RESEARCH PAPER

Decision support for efficient XAI services - A morphological analysis, business model archetypes, and a decision tree

Jana Gerlach[1] · Paul Hoppe[1] · Sarah Jagels[1] · Luisa Licker[1] · Michael H. Breitner[1]

Abstract

The black-box nature of Artificial Intelligence (AI) models and their associated explainability limitations create a major adoption barrier. Explainable Artificial Intelligence (XAI) aims to make AI models more transparent to address this challenge. Researchers and practitioners apply XAI services to explore relationships in data, improve AI methods, justify AI decisions, and control AI technologies with the goals to improve knowledge about AI and address user needs. The market volume of XAI services has grown significantly. As a result, trustworthiness, reliability, transferability, fairness, and accessibility are required capabilities of XAI for a range of relevant stakeholders, including managers, regulators, users of XAI models, developers, and consumers. We contribute to theory and practice by deducing XAI archetypes and developing a user-centric decision support framework to identify the XAI services most suitable for the requirements of relevant stakeholders. Our decision tree is founded on a literature-based morphological box and a classification of real-world XAI services. Finally, we discussed archetypical business models of XAI services and exemplary use cases.

Motivation and research needs

Artificial Intelligence (AI) has potentially far-reaching applications that can influence people's private and professional lives (Meske et al., 2022). These include the identification of diseases (Aignostics[1]; Meske et al., 2022), job recruitment (iVCV[2]; Sipior et al., 2021), public security (Intelligent Artifact[3]), and risk assessment when granting loans (Wang et al., 2019; ZEST AI[4]). The models used in these instances are often highly complex black boxes (Adadi & Berrada, 2018), meaning that the ability to understand the models' underlying AI processes—and thus the reasons for their decisions—is severely limited. This is problematic because the comprehensibility, explainability, and justification of decisions are of great importance for many applications, including in the health, finance, and energy sectors (Meske et al., 2022).

Although AI is already used for a wide range of activities and provides various benefits, many decision-makers such as managers and executive board members, remain reluctant to integrate AI technologies caused by a limited understanding (Barredo Arrieta et al., 2020). This issue can be addressed by Explainable Artificial Intelligence (XAI) methods, which emphasize the need to make complex models and algorithms understandable and reproducible to humans (Meske et al., 2022). According to Gilpin et al. (2018), the term "explainability" refers to "models that are able to summarize the reasons for neural network behavior, gain the trust of users, or produce insights about the causes of their decisions" (p.80). Moreover, it is not sufficient to gain the trust and

✉ Jana Gerlach
  gerlach@iwi.uni-hannover.de

[1] Institut für Wirtschaftsinformatik, Leibniz Universität Hannover, Königsworther Platz 1, 30167 Hannover, Germany

[1] http://www.aignostics.com.

[2] https://ivcv.eu/.

[3] https://www.intelligent-artifacts.com/

[4] https://zest.ai/.