Preface

This part of the *Journal of Automation, Mobile Robotics and Intelligent Systems* is devoted to current studies in Computer Science and Information Technology presented by young, talented contributors working in the field – it is the fourth edition of this series. Among the included papers, one can find contributions dealing with diagnosing machine learning problems, natural language processing procedures, AI classification and clustering methods, optimization tasks, and learning procedures. This part of JAMRIS was inspired by broad and interesting discussions during the Eight Doctoral Symposium on Recent Advances in Information Technology (DS-RAIT 2023), held in Warsaw, Poland, on September 17-20, 2023, as a satellite event of the Federated Conference on Computer Science and Information Systems (FedCSIS 2023). The Symposium facilitated the exchange of ideas between early-stage researchers, particularly PhD students, in Computer Science. Furthermore, the Symposium gave all participants an opportunity to obtain feedback on their ideas and explorations from the experienced members of the IT research community who had been invited to chair all DS-RAIT thematic sessions. Therefore, submitting research proposals with limited preliminary results was strongly encouraged.

Here, we highlight the contributions entitled *"Mitigating the effects of non-IID data in federated learning with a self-adversarial balancing method,"* written by Anastasiya Danilenka (Warsaw University of Technology). This paper received the Best Paper Award at DS-RAIT 2023.

This issue contains the following DS-RAIT papers in their special, extended versions.

The first paper, entitled *"Tackling Non-IID Data and Data Poisoning in Federated Learning Using Adversarial Synthetic Data,"* authored by Anastasiya Danilenka, explores crucial aspects of federated learning (FL). FL involves collaborative model training across diverse devices while safeguarding data privacy. However, managing heterogeneous data across these devices poses a significant challenge, exacerbated by the potential presence of malicious clients aiming to disrupt the training process through data poisoning. The article addresses the issue of discerning between poisoned and non-Independently and Identically Distributed (non-IID) data by proposing a technique that leverages data-free synthetic data generation, employing a reverse adversarial attack concept. This approach enhances the training process by assessing client coherence and favouring trustworthy participants. The experimental findings garnered from image classification tasks on MNIST, EMNIST, and CIFAR-10 datasets are meticulously documented and analysed, shedding light on the efficacy of the proposed method. As already mentioned, the DS-RAIT Program Committee voted this work the Best Paper of the event because of its excellent presentation of applicational aspects and promising results.

The paper entitled *"Gradient Scale Monitoring for Federated Learning Systems"* was written by Karolina Bogacka, Anastasiya Danilenka, and Katarzyna Wasielewska-Michniewska. In this paper, the authors delve into the burgeoning realm of Federated Learning amidst edge and IoT devices' expanding computational and communicational capabilities. While FL holds promise, particularly in cross-device settings, existing research often needs to be more focused on critical operationalisation and monitoring challenges. Through a case study comparing four FL system topologies, the paper uncovers periodic accuracy drops and attributes them to exploding gradients. Proposing a novel method reliant on the local computation of the gradient scale coefficient (GSC) for continuous monitoring, the study expands to explore GSC and average gradients per layer as potential diagnostic metrics for FL. By simulating various gradient scenarios, including exploding, vanishing, and stable gradients, the paper evaluates resulting visualizations for clarity and computational demands, culminating in introducing a gradient monitoring suite for FL training processes.

In their study titled *"Efficiency of Artificial Intelligence Methods for Hearing Loss Type Classification: An Evaluation,"* Michał Kassjański, Marcin Kulawiak, Tomasz Przewoźny, Dmitry Tretiakow, Jagoda Kuryłowicz, Andrzej Molisz, Krzysztof Koźmiński, Aleksandra Kwaśniewska, Paulina Mierzwińska-Dolny, and Miłosz Grono address critical issues surrounding the evaluation of hearing loss. Traditionally, hearing loss assessment relies on pure tone audiometry testing, considered the gold standard for evaluating auditory function. Once hearing loss is identified, distinguishing between sensorineural, conductive, and mixed types becomes paramount. The study compares various AI classification models using 4007 pure-tone audiometry samples meticulously labelled by professional audiologists. Models tested

range from Logistic Regression to sophisticated architectures like Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU). Furthermore, the study explores the impact of dataset augmentation using Conditional Generative Adversarial Networks and different standardisation techniques on the performance of machine learning algorithms. Remarkably, the RNN model emerges with the highest classification performance, achieving an out-of-training accuracy of 94.4% as determined by 10-fold Cross-Validation.

Finally, Marcin Sowanski, Jakub Hoscilowicz, and Artur Janicki contributed a paper titled *"Analysis of Dataset Limitations in Semantic Knowledge-Driven Multi-Variant Machine Translation."* This research explores the intricacies of dataset constraints within semantic knowledge-driven machine translation, tailored explicitly for intelligent virtual assistants (IVA). Departing from conventional translation methodologies, the study adopts a multi-variant approach to machine translation. Instead of relying on single-best translations, their method employs a constrained beam search technique to generate multiple viable translations for each input sentence. The methodology's expansion is noteworthy beyond the constraints of specific verb ontologies, operating within a broader semantic knowledge framework. This enables a more nuanced interpretation of linguistic nuances and contextual intricacies, thereby enhancing translation accuracy and relevance within the IVA domain.

We want to thank all those who participated in and contributed to the Symposium program and all the authors who submitted their papers. We also wish to thank all our colleagues and the members of the Program Committee for their hard work during the review process, their cordiality, and the outstanding local organisation of the Conference.

Editors:
Piotr A. Kowalski
Systems Research Institute, Polish Academy of Sciences
and Faculty of Physics and Applied Computer Science,
AGH University of Science and Technology

Szymon Łukasik
Systems Research Institute, Polish Academy of Sciences
and Faculty of Physics and Applied Computer Science,
AGH University of Science and Technology