

Hanna Tseran, Ph.D.

MACHINE LEARNING RESEARCHER

Tokyo, Japan (Open to relocation) • hanna.tseran@gmail.com • +81-90-7736-2572
hanna-tseran.github.io • [Google Scholar](https://scholar.google.com/citations?user=HannaTseran) • [linkedin.com/in/hanna-tseran](https://www.linkedin.com/in/hanna-tseran) • github.com/hanna-tseran

SUMMARY

Machine learning researcher with **diverse academic and practical experience** developed through postdoctoral positions, industry roles, and graduate studies. My work centers on **deep learning theory**, aiming to establish **foundational principles for modern models such as LLMs** to support safe, robust, and explainable artificial intelligence. I currently lead several projects on the **mathematical analysis of in-context learning and transformer training dynamics**. I have published in NeurIPS, ICML, and TMLR, and secured a competitive JSPS Early-Career Scientist grant. I now seek to apply my experience to high-impact challenges in industry, where access to real-world systems enables practically relevant research.

SKILLS

Technical Skills: Python, PyTorch, TensorFlow, Parallel Computing (MPI)

Research Areas: Deep Learning Theory, Large Language Models (LLMs), In-Context Learning, Optimization, Efficient Transformers

EXPERIENCE

RIKEN Center for Advanced Intelligence Project (AIP)

Postdoctoral Researcher

Tokyo, Japan

Apr 2024 – present

High-Dimensional Structure Theory Team | Team leader: Prof. Masaaki Imaizumi

- Lead several lines of research analyzing **LLMs**, with a focus on **transformer optimization, in-context learning**, and general **emergent abilities**.
- Secured a JSPS Grant-in-Aid for Early-Career Scientists to support the development of a theoretical framework for in-context learning based on singular learning theory.
- Engage in student supervision and collaboration within the joint RIKEN–University of Tokyo lab and beyond.

The University of Tokyo

Project Researcher

Tokyo, Japan

Nov 2023 – Mar 2024

Matsuo Lab

- Explored theoretical directions in **LLMs**, including work on **training dynamics** now continued at RIKEN.

Amazon

Applied Scientist Intern

Berlin, Germany

Nov 2022 – Mar 2023

Natural Language Processing Team

- Developed an **efficient memory-augmented transformer** architecture for conversational AI, enabling processing of sequences of unbounded length.

Microsoft Research

Research Software Engineer

Cambridge, UK

Dec 2018 – Dec 2019

TrueSkill project for estimating player skill based on Bayesian networks; used by Halo, the game selling 81MM+ copies.

A team of 5 people, research led by Dr. Tom Minka

- Explored system applicability to alternative game designs, obtaining quantitative results.
- Quickly ramped up in C# to enhance metrics for performance analysis and improve library design.

RIKEN Center for Advanced Intelligence Project (AIP)

Research Assistant

Tokyo, Japan

Nov 2017 – Aug 2018

Approximate Bayesian Inference Team. Team leader: Dr. M. Emtiyaz Khan

- Derived a novel continual learning method based on the approximate variational inference algorithm for Bayesian neural networks and wrote a paper accepted to a NeurIPS workshop.

Google

Site Reliability Engineering (SRE) Intern

Dublin, Ireland

Jul 2017 – Oct 2017

- Prototyped a deep learning system for identifying spam in SRE alerts based on anomaly detection techniques.

Yandex

Software Engineer

Minsk, Belarus

Dec 2014 – Mar 2016

Backend team of around 20 people working on the Yandex search engine written mostly in C++, the most popular search engine in Russia at the time, with 100MM+ daily queries

- Accelerated loading and reduced memory consumption of the search engine.
- Designed and implemented an approach to optimize data center balancing.
- Discovered and implemented a method to reduce search database size without information loss, saving approximately 1M\$ in storage costs.

CheckPoint

Software Engineer

Minsk, Belarus

Jul 2013 – Nov 2014

- Diagnosed and resolved customer-reported issues in a C codebase for enterprise media encryption and anti-malware products.

EDUCATION

Max Planck Institute for Mathematics in the Sciences (MPI MiS)

Ph.D. in Computer Science

Leipzig, Germany

Jan 2020 – Nov 2023

Thesis: Expected Complexity and Gradients of Deep Maxout Neural Networks and Implications to Parameter Initialization. Supervisor: Prof. Guido Montúfar (Group Leader at MPI MiS and Professor at UCLA).

Degree awarded by Leipzig University. Research conducted at MPI MiS with parallel enrollment in IMPRS.

- Proposed a **stable initialization method** for deep maxout networks, achieving over **40% accuracy improvement**; published in ICML.
- Proved that **expected complexity** grows polynomially with depth in maxout networks despite earlier assumptions of exponential growth; published in NeurIPS.
- Contributed to a study on **loss landscapes of ReLU networks**; published in TMLR.

The University of Tokyo

Master of Information Science and Technology

Tokyo, Japan

Sep 2016 – Sep 2018

Thesis: Variational Inference for Continual Learning by using Weight-Perturbation in Adam. Supervisor: Prof. Tatsuya Harada.

Belarusian State University

Specialist Degree in Computer Science

Minsk, Belarus

Sep 2010 – Jun 2015

Thesis: Algorithms for recognition of circular objects and elements on them (in case of coins). Supervisor: Prof. Yuri Svirid.

GRANTS & SCHOLARSHIPS

JSPS Grant-in-Aid for Early-Career Scientists for 4,810,000¥ ($\approx 33,000\$$) | Principal Investigator | Topic: Theoretical Framework for In-Context Learning Development in LLMs Based on Singular Learning Theory | Apr 2025 – Mar 2027

Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) Scholarship, one of two recipients from Belarus | Full tuition and stipend | Apr 2016 – Aug 2018

SELECTED PUBLICATIONS

Karhadkar, Kedar and Murray, Michael and **Tseran, Hanna** and Montúfar, Guido. **Mildly Overparameterized ReLU Networks Have a Favorable Loss Landscape**. *Transactions of Machine Learning Research, TMLR* (2024)

Tseran, Hanna, and Montúfar, Guido. **Expected Gradients of Maxout Networks and Consequences to Parameter Initialization**. *International Conference on Machine Learning, ICML* (2023)

Tseran, Hanna, and Montúfar, Guido. **On the Expected Complexity of Maxout Networks**. *Advances in Neural Information Processing Systems, NeurIPS* (2021)