Ya Gao

ya.gao@aalto.fi • +358 449393437 • LinkedIn • github.com/yagao403
Personal Website • Google Scholar

Education

- **Aalto University** *May 2023 - present*
  PhD, Doctoral progamme in Science.
  - Supervisor: Pekka Marttinen
- **Aalto University** *Sept 2020 - Dec 2022*
  MSc, Machine Learning, Data Science and Artificial Intelligence. (with Honours)
  - GPA: 4.8/5.0
  - Master thesis: Joint entity and relation extraction via contrastive learning on knowledge-augmented graph embeddings
- **Sichuan University** *Sept 2016 - Jun 2020*
  BEng, Computer Science
  - GPA: 89/100
  - Bachelor thesis: Bi-level attention mechanism for emotion cause extraction

Research Interests

LLM Post-training, LLM Reasoning & Agency, LLMs for Healthcare

Professional Experience

- **Doctoral Researcher** May, 2023 – Present
  Machine Learning for Health (ML4H) Group, Aalto University Espoo, Finland
  - **Training LLM Agents for Multi-Step Tool Use (Industry collab. w/ System 2 AI)**
    * Acted as **lead contributor** in training a Llama-3.1-70B agent for complex, agentic workflows. Engineered the full training pipeline, including data collection, context distillation, and refinement by human feedback. Achieved state-of-the-art results; the final agent outperformed GPT-4o on ToolQA and OfficeBench.
  - **LLMs Knowledge Editing (Industry collab. w/ System 2 AI)**
    * Developed a context distillation-based editing method for large models (Qwen3-32B), aiming at successfully updating knowledge while preserving complex reasoning abilities post-edit. Analyzed post-edit model integrity by evaluating performance differences between LLMs in the think or Non-thinking mode.
  - **LLMs Skill Composition**
    * Designed a novel student-teacher interaction framework to improve skill generalization in post-merger LLMs. Leveraged on-policy preference learning for model refinement.
  - **LLMs for Healthcare**
    * **Self-Supervised nursing note summarization**: Funded by CLISHEAT project. Proposed a self-supervised algorithm enabling small models to match GPT-4's performance on nursing note summarization.
    * **LLMs for digital phenotyping data**: Benchmarked LLM families (Llama, Gemma, Qwen) for analyzing digital phenotyping data to predict depression severity.
- **Research Intern** Oct, 2021 – Sept, 2022
  Machine Learning for Health (ML4H) Group, Aalto University Espoo, Helsinki
  Adverse Drug Event (ADE) Detection: Developed a contextualized heterogeneous graph embedding model to find causal relations in social media text. Augmented pre-trained embeddings with medical knowledge from UMLS to improve entity awareness. Achieved state-of-the-art performance on the ADE detection benchmark.

- **Startup Co-Founder** Feb, 2019 – Nov, 2019
  Artificial Intelligence for Natural insights (AiNi) Hangzhou, China
  Co-founded a startup to build a hand gesture recognition system for accessibility. Led the end-to-end development, from real-world data collection to vision model training.
- **Summer Intern** Jun, 2018 – Aug, 2018
  iFLYTEK Co., Ltd. Hefei, China
  Designed and implemented deep learning models for representation learning from semi-structured knowledge.
- **Research Assistant** Jul, 2017 – Sept, 2017
  Tsinghua University Beijing, China
  Constructed a video watermark dataset for anti-counterfeiting research. Benchmarked deep learning models for watermark extraction from steganographic images.

Publications

Equal contributor*; Advisor†

- Zhentao Zou, **Ya Gao**†, Jiarui Guan, Bin Li, Pekka Marttinen. "Adaptive Residual-Update Steering for Low-Overhead Hallucination Mitigation in Large Vision Language Models". *Under review*, 2025.
- Minttu Alakuijala*, **Ya Gao***, Georgy Ananov, Samuel Kaski, Pekka Marttinen, Alexander Ilin, Harri Valpola. "Memento No More: Coaching AI Agents to Master Multiple Tasks via Hints Internalization". *Under review*, 2025.
- Yunhao Yuan*, **Ya Gao***, Hans Moen, Erkki Isometsä, Pekka Marttinen, Talayeh Aledavood. "Leveraging Large Language Models for Digital Phenotyping: Detecting Depressive State Changes for Patients with Depressive Episodes". *Under review*, 2025.
- **Ya Gao**, Hans Moen, Saila Koivusalo, Miika Koskinen, Pekka Marttinen. "Query-Guided Self-Supervised Summarization of Nursing Notes". In *ML4H 2024*.
- **Ya Gao***, Shaoxiong Ji*, Pekka Marttinen. "Knowledge-augmented Graph Neural Networks with Concept-aware Attention for Adverse Drug Event Detection". In *LREC-COLING 2024*.
- **Ya Gao**, Shaoxiong Ji, Tongxuan Zhang, Prayag Tiwari, Pekka Marttinen. "Contextualized graph embeddings for adverse drug event detection". In *ECML-PKDD 2022*.

Skills

- **Programming:** Python, C/C++, Jax, Bash, SQL
- **Frameworks, Tools & Platforms:** PyTorch, Hugging Face (Transformers, Datasets, Accelerate, TRL), vLLM, Verl, LangChain, Git, Docker, SLURM, Weights & Biases, Google Cloud, AWS
- **Languages:** English (C1), Mandarin (native)

Professional Service & Awards

- **Reviewer:** ICLR (2025, 2026), ML4H 2025
- **Thesis Advisor:** Master's Thesis (Zhengtao Zou); Bachelor's Thesis (Jeheon Kim)
- **Teaching Assistant:** CS-E407525 Seminar on LLMs (2025); CS-E407517 Seminar on NLP (2024); CS-C3240 Machine Learning (2023, 2024, 2025)
- **Award:** Faculty's Teaching Assistants of the Year Award, 2023; Honours Program in Computer Science, 2022; International Talent Program, 2021; Dean's List Scholarships, 2021; Aalto University Scholarship (100% tuition waiver), 2020.