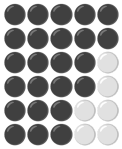
SKILLS

Python,
Tensorflow, PyTorch,
Java,
C, C++, C#,
Mathematica, Julia,
LaTeX,



ACHIEVEMENTS

Ph.D. Student Scholarship in the project *Bio-inspired artificial neural networks* carried out within the TEAM-NET programme of the Foundation of Polish Science.
October 2019 - October 2022

Scientific Scholarship granted for participation in a research project financed by the Polish National Science Centre
August 2018 - January 2019

Stipend for the best students, awarded by the Rector of Jagiellonian University
Academic years: 2018/2019, 2015/2016, 2014/2015

Graduated in top 5% students, of the first-cycle programme of study (B.Sc) in Computer Science
September 2016

LANGUAGES

English
TOEFL iBT
score: 108

Polish
German

ACTIVITIES

Reviewer, at ECML PKDD 2022, ICML 2022, ICLR 2022, ICONIP 2019/20/21

Member, of the ML in PL Association,
https://mlinpl.org/
January 2021 - present

Member, of the Group Of Machine Learning Research (GMUM) at Jagiellonian University
https://gmum.net/
July 2017 - present

Participant in the Women in Tech Camp 2018 Programme, Kielce, Poland
September 2018

Member, of the Computer Science Students Association
October 2018 - September 2019

Member, of the Applied Computer Science Students Association
February 2013 - September 2016

Volunteer, 11th Kraków's Slot Art Festival
March 2014

INTERESTS

Graphics and architecture drawing, history, history of art, playing violin

Aleksandra Irena Nowak

Kraków, Poland, **(+48) 511 160 952**
aleksandrairena.nowak@doctoral.uj.edu.pl
https://github.com/alooow

EDUCATION

Ph.D. programme in Computer Science, Jagiellonian University, *Faculty of Mathematics and Computer Science*
October 2019 - September 2024

Visiting Ph.D. Candidate, University of Twente, Netherlands,
September 2022 - January 2023
Visiting the group of dr. Decebal Mocanu, developing and analyzing dynamic sparse training solutions in deep learning.

Computer Science M.Sc., specialization: Machine Learning, Jagiellonian University, *Faculty of Mathematics and Computer Science*
October 2017 - September 2019
Diploma With The Award of Distinction

IT Analyst B.Sc., Jagiellonian University, *Theoretical Computer Science, Faculty of Mathematics and Computer Science*
October 2015 - July 2017

Computer Science B.Sc., Jagiellonian University, *Faculty of Physics, Astronomy and Applied Computer Science*
October 2013 - September 2016
Diploma With The Award of Distinction

A. Witkowski Vth High School, Kraków, *Extended program forms in: mathematics, physics, and computer science*
September 2010 - April 2013
Matriculation exam in mathematics (advanced level): 100%

EXPERIENCE

Google DeepMind (Student Researcher), Montreal, Canada
October 2023 - February 2024
Student Researcher, working on transfer learning and efficient adaptation.

IDEAS NCBR, Poland
January 2023 - September 2023
PhD Researcher, working on efficient deep neural networks and sparse architectures.

Self-Employment, Poland
August 2018 - September 2018
Internship in Data Science Team in a U.S. Financial Services Company, working on preparation and analysis of time series data.

Making Waves, Poland — *Internship*
July 2016
Working in the .NET Developer Team, preparing diagnostic solutions based on system logging information.

Estimote, Poland — *Internship*
July 2015 - September 2015
Working in the Data Science Team, developing the Indoor Localization mobile application and collecting data for beacon signal analysis.

TRAINING AND WORKSHOPS

ICLR23 SNN Workshop, Kigali, Rwanda ,
May 2023, https://www.sparseneural.net/organizers; Organizing Workshop on Sparsity in Neural Networks during ICLR 2023.

MLSS^N Summer School, Krakow, Poland ,
June/July 2022, https://mlss.mlinpl.org/; Organizing committee

From Neuroscience to Artificially Intelligent Systems (NAISys), CSHL, USA
April 2022, https://meetings.cshl.edu/abstracts.aspx?meet=NAISYS&year=22; Presenting work on analyzing neural network architectures based on random graphs.

ML in PL Conference 2021, Warsaw, Poland (online),
November 2021, https://conference2021.mlinpl.org/; Organizing committee

Google Women in Tech Mentoring Program, Poland, (online),
September-November 2021; Participated in a program for 31 selected students across Poland. The program involved regular mentoring sessions, learning best industry practices, and group sessions.

Virtual Platform Support, EEML2021 & OxML2020 Summer Schools (online),
July 2021, August 2020; Helping with designing the virtual platform for the Summer Schools conducted during the COVID-19 lockdown, https://www.eeml.eu/home, https://www.oxfordml.school/

Eastern European Machine Learning Summer School 2020, Kraków, Poland (online) ,
July 2020, https://www.eeml.eu/previous-editions/eeml2020; Organizing committee

ML in PL Conference, Warsaw, Poland ,
November 2019, https://conference.mlinpl.org/; Co-conducting workshops in reinforcement learning.

Conference on Theoretical Foundations of Machine Learning, Kraków, Poland.,
February 2019, http://tfml.gmum.net/; Organizing committee

Mediterranean School of Complex Networks, Salina, Italy, September 2017

PUBLICATIONS AND PREPRINTS

Aleksandra Nowak, Łukasz Gniecki, Filip Szatkowski, Jacek Tabor - Sparser, Better, Deeper, Stronger: Improving Sparse Training with Exact Orthogonal Initialization, ICML 2024

Abstract: *Static sparse training aims to train sparse models from scratch, achieving remarkable results in recent years. A key design choice is given by the sparse initialization, which determines the trainable sub-network through a binary mask. Existing methods mainly select such mask based on a predefined dense initialization. Such an approach may not efficiently leverage the mask's potential impact on the optimization. An alternative direction, inspired by research into dynamical isometry, is to introduce orthogonality in the sparse subnetwork, which helps in stabilizing the gradient signal. In this work, we propose Exact Orthogonal Initialization (EOI), a novel sparse orthogonal initialization scheme based on composing random Givens rotations. Contrary to other existing approaches, our method provides exact (not approximated) orthogonality and enables the creation of layers with arbitrary densities. We demonstrate the superior effectiveness and efficiency of EOI through experiments, consistently outperforming common sparse initialization techniques. Our method enables training highly sparse 1000-layer MLP and CNN networks without residual connections or normalization techniques, emphasizing the crucial role of weight initialization in static sparse training alongside sparse mask selection.*

Aleksandra Nowak, Bram Grooten, Decebal C. Mocanu, Jacek Tabor - Fantastic Weights and How to Find Them: Where to Prune in Dynamic Sparse Training, NeurIPS 2023

Abstract: *Dynamic Sparse Training (DST) is a rapidly evolving area of research that seeks to optimize the sparse initialization of a neural network by adapting its topology during training. It has been shown that under specific conditions, DST is able to outperform dense models. The key components of this framework are the pruning and growing criteria, which are repeatedly applied during the training process to adjust the network's sparse connectivity. While the growing criterion's impact on DST performance is relatively well studied, the influence of the pruning criterion remains overlooked. To address this issue, we design and perform an extensive empirical analysis of various pruning criteria to better understand their effect on the dynamics of DST solutions. Surprisingly, we find that most of the studied methods yield similar results. The differences become more significant in the low-density regime, where the best performance is predominantly given by the simplest technique: magnitude-based pruning. The code is provided at https://github.com/alooow/fantastic_weights_paper*

Mateusz Olko*, Michał Zając*, Aleksandra Nowak*, Nino Scherrer, Yashas Annadani, Stefan Bauer, Łukasz Kuciński, Piotr Miłoś - *Trust Your ∇ : Gradient-based Intervention Targeting for Causal Discovery*, NeurIPS 2023

Abstract: *Inferring causal structure from data is a challenging task of fundamental importance in science. Observational data are often insufficient to identify a system's causal structure uniquely. While conducting interventions (i.e., experiments) can improve the identifiability, such samples are usually challenging and expensive to obtain. Hence, experimental design approaches for causal discovery aim to minimize the number of interventions by estimating the most informative intervention target. In this work, we propose a novel Gradient-based Intervention Targeting method, abbreviated GIT, that 'trusts' the gradient estimator of a gradient-based causal discovery framework to provide signals for the intervention acquisition function. We provide extensive experiments in simulated and real-world datasets and demonstrate that GIT performs on par with competitive baselines, surpassing them in the low-data regime.*

Aleksandra Nowak, Romuald Janik - *Discovering wiring patterns influencing neural networks performance* *ECML PKDD*, 2022,

Abstract: *We perfome a massive evaluation of artificial neural networks with various computational architectures, where the diversity of the studied constructions is obtained by basing the wiring topology of the networks on different types of random graphs. Our goal is to investigate the structural and numerical properties of the graphs and assess their relation to the test accuracy of the corresponding neural networks. We find that none of the classical numerical graph invariants by itself allows to single out the best networks. Consequently, we introduce a new numerical graph characteristic, called quasi-1-dimensionality, which is able to identify the majority of the best-performing graph. We provide a dataset of 1020 graphs and the test accuracies of their corresponding neural networks at https://github.com/rmldj/random-graph-nn-paper*

Łukasz Maziarka, Aleksandra Nowak, Maciej Wołczyk, Andrzej Bedychaj - *On the relationship between disentanglement and multi-task learning*, *ECML PKDD*, 2022,

Abstract: *One of the main arguments behind studying disentangled representations is the assumption that they can be easily reused in different tasks. At the same time finding a joint, adaptable representation of data is one of the key challenges in the multi-task learning setting. In this paper, we take a closer look at the relationship between disentanglement and multi-task learning based on hard parameter sharing. We perform a thorough empirical study of the representations obtained by neural networks trained on automatically generated supervised tasks. Using a set of standard metrics we show that disentanglement appears naturally during the process of multi-task neural network training.*

Roumald Janik, Aleksandra Nowak - *Neural Networks Adapting to Datasets: Learning Network Size and Topology*, *preprint*, *arXiv.org:2006.12195*, *Dynamic Neural Networks, ICML 2022 Workshop*

Abstract: *We introduce a flexible setup allowing for a neural network to learn both its size and topology during the course of a standard gradient-based training. The resulting network has the structure of a graph tailored to the particular learning task and dataset. The obtained networks can also be trained from scratch and achieve virtually identical performance. We explore the properties of the network architectures for a number of datasets of varying difficulty observing systematic regularities. The obtained graphs can be therefore understood as encoding nontrivial characteristics of the particular classification tasks.*

Marcin Sendera, Jacek Tabor, Aleksandra Nowak, Andrzej Bedychaj, Massimiliano Patacchiola, Tomasz Trzcinski, Przemysław Spurek, Maciej Zieba - Non-Gaussian Gaussian Processes for Few Shot Regression, NeurIPS 2021

Abstract: *Gaussian Processes (GPs) have been widely used in machine learning to model distributions over functions, with applications including multi-modal regression, time-series prediction, and few-shot learning. GPs are particularly useful in the last application since they rely on Normal distributions and, hence, enable closed-form computation of the posterior probability function. Unfortunately,*

because the resulting posterior is not flexible enough to capture complex distributions, GPs assume high similarity between subsequent tasks - a requirement rarely met in real-world conditions. In this work, we address this limitation by leveraging the flexibility of Normalizing Flows to modulate the posterior predictive distribution of the GP. This makes the GP posterior locally non-Gaussian, therefore we name our method Non-Gaussian Gaussian Processes (NGGPs).

Przemysław Spurek, Aleksandra Nowak, Jacek Tabor, Łukasz Maziarka, Stanisław Jastrzębski - *Non-linear ICA based on Cramer-Wold metric*, International Conference on Neural Information Processing (ICONIP 2020), Lecture Notes in Computer Science, Part III, pp. 294-305, 2020

Abstract: *Non-linear source separation is a challenging open problem with many applications. We extend a recently proposed Adversarial Non-linear ICA (ANICA) model, and introduce Cramer-Wold ICA (CW-ICA). In contrast to ANICA we use a simple, closed--form optimization target instead of a discriminator--based independence measure. Our results show that CW-ICA achieves comparable results to ANICA, while foregoing the need for adversarial training.*

Łukasz Maziarka et al. - *Set Aggregation Network as a Trainable Pooling Layer*, International Conference on Neural Information Processing (ICONIP 2019), Lecture Notes in Computer Science, pp. 419-431, 2019

Abstract: *Global pooling, such as max- or sum-pooling, is one of the key ingredients in deep neural networks used for processing images, texts, graphs and other types of structured data. Based on the recent DeepSets architecture proposed by Zaheer et al. (NIPS 2017), we introduce a Set Aggregation Network (SAN) as an alternative global pooling layer. In contrast to typical pooling operators, SAN allows to embed a given set of features to a vector representation of arbitrary size. We show that by adjusting the size of embedding, SAN is capable of preserving the whole information from the input. In experiments, we demonstrate that replacing global pooling layer by SAN leads to the improvement of classification accuracy. Moreover, it is less prone to overfitting and can be used as a regularizer.*

Łukasz Struski, Jacek Tabor, Igor Podolak, Aleksandra Nowak, Krzysztof Maziarz - *Realism Index: Interpolation in Generative Models With Arbitrary Prior*, preprint, *arXiv:1904.03445*, *2019*

Abstract: *In order to perform plausible interpolations in the latent space of a generative model, we need a measure that credibly reflects if a point in an interpolation is close to the data manifold being modeled, i.e. if it is convincing. In this paper, we introduce a realism index of a point, which can be constructed from an arbitrary prior density, or based on FID score approach in case a prior is not available. We propose a numerically efficient algorithm that directly maximises the realism index of an interpolation which, as we theoretically prove, leads to a search of a geodesic with respect to the corresponding Riemann structure. We show that we obtain better interpolations then the classical linear ones, in particular when either the prior density is not convex shaped, or when the soap bubble effect appears.*

Andrzej Bedychaj, Przemysław Spurek, Aleksandra Nowak, Jacek Tabor -*WICA: nonlinear weighted ICA*, preprint, arxiv:2001.04147, 2019

Abstract: *Independent Component Analysis (ICA) aims to find a coordinate system in which the components of the data are independent. In this paper we construct a new nonlinear ICA model, called WICA, which obtains better and more stable results than other algorithms. A crucial tool is given by a new efficient method of verifying nonlinear dependence with the use of computation of correlation coefficients for normally weighted data.*

Michał Cieśla, Aleksandra Nowak - *Managing numerical errors in random sequential adsorption*, Surface Science, v.651 p.182 - 186, 2016

Abstract: *Aim of this study is to examine the influence of a finite surface size and a finite simulation time on a packing fraction estimated using random sequential adsorption simulations. The goal of particular interest is providing hints on simulation setup to achieve desired level of accuracy. The analysis is based on the properties of saturated random packing of disks on continuous and flat surfaces of different sizes.*