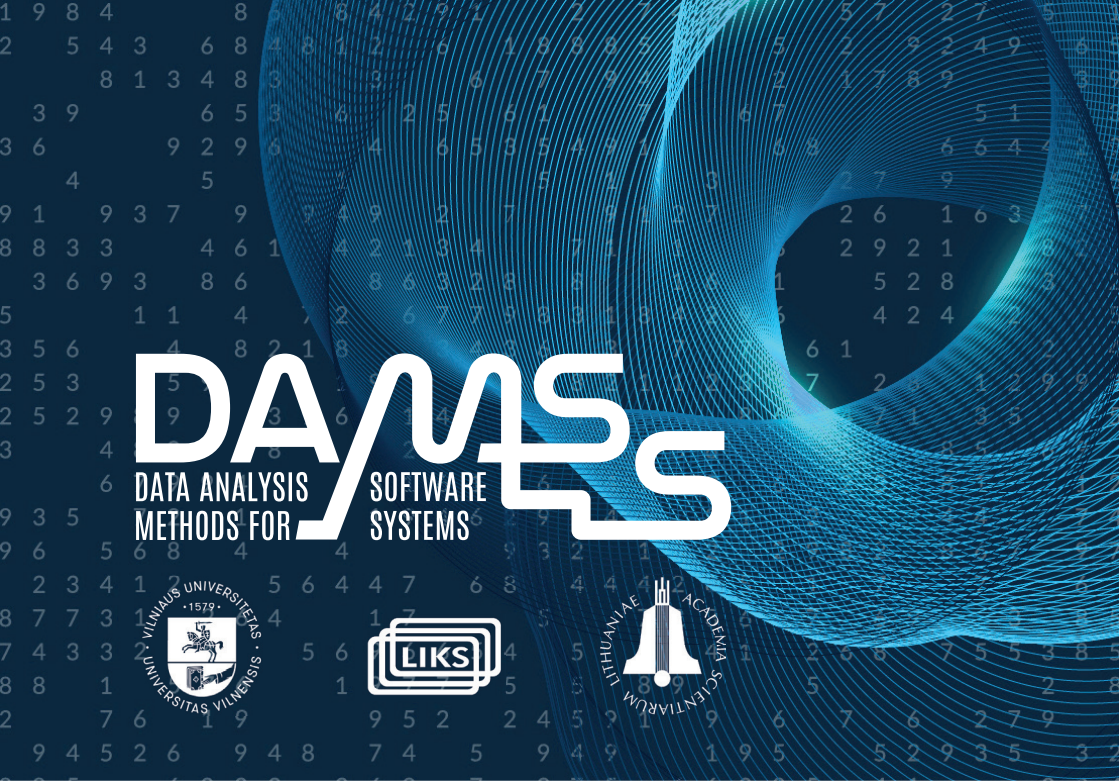

15th Conference on

DATA ANALYSIS METHODS
for Software Systems

November 28–30, 2024

Druskininkai, Lithuania, Hotel "Europa Royale"
https://www.mii.lt/DAMSS

Contacts:

Dr. Jolita Bernatavičienė
jolita.bernataviciene@mif.vu.lt
Tel. (+370 5) 2109 315
Prof. Olga Kurasova
olga.kurasova@mif.vu.lt

https://doi.org/10.15388/DAMSS.15.2024
ISBN 978-609-07-1112-5 (digital PDF)

Transforming Black-Box Models into Explainable AI for Breast Cancer Recognition

Sobia Dastgeer, Povilas Treigys

Institute of Data Science and Digital Technologies
Vilnius University

Sobia.dastgeer@mif.stud.vu.lt

Artificial intelligence (AI) with deep learning (DL) models using medical imaging has sparked a revolutionary wave in healthcare in recent years. DL models often function as black boxes, making their decision-making process difficult to interpret. These considerations get more significant when people's health is at risk, as in the involvement of AI applications used in healthcare. The situation is further worsened by the fact that, unlike conventional machine learning algorithms, state-of-the-art deep learning algorithms consist of complex connected structures, millions of parameters, and a black box mentality that provides no insight into how they operate inside. A recent field of study in machine learning called Explainable Artificial Intelligence (XAI) aims to understand how AI systems make judgement that are hidden from visibility. The idea of explainable AI (XAI) emerged as a response to the issue of AI black boxes, with the goal of offering explanations that are transparent, believable, and sensitive to human understanding. For improved comprehension, a case study regarding the function of XAI in the identification of Breast cancer (BC) with their performances are discussed in this study. BC is a prevalent and extremely deadly disease. It is currently the second largest cause of cancer-related mortality among females globally. This study focuses on the integration of several XAI techniques like Grad-CAM, LIME, SHAP, PDPs and other deep learning models used for breast cancer classification and detection and how these explainability strategies are the better way to establish trustworthiness when utilizing AI systems in the healthcare industry. We will examine several key ideas in XAI, outline various challenges related to XAI in the healthcare industry, and explore

whether XAI can actually progress the field by fostering greater understanding and trust among other things.

Acknowledgements: This research is funded under the Programme "University Excellence Initiatives" of the Ministry of Education, Science and Sports of the Republic of Lithuania (Measure No. 12-001-01-01-01 "Improving the Research and Study Environment").