Artificial Intelligence Review (2023) 56:11149–11296
https://doi.org/10.1007/s10462-023-10415-5

Trustworthy artificial intelligence in Alzheimer's disease: state of the art, opportunities, and challenges

Shaker El-Sappagh[1,2,3] · Jose M. Alonso-Moral[4] · Tamer Abuhmed[3] · Farman Ali[5] · Alberto Bugarín-Diz[4]

Accepted: 31 January 2023 / Published online: 12 March 2023

Abstract
Medical applications of Artificial Intelligence (AI) have consistently shown remarkable performance in providing medical professionals and patients with support for complex tasks. Nevertheless, the use of these applications in sensitive clinical domains where high-stakes decisions are involved could be much more extensive if patients, medical professionals, and regulators were provided with mechanisms for trusting the results provided by AI systems. A key issue for achieving this is endowing AI systems with key dimensions of Trustworthy AI (TAI), such as fairness, transparency, robustness, or accountability, which are not usually considered within this context in a generalized and systematic manner. This paper reviews the recent advances in the TAI domain, including TAI standards and guidelines. We propose several requirements to be addressed in the design, development, and deployment of TAI systems and present a novel machine learning pipeline that contains TAI requirements as embedded components. Moreover, as an example of how current AI systems in medicine consider the TAI perspective, the study extensively reviews the recent literature (2017–2021) on AI systems in a prevalent and high social-impact disease: diagnosis and progression detection of Alzheimer's Disease (AD). The most relevant AI systems in the AD domain are compared and discussed (such as machine learning, deep learning, ensembles, time series, and multimodal multitask) from the perspective of how they address TAI in their design. Several open challenges are highlighted, which could be claimed as one of the main reasons to justify the rare application of AI systems in real clinical environments. The study provides a roadmap to measure the TAI status of an AI systems and highlights its limitations. In addition, it provides the main guidelines to overcome these limitations and build medically trusted AI-based applications in the medical domain.

Keywords Trustworthy AI · AI for Alzheimer's disease diagnosis and progression detection · Machine learning in medicine · Responsible AI · Fairness, accountability, and transparency in AI

✉ Tamer Abuhmed
tamer@skku.edu

Extended author information available on the last page of the article