

# Augustas Macijauskas

Email: [august.macijauskas@gmail.com](mailto:august.macijauskas@gmail.com) • Phone: +37061001317 • LinkedIn: [linkedin.com/in/augustas-macijauskas](https://www.linkedin.com/in/augustas-macijauskas)  
• GitHub: [github.com/AugustasMacijauskas](https://github.com/AugustasMacijauskas) • Website/blog: [augustasmacijauskas.github.io/personal-website](https://augustasmacijauskas.github.io/personal-website)

## EDUCATION

---

**University of Cambridge** (October 2022 – September 2023; Cambridge, United Kingdom)

- Machine Learning and Machine Intelligence (MPhil, 77.66%, **distinction**).
- Notable topics studied: **Deep Learning; Neural Machine Translation; Reinforcement Learning; GNNs**.
- Thesis on interpretability and alignment of LLMs. Supervised by Dr Samuel Albanie and Herbie Bradley.

**The University of Manchester** (September 2019 – June 2022; Manchester, United Kingdom)

- Mathematics (BSc, 91.7%, 1st, 4th rank overall).
- Final project: **Numerical Solutions to the Navier-Stokes Equations**. Supervised by Dr Matthias Heil.

## WORK EXPERIENCE and PROJECTS

---

**Spike Technologies, Inc.** (*Machine Learning Engineer*, January 2024 – present, remote)

- Contributed to developing an advanced agentic chatbot designed to answer user inquiries about health-related topics.
  - Created evaluation pipelines and conducted rigorous experiments to enhance the reliability of individual subagents.
  - Implemented new features that improved chatbot functionality (RAG, guardrails), ensuring the system was production-ready and capable of handling real-world health queries with near-perfect accuracy and grounding.
- Led the development of a multimodal model-based food scanner API designed to simplify meal logging by automatically identifying food items and calculating their nutritional values, currently used by thousands of users.
  - Optimized the system for accuracy, robustness to diverse inputs, and efficiency before deploying it in production.

**Thesis Project: Eliciting Latent Knowledge from Lange Reward Models** (May 2023 – September 2023, Cambridge, UK)

- Created a method that allows using linear classifiers trained on top of a model's activations (referred to as *discovering latent knowledge* (DLK)) **to build reward models that promote truthfulness** (in a narrow sense).
- Utilized the trained reward models to fine-tune pre-trained *large language models* (LLMs) to be more truthful by using the *proximal policy optimization* (PPO) *reinforcement learning* (RL) algorithm.
  - Adopted **efficient fine-tuning and evaluation strategies**, such as distributed *data-parallel training* (DDP), *low-rank adaptation* (LoRA), and *quantization*.
  - Created scripts to **automatically launch training jobs** on a computing cluster equipped with **SLURM**.
- Improved the truthfulness of pre-trained LLMs by **up to 1.6%**, as measured by the TruthfulQA benchmark, **without compromising the models' performance on general NLP tasks**.

**Baltic Institute of Advanced Technology (BPTI)** (*Research Assistant*, July 2020 – September 2022; Vilnius, Lithuania)

- Investigation of object **3D geometry reconstruction** using **neural radiance fields**.
  - Read papers, browsed code repositories and optimized them for both training and inference efficiency.
  - Achieved **satisfactory neural view synthesis and reconstruction quality for reflective objects**.
  - Summarized all the successes and learnings in a scientific report and presented it to relevant stakeholders.
- Replicated the Point Transformer architecture for 3D point cloud classification and segmentation.
  - Tweaked the above model to segment out artificially added noise in an efficient way.
- Developed a PyTorch model that utilizes **similarity learning using Triplet loss to perform real-world visa stamp recognition**, i.e. classifying the country and direction of travel (arXiv preprint: <https://arxiv.org/abs/2112.00348>).
  - Achieved **93% accuracy** on unseen validation data using a ResNet-18 Siamese network architecture.
  - Wrote an API that allowed the team to deploy the trained model for demonstration purposes.

## SKILLS

---

**Programming Languages:** Python, JavaScript, MATLAB, C++.

**Frameworks, libraries and tools:** PyTorch, transformers, datasets, trl, accelerate, AWS, PyTorch Lightning, scikit-learn.

**Soft skills:** Leadership, communication, pitching, teaching.

**Languages:** native in Lithuanian, fluent in English.