INTERNATIONAL NEURAL NETWORK SOCIETY

Nov 2021

INNS Members News

https://www.inns.org/

INSIDE

INNS is proud to announce the following awardees:

Gabor Award: **Plamen Angelov**
Helmholtz Award: **Tianzi Jiang**
Aharon Katzir Young Investigator Award: **Shujian Yu and Xiangnan Zhong**
INNS Doctoral Dissertation Award: **Leonardo Enzo Brito da Silva**

Happy Holidays

– By Chrisina Jayne, INNS President

On behalf of the International Neural Networks Society, we would like to thank all our INNS members for their continued support over this past year. We appreciate your membership and engagement in the society and wish you and your families a joyous holiday season.

When AI Instabilities are Expected: The Feasibility and Inevitability of Stealth Attacks

By Ivan Tyukin, University of Leicester, School of Computing and Mathematical Sciences, United Kingdom

One of the key questions and challenges for modern data-driven high-dimensional AI based on large-scale deep learning or more shallow models is the question of their stability. These questions are becoming particularly acute in view of the global push driven by governments and markets to use such data-driven AI models in high-stake applications including in autonomous self-driving cars, health and healthcare technologies, and finance and security sectors. Several notions of AI stability have been developed to date. One group of stability definitions focuses on AI sensitivity to data perturbations which are, on the one hand, imperceivable to a relevant observer but, on the other hand, are capable of altering or corrupting performance of trained models. These perturbations are often referred to as adversarial data. The other group of stability notions focus on the model's capability to retain its performance and functionality in presence of perturbations to its parameters and structure.

Recent work by Ivan Tyukin, Desmond Higham, Eliyas Woldegeorgis, and Alexander Gorban https://arxiv.org/pdf/2106.13997v2.pdf developed and studied adversarial perturbations of the second type. These perturbations enable an attacker to gain control over decisions in generic Artificial Intelligence (AI) systems including deep learning neural networks. In contrast to adversarial data modification, the attack mechanism considered involves alterations to the AI system itself. Such a stealth attack, which could in some cases be classified as a backdoor attack, could be conducted by a mischievous, corrupt or disgruntled member of a software development team. It could also be made by those wishing to exploit a "democratization of AI" agenda, where network architectures and trained parameter sets are shared publicly.

The work reveals a range of implementable attack strategies with accompanying analysis, showing that with high probability these attacks can be made transparent (and hence the name – stealth attacks), in the sense that system performance is guaranteed to be unchanged on a fixed validation set which is unknown to the attacker, while evoking any desired output on a trigger input of interest. The attacker only needs to have estimates of the size of the validation set and the spread of the AI's relevant latent space. In the case of deep learning neural networks, it has been shown that a one neuron attack is possible—a modification to the weights and bias associated with a single neuron—revealing a vulnerability arising from over-parameterization. These concepts and results are illustrated with examples. Guided by the theory and computational results, the work proposes strategies to guard against stealth attacks.

https://arxiv.org/pdf/2106.13997v2.pdf

INNS Members News

ICANN 2021 organization

– By Igor Farkaš, Comenius University in Bratislava

As a proud member of both INNS and ENNS, I had a great opportunity, as a general chair, to contribute to organization of ICANN 2021 planned in Bratislava, Slovak Republic. The online version of the conference that we had to decide for shifted the nature of the concept inevitably, but this is something we all had to get used to, trying to see the pros and not so much cons, like the lack of physical contact and socializing. We enjoyed a significantly increased number of submissions, and online attendance, and believe it was a successful online event, as confirmed by numerous feedbacks. On the other hand, we were sorry we had had bad luck because of the pandemics. Anyway, this was a good experience for our organizing team to whom I am thankful and the time will show how these online events find their place in the community in the post-covid future.

Announcement: Prof. Emeritus, Toyohashi Univ. of Technology

Professor Shiro Usui has retired from Toyohashi Univ. of Technology and becomes Professor Emeritus. https://www.linkedin.com/in/shiro-usui-916330b3/?originalSubdomain=jp

CALL FOR PAPERS

The Thirty-Fifth International Conference on Industrial, Engineering & Other Applications of Applied Intelligent Systems (IEA/AIE-2022) – Kitakyushu, Japan – July 19- July 22, 2022

IEA/AIE-2022 Conference website: https://ieaaie2022.wordpress.com/
Sponsored by: International Society of Applied Intelligence (ISAI) and
In Cooperation with: Assoc. for the Advancement of Art. Intell. (AAAI) / Assoc. for Computing Machinery (ACM/SIGAI) / Austrian Assoc. for Art. Intelligence (OEGAI) / Catalan Assoc. for Art. Intelligence (ACIA) / Graz University of Technology / Italian Art. Intelligence Assoc. (AI*IA) / Iwate Prefectural University, Japan / Japanese Society for Art. Intelligence (JSAI) / Lithuanian Computer Soc. – AI Section (LIKS-AIS) / Spanish Society for Art. Intelligence (AEPIA) / Soc. for the Study of AI and the Sim. of Behav. (AISB) /Taiwanese Association for Consumer Electronics (TACE) / Taiwanese Association for Artificial Intelligence (TAAI) / Texas State University, USA / University of Klagenfurt, Austria
IEA/AIE 2022 continues the tradition of emphasizing applications of applied intelligent systems to solve real-life problems in all areas including engineering, science, industry, automation & robotics, business & finance, medicine and biomedicine, bioinformatics, cyberspace, and human-machine interactions. IEA/AIE-2022 will include oral presentations, invited speakers, and special sessions. Paper submission is required by December 15, 2021. Submission instructions and additional details may be obtained from the website: https://ieaaie2022.wordpress.com/