Chapter 5
Explainable AI in Healthcare Application

Siva Raja Sindiramutty
Taylor's University, Malaysia

Wee Jing Tee
Taylor's University, Malaysia

Sumathi Balakrishnan
Taylor's University, Malaysia

Sukhminder Kaur
https://orcid.org/0000-0003-1616-3261
Taylor's University, Malaysia

Rajan Thangaveloo
https://orcid.org/0000-0002-0402-6019
Universiti Malaysia Sarawak, Malaysia

Husin Jazri
https://orcid.org/0009-0005-3940-2351
Taylor's University, Malaysia

Navid Ali Khan
Taylor's University, Malaysia

Abdalla Gharib
Zanzibar University, Tanzania

Amaranadha Reddy Manchuri
https://orcid.org/0000-0002-3873-0469
Kyungpook National University, South Korea

ABSTRACT

Given the inherent risks in medical decision-making, medical professionals carefully evaluate a patient's symptoms before arriving at a plausible diagnosis. For AI to be widely accepted and useful technology, it must replicate human judgment and interpretation abilities. XAI attempts to describe the data underlying the black-box approach of deep learning (DL), machine learning (ML), and natural language processing (NLP) that explain how judgments are made. This chapter provides a survey of the most recent XAI methods employed in medical imaging and related fields, categorizes and lists the types of XAI, and highlights the methods used to make medical imaging topics more interpretable. Additionally, it focuses on the challenging XAI issues in medical applications and guides the development of better deep-learning system explanations by applying XAI principles in the analysis of medical pictures and text.

DOI: 10.4018/978-1-6684-6361-1.ch005

INTRODUCTION

The healthcare sector is recognized as having significant potential for the implementation of artificial intelligence (AI), making it one of the largest sectors in this field. Clinical decision support systems (CDSSs) have been developed as productive systems over the last few decades, leading to a progressive transformation of the healthcare sector through AI solutions (Yu et al., 2018; M. Shafiq et al., 2021). Collaboration between doctors and computational scientists has been instrumental in enhancing healthcare facilities by implementing computerized processes that can increase the reliability of illness detection and prescribing appropriate therapies (Albahri et al., 2021). Notably, AI applications designed specifically for the medical field have assisted physicians in formulating diagnostic hypotheses, providing justifications for clinical decision-making, and choosing the best therapies (Rong et al., 2020; Priyadarshini et al., 2021). These technologies have had a significant impact on clinical settings worldwide, enabling the recognition and detection of health hazards. Furthermore, the FDA of the United States has approved the use of AI solutions in the healthcare sector to enhance healthcare services and reduce health hazards (Yu et al., 2018; Kok et al., 2020). Medical centers and hospitals have recently embraced novel innovations and advancements at a rapid pace, aiming to create intelligent things that are unmatched in the healthcare industry (Amann et al., 2020; Humayun et al., 2022).

The development of intelligent apps for medical diagnosis and therapy, taking ethical concerns like privacy, openness, safety, and accountability into account, is being actively supported by the FDA and the EU. In this context, medical innovations that demonstrate the adoption of AI in healthcare are of utmost importance and value and have been successfully applied in fields such as translational medical studies, basic biomedical studies, clinical practice, and medical image assessment devices (Hayden, 2014). The use of deep learning (DL) and machine learning (ML) techniques in current AI research in the healthcare industry has enabled the recognition of data patterns and comprehension of complicated relationships (Santamaría et al., 2011). Moreover, the model for AI in medical applications has been transformed by DL and ML techniques (Deo, 2015). As per the International Data Corporation, investment in AI systems will reach US$97.9 billion by 2023 (Markus et al., 2021; Lim et al., 2019). However, due to the transformation of decision norms and corporate ethics, platforms, and uses of AI aimed at the medical sector are vulnerable (Yu et al., 2018; Gouda et al., 2022). To mitigate this, an emerging field is aiming to provide insights into how explainable AI (XAI) algorithms work and how their decisions are made. In healthcare, XAI is being increasingly used to develop and deploy AI models that can support clinical decision-making and improve patient outcomes. Various resources, such as digital health records (EHRs), wearable technology, and medical imaging, are used to gather medical information. This data is then preprocessed to ensure it is clean and arranged in a format that AI systems can use. A variety of methods, including DL, ML, and natural language processing (NLP), are used to create AI systems, which are then trained using preprocessed medical information to discover trends and make forecasting. The performance measures of precision, recall, and accuracy are examined to evaluate the AI models. Explainability techniques enable human-AI interaction, allowing physicians and other medical professionals to better comprehend and trust the conclusions reached by the AI system. The AI algorithm's outputs support medical decision-making, which can be used to predict illness development, detect patients at greater risk, and suggest treatment strategies. Ultimately, XAI in healthcare aims to improve patient outcomes by enabling more precise and personalized medical decision-making (Reddy et al.,2019; Gunning & Aha, 2019; Rajkomar et al., 2018).

Related Content