctors from the Molinette hospital in Turin and to Dedalus Italia S.p.A., receiving positive feedbacks in both the situations.

Moreover, YouForAll deals with another concern that is very critical for the healthcare ecosystem: interoperability. Indeed, allowing data sharing and compatibility among clinical infrastructures could boost healthcare efficiency. With this awareness, YouForAll relies on common standards for healthcare data representation, namely FHIR (Fast Healthcare Interoperability Resources) and CQL (Clinical Quality Language). In addition to the creation of NEAR, the project involves the translation into these clinical standards of the integrated care pathways used as guidelines for patient treatment. Thanks to this, clinical guidelines can be integrated in the currently adopted software tools (as the ones provided by the external partner Dedalus) and automatically provide suggestions to doctors during clinical decision-making. Moreover, also the NEAR score, thanks to its explainable nature, can be easily implemented using these standards, therefore allowing its integration into the existing clinical infrastructures.

Main bibliographic references

F. D'Ascenzo, O. De Filippo, G. Gallone, G. Mittone, M. A. Deriu, M. Iannaccone, A. Ariza-Solé, C. Liebetrau, S. Manzano-Fernàndez, G. Quadri, T. Kinnaird, G. Campo, J. P. S. Henriques, J. M. Hughes, A. Dominguez-Rodriguez, M. Aldinucci, U. Morbiducci, G. Patti, S. Raposeiras-Roubin, E. Abu-Assi, and G. M. De Ferrari, Machine learning-based prediction of adverse events following an acute coronary syndrome (PRAISE): a modelling study of pooled datasets, Lancet, 397:10270, 199-207, 2021.

P. J. Phillips, C. A. Hahn, P. C. Fontana, D. A. Broniatowski, and M. A. Przybocki, Four principles of explainable artificial intelligence, NIST (National Institute of Standards and Technology) Interagency/Internal Report (NISTIR), 2020.