

LITHUANIAN COMPUTER SOCIETY

VILNIUS UNIVERSITY, INSTITUTE OF DATA SCIENCE AND DIGITAL TECHNOLOGIES

LITHUANIAN ACADEMY OF SCIENCES



16th Conference on

DATA ANALYSIS METHODS for Software Systems

November 27–29, 2025

Druskininkai, Lithuania, Hotel “Europa Royale”

<https://www.mii.lt/DAMSS>

VILNIUS UNIVERSITY PRESS

Vilnius, 2025

Co-Chairs:

Dr. Saulius Maskeliūnas (Lithuanian Computer Society)

Prof. Gintautas Dzemyda (Vilnius University, Lithuanian Academy of Sciences)

Programme Committee:

Dr. Jolita Bernatavičienė (Lithuania)

Prof. Juris Borzovs (Latvia)

Prof. Janusz Kacprzyk (Poland)

Prof. Ignacy Kaliszewski (Poland)

Prof. Bożena Kostek (Poland)

Prof. Tomas Krilavičius (Lithuania)

Prof. Olga Kurasova (Lithuania)

Assoc. Prof. Tatiana Tchemisova (Portugal)

Assoc. Prof. Gintautas Tamulevičius (Lithuania)

Prof. Julius Žilinskas (Lithuania)

Organizing Committee:

Dr. Jolita Bernatavičienė

Prof. Olga Kurasova

Assoc. Prof. Viktor Medvedev

Laima Paliulionienė

Assoc. Prof. Martynas Sabaliauskas

Prof. Povilas Treigys

Contacts:

Dr. Jolita Bernatavičienė

jolita.bernatavicienne@mif.vu.lt

Prof. Olga Kurasova

olga.kurasova@mif.vu.lt

Tel. (+370 5) 2109 315

Copyright © 2025 Authors. Published by Vilnius University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

<https://doi.org/10.15388/DAMSS.16.2025>

ISBN 978-609-07-1200-9 (digital PDF)

© Vilnius University, 2025

Evaluation Metrics for Explainable AI

Rokas Gipiškis, Olga Kurasova

Institute of Data Science and Digital Technologies

Vilnius University

rokas.gipiskis@mif.vu.lt

Explainable Artificial Intelligence (XAI) has become one of the key components in ensuring transparency and accountability in AI systems. However, the evaluation of XAI methods remains inconsistent and fragmented. Here, we analyse the categorisations, methodologies, and properties of various evaluation metrics to highlight key differences in evaluation approaches and the lack of consensus across studies. Our findings emphasise challenges such as inconsistent definitions, limited implementation, and the absence of standardised evaluation frameworks. By comparing existing taxonomies and metric classifications, we identify gaps and provide insights to support the development of more robust and comparable XAI evaluation methodologies. Finally, we highlight underexplored application areas beyond classification, such as semantic segmentation, as promising directions for evaluation research.