

Editorial: The First Workshop on Explainable Artificial Intelligence for the medical domain - EXPLIMED@ECAI2024

Gianluca Zaza^{1,*}, Gabriella Casalino¹ and Giovanna Castellano¹

¹Computer Science Department, University of Bari Aldo Moro Bari, Italy

Abstract

The 2024 First Workshop on Explainable Artificial Intelligence for the Medical Domain (EXPLIMED) marks its inaugural edition in conjunction with the 27th European Conference on Artificial Intelligence (ECAI 2024) in Santiago de Compostela. This workshop brought together experts in Artificial Intelligence to deepen the latest innovations and best practices in Explainable AI (XAI) within the medical field. Participants engaged in discussions that covered recent trends, research initiatives, and emerging developments in XAI as they pertain to healthcare applications, emphasizing a multifaceted approach to understanding how these advancements can enhance medical practice and patient outcomes.

Keywords

Explainable Artificial Intelligence, Transparent models, Interpretable models, e-health, bioinformatics

Introduction

The EXPLIMED workshop is a pivotal forum that integrates Explainable Artificial Intelligence (XAI) within the medical field, focusing on the latest research, methodologies, and practical case studies. In an era where AI is becoming increasingly central to healthcare decision-making, the workshop aims to cultivate a collaborative platform for researchers, practitioners, and policymakers to exchange insights on enhancing transparency, interpretability, and trust in medical AI systems [1, 2]. The primary objective of EXPLIMED is to underscore the critical importance of XAI in ensuring that medical professionals and patients can fully understand and trust the outcomes generated by AI [3]. With the opacity of AI models posing significant concerns regarding accurate diagnoses, treatments, and patient care, this workshop emphasizes XAI's vital role in delivering understandable insights that empower healthcare professionals and patients alike [4]. Furthermore, EXPLIMED highlights the necessity of transparently communicating AI-generated outcomes to patients, enabling them to engage more actively in their healthcare decisions [5]. Addressing ethical considerations related to biases in AI systems, the workshop advocates for fair and equitable healthcare practices by elucidating the decision-making processes inherent in AI technologies [6]. Covering a broad spectrum

EXPLIMED - First Workshop on Explainable Artificial Intelligence for the medical domain - 19-20 October 2024, Santiago de Compostela, Spain

*Corresponding author.

✉ gianluca.zaza@uniba.it (G. Zaza); gabriella.casalino@uniba.it (G. Casalino); giovanna.castellano@uniba.it (G. Castellano)

🆔 0000-0003-3272-9739 (G. Zaza); 000-0003-0713-2260 (G. Casalino); 0000-0002-6489-8628 (G. Castellano)

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).



of topics—including post-hoc and ante-hoc methods for explainability, uncertainty modeling, and applications of XAI in medical imaging and video—EXPLIMED is committed to fostering interdisciplinary collaboration. By bringing together researchers, clinicians, and policymakers, the workshop promotes the adoption of XAI in healthcare, ultimately contributing to developing reliable and interpretable AI systems that enhance decision-making in clinical environments.

Workshop's Contributions

We received 24 submissions for the EXPLIMED workshop, and after a thorough review process, 18 were accepted for presentation. Researchers from 12 countries (Brazil, Germany, India, Ireland, Israel, Italy, Netherlands, Portugal, Spain, Taiwan, and Türkiye) attended the workshop, enriching the exchange of ideas and insights in Explainable Artificial Intelligence in medicine.

EXPLIMED featured several impactful presentations, opened by the keynote by Prof. Hani Hagra (University of Essex) on "True Explainable Artificial Intelligence for Health Applications." Prof. Hagra discussed how advances in computing power and the exponential growth of data have renewed interest in AI, yet the complexity of many AI algorithms creates opacity, often termed "black-box" models. He emphasized that for AI to earn trust and broad adoption, especially in healthcare, it must offer transparency through Explainable AI (XAI). XAI aims to build models that both explain individual decisions and help users understand a system's capabilities, predict behaviors, and identify improvements. Prof. Hagra highlighted XAI's transformative potential for healthcare, advocating for systems that are accessible, understandable, and supportive of human augmentation in decision-making.

The paper "Integrating Graph Neural Networks and Fuzzy Logic to Enhance Deep Learning Interpretability" introduced a methodology integrating Graph Neural Networks (GNNs) with Fuzzy Logic to enhance interpretability in deep learning models, demonstrating how this combination can address the complexities of structured data while maintaining transparency and reliability in AI systems. The paper "ProtoAL: Interpretable deep active learning with prototypes for medical imaging" focused on ProtoAL, presented a deep active learning model utilizing prototypes to foster interpretability in medical imaging, achieving commendable accuracy while reducing data requirements. In a significant advancement for patient privacy, the paper "Latent diffusion models for Privacy-preserving Medical Case-based Explanations" discussed using Latent Diffusion Models to create privacy-preserving, case-based explanations for medical diagnoses, effectively balancing visual quality with anonymity. Furthermore, the paper "Explaining Bayesian Networks in Natural Language using Factor Arguments. Evaluation in the medical domain" explored how Bayesian Networks can be explained in natural language through factor arguments, presenting a novel approach that aids users in understanding the reasoning process behind probabilistic inference in medical contexts. The paper "VISE: Validated and Invalidated Symbolic Explanations for Knowledge Graph Integrity" proposed a hybrid method that combines symbolic and numerical learning techniques for Knowledge Graphs, ensuring integrity and improving predictive performance while generating meaningful insights. Instead, the paper "Prediction of Continuous Targets by Explainable Imbalanced Regression from Omics Data in Childhood Obesity" addressed the challenge of imbalanced regression in predicting health metrics related to childhood obesity, employing explainable models that

improve prediction accuracy and elucidate meaningful biological relationships.

The paper "Explainable skin lesion classification with multitask learning" introduced a multi-task learning framework for skin lesion classification that utilizes optical coherence tomography to analyze cell nuclei and skin layers. This approach enhances model interpretability while accurately identifying skin conditions. Following this, the paper entitled "An Explainable Convolutional Neural Network for the Detection of Drug Abuse" study on drug abuse detection utilized a convolutional neural network (CNN) to analyze lateral-flow tests. The authors highlighted the model's ability to explain its predictions, providing crucial insights for real-world applications. Instead, the paper "Towards Explainable Federated Learning in Healthcare: A Focus on Heart Arrhythmia Detection" tackled the integration of explainable federated learning in detecting heart arrhythmias, employing a temporal convolutional network with attention mechanisms to ensure both privacy and interoperability. Another notable contribution entitled "Towards Explainable Deep Learning in Oncology: Integrating EfficientNet-B7 with XAI techniques for Acute Lymphoblastic Leukaemia" examined the explainability of electrocardiogram-based algorithms using Shapley attribution under a counterfactual reasoning setup, demonstrating the importance of understanding model predictions in cardiovascular diagnostics. A further study, called "Explainability by Shapley attribution for electrocardiogram-based algorithmic diagnosis under subtractive counterfactual reasoning setup", proposed a framework combining EfficientNet-B7 with various XAI techniques for diagnosing Acute Lymphoblastic Leukaemia, emphasizing the importance of explainability to enhance trust in AI-driven diagnostics. Instead, the paper "AI Readiness in Healthcare through Storytelling XAI" aims to tailor explanations to diverse audience needs, enhancing user trust and understanding of AI systems.

The paper "Mechanistic Causal Models for Explainable AI in Medicine: Coupling Respiratory and Immunological Systems for In Silico Medicine Simulations" proposed a novel approach utilizing mechanistic causal models that integrate known physiological principles to enhance understanding of complex medical conditions, mainly focusing on the dynamics of respiratory and immunological systems during cytokine storms. Instead, the paper "Identifying Candidates for Protein-Protein Interaction: A Focus on NKp46's Ligands" introduced a method for identifying candidates for protein-protein interactions (PPIs) using a deep learning model called DSCRIPT. This study emphasized the model's self-explanatory capabilities to streamline the screening process for potential interacting proteins. Furthermore, the paper "Evaluating Machine Learning Models against Clinical Protocols for Enhanced Interpretability and Continuity of Care" examined how machine learning models can be evaluated against established clinical protocols to enhance interpretability and continuity of care. The authors proposed metrics for comparing model predictions with clinical rules, ensuring that AI tools align with existing medical practices. In addition, the paper "Towards Explainable General Medication Planning" focused on a framework for explainable medication planning, which aimed to clarify the decision-making process in personalized medication administration by employing visualization techniques. The paper entitled "Reliable central nervous system tumor diagnosis on MRI images with Deep Neural Networks and Conformal Prediction" addressed the reliable differentiation of central nervous system tumors using deep neural networks coupled with conformal prediction, which provided not only accurate classifications but also quantifiable confidence measures. Finally, the paper "Explaining Predictions of Hypertension Disease through Anchors" discussed the use of the Anchors algorithm for explaining predictions in hypertension diagnosis, demonstrating

that optimal feature selection could enhance classification accuracy and explanation clarity.

Organizing Committee



Gianluca Zaza is an assistant professor at the University of Bari Aldo Moro and is a member of the Computational Intelligence Laboratory (CILab). He is working on "Understandability of AI systems" within the NRRP project "FAIR - Future Artificial Intelligence Research," Spoke 6 - Symbiotic AI. He is the project coordinator for the research project titled "Computational Models based on Fuzzy Logic for eXplainable Artificial Intelligence," which is funded for one year under the "Research Projects GNCS 2023" grant. He is a Guest Co-Editor of the Special Issue "Computational Intelligence in Healthcare" in *Bioengineering* (MDPI) and an Associate Editor for the *Journal of Intelligent & Fuzzy Systems* (IOS Press). He is a reviewer for several international journals published by leading publishers, including Elsevier and Springer.



Gabriella Casalino is currently an Assistant Professor (Tenure Track) at the Computational Intelligence Laboratory (CILab) of the Informatics Department of the University of Bari. Her research is focused on Computational Intelligence methods for interpretable data analysis. She is actively involved in eHealth, Data Stream Mining, and eXplainable Artificial Intelligence. Her work primarily focuses on the medical domain. She holds membership in the IEEE Task Force on Explainable Fuzzy Systems and the Interdepartmental Center for Telemedicine of the University of Bari-CITEL. She is an active member of the computer science community and contributes by organizing committees of workshops and special sessions in prestigious international conferences such as ECAI and IEEE WCCI. She is an Associate Editor for the international journals "IEEE Transactions on Computational Social Systems" and "Soft Computing". She is a Guest Editor for several special issues in IEEE SMC Magazine, IEEE Transactions on Computational Social Systems, and IEEE Systems Journal. She is a Senior member of the IEEE Society and has received several awards for her research, including the prestigious FUZZ-IEEE Best Paper award in 2022.



Giovanna Castellano is an Associate Professor at the Department of Computer Science, University of Bari Aldo Moro, where she coordinates the Computational Intelligence Laboratory (CILab). Her research interests are in the area of Computational Intelligence and Computer Vision. She has been responsible for the local unit of several research projects and is currently the Principal Investigator of the WP 6.4 "Understandability of AI systems" in the NRRP "FAIR - Future Artificial Intelligence Research" project, Spoke 6 - Symbiotic AI. She is an Associate Editor of several international journals. She has been a Guest Editor of special issues and participated in organizing scientific events. She is a reviewer for several international journals published by leading publishers, including Elsevier, IEEE, and Springer, and a member of the program committee of several international conferences. She is a member of the IEEE Society, EUSFLAT Society, INDAM-GNCS Society, IAPR Technical Committee 19

(Computer Vision for Cultural Heritage Applications), CINI-AIIS laboratory, CINI-BIG DATA laboratory, CITEL telemedicine research center, GRIN, MIR laboratories. She is also a member of the IEEE CIS Task Force on Explainable Fuzzy Systems.

Acknowledgments

The EXPLIMED organizers would like to thank the organizing committee of the 27th European Conference on Artificial Intelligence (ECAI 2024) for hosting this first edition of the workshop. The EXPLIMED workshop was organized with support from the CILAB (Computational Intelligence Lab) at the Department of Computer Science, University of Bari, and patronized by Fondazione FAIR through the PNRR project, FAIR - Future AI Research (PE00000013), Spoke 6 - Symbiotic AI (CUP H97G22000210007), under the NRRP MUR program funded by NextGenerationEU. The workshop organizers are members of the INdAM GNCS research group. Gabriella Casalino and Giovanna Castellano are members of the CITEL - Centro Interdipartimentale della ricerca in Telemedicina, of the University of Bari Aldo Moro.

References

- [1] A. B. Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. García, S. Gil-López, D. Molina, R. Benjamins, et al., Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai, *Information fusion* 58 (2020) 82–115.
- [2] A. S. Albahri, A. M. Duhaim, M. A. Fadhel, A. Alnoor, N. S. Baqer, L. Alzubaidi, O. S. Albahri, A. H. Alamoodi, J. Bai, A. Salhi, et al., A systematic review of trustworthy and explainable artificial intelligence in healthcare: Assessment of quality, bias risk, and data fusion, *Information Fusion* 96 (2023) 156–191.
- [3] H. W. Loh, C. P. Ooi, S. Seoni, P. D. Barua, F. Molinari, U. R. Acharya, Application of explainable artificial intelligence for healthcare: A systematic review of the last decade (2011–2022), *Computer Methods and Programs in Biomedicine* 226 (2022) 107161.
- [4] I. Stepin, M. Suffian, A. Catala, J. M. Alonso-Moral, How to build self-explaining fuzzy systems: from interpretability to explainability [ai-explained], *IEEE Computational Intelligence Magazine* 19 (2024) 81–82.
- [5] Y. Jia, J. McDermid, T. Lawton, I. Habli, The role of explainability in assuring safety of machine learning in healthcare, *IEEE Transactions on Emerging Topics in Computing* 10 (2022) 1746–1760.
- [6] S. Bharati, M. R. H. Mondal, P. Podder, A review on explainable artificial intelligence for healthcare: why, how, and when?, *IEEE Transactions on Artificial Intelligence* (2023).