IAF SYMPOSIUM ON INTEGRATED APPLICATIONS (B5)
Tools and Technology in Support of Integrated Applications (1)

Author: Ms. Maren Hülsmann
Universität der Bundeswehr München, Germany, maren.huelsmann@unibw.de

Prof.Dr. Roger Förstner
Universität der Bundeswehr München, Germany, roger.foerstner@unibw.de

EXPLAINABLE AI IN SPACE – A REVIEW

Abstract

Artificial Intelligence (AI), and in particular Machine Learning (ML) and Deep Learning (DL), has gained increasing attention in the space domain. However, there is some reluctance to deploy AI-based space systems beyond research purposes. Users and developers alike state that especially DL models especially behave like black boxes, are too complex, and their decision-making is not understandable. As a result, AI-based space systems are often considered to not be reliable enough for routine operations.

In this context, the relatively new framework of Explainable Artificial Intelligence (XAI) and its methods for post-hoc explanation have the potential to provide a deeper understanding about the inner workings of DL methods. These methods aim to visualize and extract the feature relevance of parameters using only the results of the DL-model in question. This framework is of interest for AI-based space systems on ground and in orbit to finally deploy ML and DL beyond the research domain and thus further increase the level of autonomy.

This study provides an overview of the XAI paradigm and its potential applications for AI in space. A review of the current state of XAI in space is presented and the study concludes with a mapping of XAI methods and space-related use cases for Earth Observation, Spacecraft Operations, and beyond.