

Explainability in Medical AI: Unpacking Deep Learning for Diagnostic Accuracy and Trust

Shekhar Katukoori¹ and Dr. Sandeep Chahal²

¹Research Scholar, Department of Computer Science and Engineering

²Associate Professor, Department of Computer Science and Engineering
NIILM University, Kaithal, Haryana, India

Abstract: *AI and Deep Learning (DL) may improve medical diagnosis across specializations. Due to the importance of healthcare choices, these models' complexity and opacity make clinical adoption difficult. This study develops interpretable deep learning models for medical diagnosis utilizing Explainable AI (XAI) methodologies to improve transparency, accountability, and confidence in AI-assisted medical decision-making. We study XAI methods, their use in diverse medical fields, and their effects on diagnostic accuracy and clinical interpretability. A thorough review of case studies shows how explainable models may sustain excellent diagnostic performance and give useful insights into their decision-making processes, possibly redefining the synergy between AI and human knowledge in healthcare.*

Keywords: Explainable AI; Deep Learning; Medical Diagnosis; Interpretability; Healthcare; Artificial Intelligence