



## **MACHINE LEARNING REGRESSION MODELS' ANALYSIS: PIEZOMETRIC WATER LEVEL PREDICTION - CASE STUDY**

**Vukasin Ćirović<sup>1</sup>, Vesna Ranković<sup>2</sup>, Nikola Milivojević<sup>1</sup>**

<sup>1</sup> Jaroslav Cerni Water Institute, Jaroslava Cernog 80, 11226 Pinosava, Belgrade, Serbia  
e-mail: vukasin.cirovic@jcerni.rs, nikola.milivojevic@jcerni.rs

<sup>2</sup> Faculty of Engineering, University of Kragujevac, Sestre Janjić 6, 34000 Kragujevac, Serbia  
e-mail: vesnar@kg.ac.rs

### **Abstract:**

Recent development of artificial intelligence, machine learning and deep learning, in particular, resulted in the increase in the use of data-based models in various fields; among others, in the field of dam safety. Neural networks are the most frequently used machine learning technique which has been applied to various problems. Other machine learning techniques are used for the analysis and interpretation of dam structural behaviour. In this paper, an analysis is conducted exhibiting how novel machine learning techniques can be used for piezometric water level prediction. Results from different techniques are presented and discussed. At the same time, the performance of the previously developed neural network model is analysed with the extended dataset, since additional measurements have been collected in the meantime. Although only one representative piezometer is considered, the proposed methodology may be generally applicable. Finally, some recommendations are given on how predictive models that are very similar at first glance may differ by additional analyses.

**Keywords:** neural networks, machine learning, deep learning, dam safety

### **1. Introduction**

In dam safety systems, the response of behaviour models is of great importance for daily operation as well as long-term evaluation. Finite element-based models (FEM) are widely used in dam safety analysis because of its physically based nature, transparency, and interpretation of results. The main disadvantage of those models, in daily operation, is its slowness. Additionally, as dam monitoring system develops further (measurement automatization, new sites, etc.), there is a growing influx of data that requires the adaptation of current model. On the other hand, some problems in dam safety analysis are of local character (seepage, local pressure increase) [1] and are often quite difficult to model. Statistical models [2] in analysis allow the creation of predictive models using large amount of available data. But, the tendency to extract as much information as possible from data related to dam safety sets up the limits to statistical models [3]. However, development in the field of machine learning (ML) in recent years has enabled application of data-based modes in various fields such as medicine, e-commerce, business intelligence, and dam safety as well. Some authors are focused on dam behaviour [4-5], prediction of displacement [6-7], shape optimization [8], crack detection [9], flow prediction [10] or piezometric water level prediction [11].

The main objective of this study is to analyse certain novel machine learning regression models in the context of applicability to the piezometric water level prediction. Additionally, the usability of the existing model presented in [11] is checked particularly in the domain of accuracy, since after almost a decade, new data have been collected. The results of the

comparative analysis of predictions of water level obtained from different models are also presented.

## 2. Dataset

Dataset consists of the piezometric water level acquired in the period from 1999 to 2020 from piezometers labelled as FP-13A, located on the non-overflow dam of the Iron Gate II and downstream water level measured in the same period. The total amount of data per piezometer is about 490 which 3 times more than the dataset used in the related study. To compare results with those in [11], the same record of data per piezometer is used. Single record contains piezometer water level (current day), and 3 downstream water levels labeled as  $h_t$ ,  $h_{t-1}$  and  $h_{t-2}$  which refer to the current day, day before and two days before, respectively. Some basic statistics of the data sets are presented in Table 1.

Variable	Min		Max		Average		Std		Cor <sub>FP-13</sub>		Cor <sub>FP-29</sub>	
	Now	Prev.	Now	Prev.	Now	Prev.	Now	Prev.	Now	Prev.	Now	Prev.
$h_t$	28.6	28.6	38.25	37.16	32	32.16	1.8498	1.669	0.9564	0.971	0.9642	0.9715
$h_{t-1}$	28.71	28.71	38.21	36.93	31.98	32.12	1.852	1.6603	0.9522	0.9611	0.9628	0.9723
$h_{t-2}$	28.79	28.87	38.23	37	31.99	32.13	1.8563	1.663	0.938	0.971	0.9534	0.9634
FP-13	29.8	29.96	38.38	37.26	32.29	32.39	1.7108	1.5077	1	1	-	-

**Table 1.** Basic statistics of datasets

## 3. Machine learning models

In recent years, there were a lot of examples of using machine learning models in the field of dam safety. Some of them used Support Vector Machine (SVM) [12-13], while others used Gaussian Process Regressors (GPR) [14-15]. The main goal of this paper is to develop, check performance, and apply to water level prediction problem the regression models belonging to SVM, GPR and Regression Trees as well as deep learning Long-ShortTermMemory (LSTM) model. To achieve this goal, two software packages are used. The first one is Matlab and its Regression Learner application, while the second one is ML.NET library. While Matlab is well-known software package in scientific community with plenty of regression models through Regression Learner application, ML.NET is a relatively new, open source, machine learning library developed by Microsoft for C# language [16]. The entire library contains a lot of models not only for regression, but also for classification and clusterization. To our knowledge, there are no papers in literature related to implementation of this library in dam safety analysis. The total number of selected models from both software packages is 23: 15 from Matlab (SVM:6, Regression Tree:3, Ensemble:2, GPR:4) and 8 from ML.NET. Simple LSTM model is generated in Matlab.

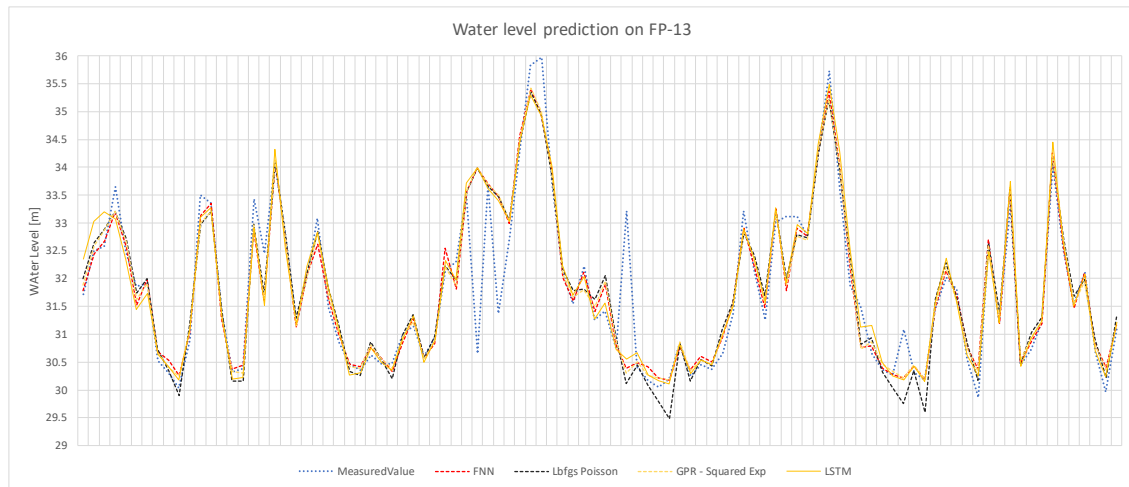
Dataset is divided into train and test data by the ratio 80:20 which means 80 percent of data is used for training, and 20 percent of data is used for testing models. In order to compare the results from different models, standard metrics is used: correlation coefficient  $r$ , *R-squared* ( $r^2$ ), *Mean Absolute Error (MAE)* and *Mean Square Error (MSE)*. The best models according to proposed metrics are presented in Table 2. The first row is the best Matlab model, the second row is the best ML.NET model, the third one is LSTM, and the fourth row contains results obtained by FNN model proposed in [11].

Piezometer FP-13	r		R <sup>2</sup>		MAE		MSE	
	Training	Test	Training	Test	Training	Test	Training	Test
GPR - Squared Exponential	0.97	0.9	0.94	0.81	0.2837	0.2911	0.1865	0.3344
LbfgsPoisson	0.97	0.9	0.94	0.82	0.31	0.33	0.19	0.37
LSTM	0.98	0.9	0.96	0.82	0.2543	0.3138	0.1188	0.3351
FNN	0.99	0.95	0.97	0.91	0.285	0.31	0.18	0.34

**Table 2.** The best models for FP-13

#### 4. Results and discussion

Graphic representation of results obtained using different ML models for piezometer FP-13 is shown in Figure 1.



**Figure 1.** Prediction of water level on FP-13

Briefly, model results are almost identical or very similar. All of them have similar dynamics, none of them reaches some measured peaks. There are no significant deviations in the model response. Considering numerical results shown in Table 2, the best model is GPR – Squared Exponential regression model according to MAE and MSE for test case. Generated LSTM model achieved the best performance within training data, but the results with test data are slightly worse. Lbfgs Poisson model of FP-13 has pronounced spikes i.e. it is sensitive to sharp changes. Other regression models do not have such characteristics.

In order to distinguish almost similar regression models, additional experiment has been made. Absolute error values obtained from the testing period have been divided into 3 groups and have been counted. The results are shown in Table 3.

Piezometer FP-13	< 0.15	[0.15,0.5]	> 0.5
FNN	55	33	10
GPR - Squared Exponential	36	49	13
LbfgsPoisson	47	41	10
LSTM	45	38	15

**Table 3.** The best models for FP-13

Most errors for FNN are below 0.15, while for GPR is in between 0.15 and 0.5 but, according to metric parameter MAE, GPR is slightly better than FNN.

#### 5. Conclusion

Neural networks are certainly the most used ML technique in dam safety analysis while other techniques are less common. In this paper, it is shown that all ML models presented are suitable for prediction of piezometric water level. Moreover, deep learning model LSTM had the best performance on the training dataset, but was not so beneficial with the test dataset which requires additional tuning process. Although only one piezometer is considered in this study, the presented methodology could be used for all piezometers of non-overflow dam and their number is significant.

Like any other tool, ML must be used by specialists with a broad knowledge of how it works.

But, there is still the question of a suitable technique to be used for some problem not only in the sense of accuracy but also in the sense of interpretability. For example, it is known

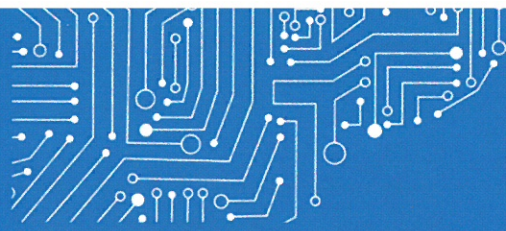
that SVM is more interpretable than NN because of the kernel function. Of course, we should strive to develop and use as simple models as possible taking care not to reduce the quality of the results obtained. Because of that, it is strongly recommended to use more than one predictive model and compare the results obtained.

## References

- [1] G. Lombardi (2004). Advanced data interpretation for diagnosis of concrete dams. Technical report, CISM.
- [2] Chouinard Luc, Bennett David and Feknous Nadia. (1995). Statistical Analysis of Monitoring Data for Concrete Arch Dams. *Journal of Performance of Constructed Facilities - J PERFORM CONSTR FACIL.* 9. 10.1061/(ASCE)0887-3828(1995)9:4(286).
- [3] F. Salazar, M. A. Toledo, E. Onate, R. Moran Data-Based Models for the Prediction of Dam Behaviour: A Review and Some Methodological Considerations, *Journal: Archives of Computational Methods in Engineering*, : 2015, ISSN: 1886-1784
- [4] Belmokre, A., Mihoubi, M.K. & Santillán, D. Analysis of Dam Behavior by Statistical Models: Application of the Random Forest Approach. *KSCE J Civ Eng* 23, 4800–4811 (2019).
- [5] B. Stojanovic, M. Milivojevic, M. Ivanovic, N. Milivojevic, D. Divac (2013). Adaptive system for dam behavior modeling based on linear regression and genetic algorithms. *Advances in Engineering Software*, 65:182190.
- [6] Mata J (2011) Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models. *Eng Struct* 33(3):903–910
- [7] Rankovic V, Grujovic N, Divac D, Milivojevic N, Novakovic A (2012) Modelling of dam behaviour based on neuro-fuzzy identification. *Eng Struct* 35:107–113
- [8] Gholizadeh S, Seyedpoor SM (2011) Shape optimization of arch dams by metaheuristics and neural networks for frequency constraints. *Sci Iran* 18(5):1020–1027
- [9] Wang BS, He ZC (2007) Crack detection of arch dam using statistical neural network based on the reductions of natural frequencies. *J Sound Vib* 302(4–5):1037–1047
- [10] Tayfur G, Swiatek D, Wita A, Singh VP (2005). Case study: finite element method and artificial neural network models for flow through JeziorskoEarthfill Dam in Poland. *J HydraulEng* 131(6):431–440
- [11] V. Ranković, A. Novaković, N. Grujović, D. Divac, N. Milivojević *Predicting piezometric water level in dams via artificial neural networks*, *Neural Computing and Applications*, 24(5), 1115-1121, 2014.
- [12] V. Ranković, N. Grujović, D. Divac, N. Milivojević (2014). Development of support vector regression identification model for prediction of dam structural behaviour. *Structural Safety*, 48:33–39.
- [13] Su Huaizhi, Chen Zhexin, Wen Zhiping. (2015). Performance improvement method of support vector machine-based model monitoring dam safety. *Structural Control and Health Monitoring*. 23. 10.1002/stc.1767.
- [14] Kang, Fei and Li, Junjie. (2020). Displacement Model for Concrete Dam Safety Monitoring via Gaussian Process Regression Considering Extreme Air Temperature. *Journal of Structural Engineering*. 146. 05019001. 10.1061/(ASCE)ST.1943-541X.0002467.
- [15] Lin Chaoning, Li Tongchun, Chen Siyu, Liu Xiaoqing, Lin Chuan Liang Siling. (2019). Gaussian process regression-based forecasting model of dam deformation. *Neural Computing and Applications*. 31. 10.1007/s00521-019-04375-7.

- [16] Ahmed, Zeeshan et al. (2019). Machine Learning at Microsoft with ML.NET. 2448-2458. 10.1145/3292500.3330667.





First Serbian International  
Conference on

**APPLIED**  
**ARTIFICIAL**  
**INTELLIGENCE**

# Book of Abstracts



May 19-20, 2022, Kragujevac, Serbia



**The First Serbian International Conference on Applied Artificial Intelligence,  
Kragujevac – Book of abstracts**

**Editor**

Professor Nenad Filipović

**Technical Editor**

Đorđe Dimitrijević

**Proofreaders**

Neda Vidanović Miletić

Aleksandra Stojanović

Nikola Šipić

Milena Đorđević

**Publisher**

University of Kragujevac, Serbia

**Press**

GRAFTEK doo Beograd

**Impression**

100 copies

**Year of publication**

2022

CIP - Каталогизација у публикацији  
Народна библиотека Србије, Београд

004.8(048)

**SERBIAN International Conference on Applied Artificial Intelligence (1 ; 2022 ;  
Kragujevac)**

Book of abstracts / First Serbian International Conference on Applied Artificial  
Intelligence, AAI, May19-20, 2022, Kragujevac, Serbia ; [organizers University of Kragujevac  
[and] Serbian Academy of Sciences and Arts] ; [editor Nenad Filipović]. - Kragujevac :  
University, 2022 (Beograd : Graftek). - 116 str. ; 30 cm

Tiraž 100. - Str. 4-5: Welcome message / Nenad Filipović.

ISBN 978-86-81037-71-3

a) Вештачка интелигенција -- Апстракти б) Машинско учење -- Апстракти

COBISS.SR-ID 65855753

## Organizers

- **University of Kragujevac**
- **Serbian Academy of Sciences and Arts**



with the support of

- **Ministry of Education, Science and Technological Development of the Republic of Serbia**



MINISTRY OF EDUCATION,  
SCIENCE AND TECHNOLOGICAL DEVELOPMENT

## Technical Co-sponsor

- **Serbia and Montenegro Chapter of IEEE Computer Society**





# Contents

Welcome Message .....	4
Organizing Committee.....	6
Keynote Speakers.....	9
Program at a Glance.....	16
Technical Program .....	18
Book of Abstracts .....	27

# Welcome Message

Dear colleagues and students,

On behalf of the Organizing Committee, it is a pleasure to welcome you at the First Serbian International Conference on Applied Artificial Intelligence which takes place in Kragujevac, Serbia, on May 19<sup>th</sup>-20<sup>th</sup>, 2022 at the University of Kragujevac.

Artificial intelligence and related technologies are increasingly prevalent in business and industry and are beginning to be applied to healthcare, social science and arts. These technologies have the potential to transform many areas. Artificial intelligence has revolutionized information technologies. The new economy of information technologies has shaped the way we live today.

The First Serbian International Conference on Applied Artificial Intelligence 2022 (AAI2022) will provide an excellent international forum for sharing knowledge and results in theory, methodology and application of *Artificial Intelligence* and *Machine Learning* in academia and industry. It is jointly organized by the University of Kragujevac and Serbian Academy of Sciences and Arts.

Nowadays, artificial intelligence has been used in every company where intelligence elements are embedded inside sensors, devices, machines, computers and networks. The conference organizers aim to gather the attention of and contribution from researchers, academicians, and scientists from various fields of *Artificial Intelligence* to create an integrated approach towards global exchange of information on technological advances, scientific innovations, and the effectiveness of various regulatory programs towards *Artificial Intelligence* application.

It also aims to:

- provide early-stage researchers with an inspiring event allowing them to connect to relevant experts in related fields;
- provide an exciting venue for researchers to network and establish national and international collaborations;
- bring together leading experts from all relevant scientific domains to enhance the understanding of *Artificial Intelligence*;
- make historical shift for the very beginning of this area in Serbia and at the University of Kragujevac.

There are five different Mini-symposiums:

- **Mini-symposium I – Soft Computing Techniques in Multi-Criteria Optimization Problems – Organisers: Mališa Žižović and Dragan Pamučar**
- **Mini-symposium II – Artificial Intelligence in Medicine: Current Situation and Future Trends – Organisers: Tijana Šušteršič and Nenad Filipović**
- **Mini-symposium III – Artificial Intelligence Applications in Social Sciences – Organiser: Branko Urošević**
- **Mini-symposium IV – Artificial Intelligence and Entrepreneurship – Organisers: Dubravko Čulibrk and Branislav Kisačanin**

- **Mini-symposium V – Supercomputing for Artificial Intelligence – Organiser: Vladimir Srdanović**

As well as six world renowned plenary speakers in the area of applied artificial intelligence:

- **Prof. Michalis Zervakis** – Technical University of Crete, Greece
- **Prof. Zoran Obradović** – Temple University, USA
- **Prof. Veljko Milutinović** – Purdue University, USA
- **Prof. Themis Exarchos** – Ionian University, Greece
- **Prof. Sašo Džeroski** – IJS Ljubljana, Slovenia
- **Prof. Priyanka Harjule** – Malaviya National Institute of Technology Jaipur, India

We have received more than 120 high-quality research papers. As a result of the strict review process and evaluation, the committee selected 91 papers as extended abstracts.

After the review, full papers from the AAI2022 conference will be published by Springer Verlag in the series “Learning and Analytics in Intelligent Systems” under the title “Applied Artificial Intelligence”.

We must also admit that the conference certainly would not have been so successful without the efforts of many people who were actively engaged in organization of such a major academic event. We express gratitude to the members of the program and scientific review committee as well as to all the chairs, organizers and committee members for their dedication and support.

On behalf of the Organizing Committee, we wish you all a pleasant stay in Kragujevac and a productive conference.

*Prof. Nenad Filipović*, Conference Program Chair

# Organizing Committee

## Co-chairs:

- Nenad Filipović, University of Kragujevac
- Miloš Kojić, Serbian Academy of Sciences and Arts
- Miloš Đuran, Serbian Academy of Sciences and Arts
- Vesna Ranković, University of Kragujevac
- Vladimir Ranković, University of Kragujevac
- Dragan Đurčić, University of Kragujevac
- Dušan Teodorović, Serbian Academy of Sciences and Arts
- Zoran Lj. Petrović, Serbian Academy of Sciences and Arts
- Steva Todorčević, Serbian Academy of Sciences and Arts
- Vladan Devedžić, Serbian Academy of Sciences and Arts

## Local Organizers:

- Tijana Šušteršič, University of Kragujevac
- Smiljana Tomašević, University of Kragujevac
- Aleksandra Vulović, University of Kragujevac
- Anđela Blagojević, University of Kragujevac
- Nikola Radovanović, University of Kragujevac
- Đorđe Dimitrijević, University of Kragujevac
- Dalibor Nikolić, University of Kragujevac
- Igor Saveljić, University of Kragujevac
- Milica Kaplarević, University of Kragujevac
- Marija Gačić, University of Kragujevac
- Neda Vidanović Miletić, University of Kragujevac
- Branko Urošević, School of Computing, Union University
- Boban Stojanović, University of Kragujevac
- Miloš Ivanović, University of Kragujevac
- Branko Arsić, University of Kragujevac

## Program Committee:

- Martin Aleksandrov (TU Berlin, Germany)
- Sandra Avila (University of Campinas (Unicamp), Brazil)
- Christian Blum (Spanish National Research Council (CSIC), Spain)
- Carlos Cardonha (University of Connecticut, United States)
- Vinay Chaudhri (United States)
- John Chinneck (Carleton University, Canada)
- Andy Chun (City University of Hong Kong, Hong Kong)
- Andre Augusto Cire (University of Toronto, Canada)
- Bradley Clement (Jet Propulsion Laboratory, United States)
- Dubravko Čulibrk (University of Novi Sad, Serbia)
- Veljko Milutinović (University of Kragujevac and University of Belgrade, Serbia)
- Diane Cook (Washington State University, United States)



- Gabriella Cortellessa (CNR-ISTC, National Research Council of Italy, Italy)
- Lizhen Cui (Shandong University, China)
- Akay Metin (University of Houston, USA)
- Allen Robert (University of Southampton, UK)
- Zoran Bosnić (University of Ljubljana, Slovenia)
- Zlatan Car (University of Rijeka, Croatia)
- Ciaccio Edward (Columbia University, USA)
- Themis Exarchos (University of Ioannina, Greece)
- Dimitrios Fotiadis (University of Ioannina, Greece)
- Nikola Jorgovanović (University of Novi Sad, Serbia)
- Zoran Marković (Institute of Information Technologies, Serbia)
- Michalopoulos George (University of Pittsburgh, USA)
- Nikita Konstantina (National Technical University of Athens, Greece)
- Zoran Obradović (Temple University, USA)
- Ouzounis Christos (King's College, UK)
- Pattichis Constantinos (University of Cyprus, Cyprus)
- Sheu Phillio (University of California, USA)
- Stojanović Radovan (University of Montenegro, Montenegro)
- Miroslav Trajanović (University of Niš, Serbia)
- Tsiknakis Manolis (Hellenic Mediterranean University, Greece)
- Yang Guang-Zhong (Imperial College London, UK)
- Zervakis Michalis (University of Crete, Greece)
- Andre de Carvalho (University of São Paulo, Brazil)
- Luca Di Gaspero (DPIA – University of Udine, Italy)
- Matthew Gaston (Carnegie Mellon University, United States)
- Carmen Gervet (Université de Montpellier, France)
- Odd Erik Gundersen (Norwegian University of Science and Technology, Norway)
- Koen Hindriks (Vrije Universiteit Amsterdam, Netherlands)
- Neil Jacobstein (Singularity University, United States)
- Binbin Jia (Southeast University, China)
- Elias Khalil (Georgia Institute of Technology, United States)
- Lars Kotthoff (University of Wyoming, United States)
- Hoong Chuin Lau (Singapore Management University, Singapore)
- Jimmy Lee (The Chinese University of Hong Kong, Hong Kong)
- Lee McCluskey (University of Huddersfield, United Kingdom)
- Felipe Meneguzzi (Pontifical Catholic University of Rio Grande do Sul, Brazil)
- Mitra Nasri (Delft University of Technology, Netherlands)
- Barry O'Sullivan (University College Cork, Ireland)
- Michael Orosz (University of Southern California Information Sciences Institute, United States)
- Simon Parsons (University of Lincoln, United Kingdom)
- Andrew Perrault (Harvard University, United States)
- David Pynadath (University of Southern California, United States)
- Claude-Guy Quimper (Laval University, Canada)

- Howard Shrobe (Massachusetts Institute of Technology, United States)
- Madhav Sigdel (University of Alabama in Huntsville, United States)
- David Stracuzzi (Sandia National Laboratories, United States)
- Dimitris Stripelis (University of Southern California, United States)
- Nirmalya Thakur (University of Cincinnati, United States)
- Kevin Tierney (Bielefeld University, Germany)
- Michael Trick (Carnegie Mellon University, United States)
- Pradeep Varakantham (Singapore Management University, Singapore)
- Deng-Bao Wang (Southeast University, China)
- Shinjae Yoo (Brookhaven National Laboratory, United States)
- Yingqian Zhang (Eindhoven University of Technology, Netherlands)
- Jovan Stojanović (Serbian AI Society, Serbia)
- Stefan Badža (Serbian Government, Serbia)