

CLEMENT OU

650-942-4296 | clement.h.ou@gmail.com | [linkedin.com/in/clementou](https://www.linkedin.com/in/clementou) | github.com/clementou | Pittsburgh, PA

EDUCATION

- **Carnegie Mellon University** Pittsburgh, PA
May. 2025
B.S. in Statistics and Machine Learning, Minor in Machine Learning
 - **Relevant Coursework:** Computer Systems, Distribute Systems, Parallel Computing & Architecture, Algorithms & Data Structures, Cloud Computing, Software Engineering, Distributed ML, Deep Learning, Probability and Inference
 - **Activities:** Students Using Data for Social Good (Data Project Chair), Data Science Club (Competitions & Outreach Director)
 - **Awards:** AI Governance Best Team Project Award (Zoom AI Companion)

EXPERIENCE

- **Quantifind** Palo Alto, CA
Jan. 2025 - Present
Part-time Data Science Intern
 - Developing seq2seq transformer model for Russian-English organization name translation.
- **CMU Software and Societal Systems Department** Pittsburgh, PA
Aug. 2024 - Dec. 2024
Teaching Assistant (17-214 Software Engineering)
 - Reviewed student projects for OOP principles and design patterns, providing detailed feedback on code quality and style.
 - Led weekly recitations and office hours for 30+ students, mentoring on software engineering concepts and best practices.
- **Quantifind** Palo Alto, CA
Jun. 2024 - Aug. 2024
Data Science Intern
 - Engineered ETL pipeline in Scala, cleaning and integrating 130k entities and 300k structured sanctions into core product.
 - Built scalable boosted tree model using PySpark, achieving 95% precision classifying risk factors in 50M documents.
- **CMU Language Technologies Institute** Pittsburgh, PA
Dec. 2023 - May. 2024
Research Assistant
 - Designed and implemented RESTful API for Sotopia platform using LangChain, facilitating AI agent benchmarking.
 - Conducted experiments to quantify social intelligence of AI agents, leveraging OpenAI API for data collection and analysis.
- **Jabil** San Jose, CA
Sep. 2022 - Apr. 2023
Software Engineer
 - Developed full-stack React/Django web app for inventory and test flow management, improving efficiency by 30%.
 - Created automated testing scripts in C++ and Python, reducing manual testing time by 50%.
 - Led testing team, established Agile methodologies, and liaised with client project managers.

PROJECTS

- **Poker AI Competition:** Founded and Led Cloud-Native AI Competition Platform (\$100,000 prize pool)
 - Led team of 10+ developers and built serverless-first AWS infrastructure, scaling to 500+ users and 10,000+ daily matches.
 - Implemented full-stack platform using Remix, Prisma, and Clerk, with match spectating and notifications through SSE.
 - Built event-driven architecture with SNS/SQS/EventBridge for match orchestration and automated user notifications.
 - Engineered developer-friendly bot submission system with integrated testing and logging, improving participant experience.
 - Managed all competition logistics including documentation, support, and participant engagement for 500+ users.
- **Twitter PageRank:** Distributed Graph Processing System
 - Implemented PageRank algorithm in Scala Spark, processing graph with 1M+ vertices and 500M+ edges.
 - Optimized data movement and storage patterns, improving performance by 70% through caching and shuffle reduction.
- **GPT-2 Training Optimization:** Large Language Model Training Pipeline
 - Implemented systematic hyperparameter optimization strategy for GPT-2 (31M params) training on 10B tokens.
 - Engineered distributed training infrastructure, achieving 0.33 HellaSwag score through automated experimentation.
- **MuZero Implementation:** Deep Reinforcement Learning System
 - Implemented MuZero algorithm with MCTS (Monte Carlo Tree Search), achieving optimal performance on RL tasks.
 - Optimized training through GPU acceleration and tensor management, improving performance by 30% over baseline.

TECHNICAL SKILLS

- **Languages:** Python, Typescript/Javascript, Scala, Java, SQL, C/C++, HTML/CSS
- **Frameworks & Libraries:** React, Remix, Node.js, Django, Spark, PySpark, PyTorch, LangChain
- **Tools & Platforms:** AWS, Docker, Kubernetes, Terraform, Git, GCP, Linux, Databricks, Hadoop