



SPECIAL SESSION ON

EXPLAINABLE ARTIFICIAL INTELLIGENCE FOR UNVEILING THE BRAIN: FROM THE BLACK-BOX TO THE GLASS-BOX (XAIB)

ORGANIZER

Dr. Alessia SARICA, Assistant Professor, Neuroscience Research Center, Magna Graecia University of Catanzaro, Italy (sarica@unicz.it).

[Alessia SARICA](#) is Assistant Professor at the Neuroscience Research Center, Magna Graecia University of Catanzaro, Italy. She received the Ph.D. degree in Biomedical and Computer Science Engineering in 2015, and her main research field is Machine Learning on Neuroimaging and Neuropsychological data for the early and differential diagnosis of neurodegenerative diseases. In particular, she is an expert in decision trees, ensemble learning, Random Forests, and feature selection, which she applied for the prediction of Alzheimer's disease, Parkinsonisms and Motor Neuron Diseases. In 2018, She was the organizer and Guest Editor of a Special Issue hosted by the Journal of Neuroscience Methods, titled "[A Machine learning neuroimaging challenge for automated diagnosis of Mild Cognitive Impairment](#)" (10 contributions) and more recently, She was the Guest editor of another Special Issue "[Machine Learning in Healthcare and Biomedical Application](#)" hosted by the Journal Algorithms MDPI (5 contributions).

CALL FOR PAPERS AND ABSTRACTS

Nowadays, Artificial Intelligence (AI) and Machine Learning (ML) are widely used for the exploration of the Brain and their application ranges from the processing and analysis of neuroimages to the automatic diagnosis of the neurodegenerative diseases. However, without an explanation of the ML findings, the automatic medical and clinical decisions are still hard to be trusted. Indeed, the *black-box* nature of most algorithms, although providing high accuracy, makes the interpretation of the predictions hard. Indeed, in recent years the need of interpretable and explainable AI, especially in Healthcare, got stronger, as well as the need of *glass-box* models able to show a trade-off between intelligibility and optimal performance.

The aim of this Special Session is to collect scientific works devoted to the new challenge of Explainable Artificial Intelligence applied on Neuroscience, Neuroimaging and Neuropsychological data for unveiling the Brain. Researchers are encouraged to submit high quality papers or abstracts on novel or state-of-the-art intelligible, interpretable, and understandable AI approaches, such as post-hoc explainability techniques both model-agnostic (e.g., *lime*, *shap*) and model-specific (e.g., CNN, SVM, Random Forests), and transparent models (i.e., linear/logistic regression, decision trees, GAM), with special attention to *global* and *local* explanations. Systematic reviews and meta-analyses are also welcome.

POTENTIAL PROGRAM COMMITTEE MEMBERS AND INVITED SPEAKERS

- Javier Del Ser, "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI" (javier.delser@tecnalia.com).
- Alex John London, "Artificial Intelligence and Black-Box Medical Decisions: Accuracy versus Explainability" (ajlondon@andrew.cmu.edu).

- Muhammad Aurangzeb Ahmad, “**Interpretable Machine Learning in Healthcare**” (muhammad@kensci.com).
- Rich Caruana, “**Intelligible Models for HealthCare: Predicting Pneumonia Risk and Hospital 30-day Readmission**” (rcaruana@microsoft.com).
- Michele Ferrante, “**Explainable Artificial Intelligence for Neuroscience: Behavioral Neurostimulation**” (ferrantem@nih.gov).
- Hamed Asadi, “**Peering Into the Black Box of Artificial Intelligence: Evaluation Metrics of Machine Learning Methods**” (hamed.asadi@monash.edu).
- J. Eric T. Taylor, “**Artificial cognition: How experimental psychology can help generate explainable artificial intelligence**” (j.eric.t.taylor@gmail.com).