VICTOR MAY

Brooklyn, NY

mrcabbage972.github.io

SUMMARY

I'm an engineer with end-to-end experience in building ML-centric products—from rapid prototyping to production-grade systems and low-level optimization. I've led teams in both industry and the open-source community, and have worked at companies ranging from early-stage startups to Google. Recently, I've developed a deeper focus on research and am actively building my publication record.

WORK EXPERIENCE

- **Staff ML Engineer** New York City
  Google *June 2024-Now*
  - Working on SWE agent quality and evaluations for Gemini CLI.
  - Led a project to benchmark Java Migration capabilities of SWE agents. The resulting work had been submitted to a ML conference.

- **Advisory Board Member** New York City
  Ontocord.ai *January 2024-December 2024*
  - Advising on tech roadmap, team building and fundraising.
  - Led on the Aurora-M model training project and paper (see Publications section).
  - Leading on a project for an upcoming permissively licensed dataset for LLM pretraining.

- **ML Engineering Manager** New York City
  Chegg *June. 2023 - May 2024*
  - Established and led a new cross-functional team that was researching using Vision-Language Models (VLMs) for Q&A in the multi-modal setting.
  - Responsible for the planning of quarterly goals, resources and high-level technical roadmaps.
  - People manager of 5 MLE's, responsible for career growth, performance reviews and compensation.

- **Staff ML Engineer** New York City
  Chegg *May 2022 - June 2023*
  - Initiated and led an effort to fine-tune domain-specific LLMs using Chegg's proprietary data, building out data preparation pipelines, training infra and evaluation processes.

- **Senior ML Engineer, NLP** Tel-Aviv, New York City
  Taboola *Mar. 2020 - April 2022*
  - End-to-end development of a system for extracting high-level concepts from multilingual short texts. The project included research and prototyping to meet the desired quality goals, building multiple data pipelines and micro-services and optimizing GPU utilization via CUDA optimization. This system and also the one described below are processing millions of texts per day in production.
  - Designed and implemented a data pipeline that orchestrates all of Taboola's NLP processing.
  - Developed a proof-of-concept of a highly accurate multilingual Named Entity Recognition model using stacked PEFT adapters for transformers and multi-task learning. Demonstrated that the NER task can be decoupled from language, enabling improved performance in long-tail languages using unlabeled data and multilingual transfer.

- **Senior ML Engineer, Recommender Systems** Tel-Aviv, New York City
  Taboola *Mar. 2015 - March 2020*

- Designed and implemented Taboola's first real-time personalized recommender system for ad ranking, including the MLOps infrastructure and the first models running on production traffic. The system was rolled out on all of Taboola's traffic, serving around 20B pageviews per day.
- Built a system for driving the subscription rates of web publishers by a modeling-based personalization of a user's browsing experience, using recommenders and multi-arm bandits. The system was shipped and a dozen B2B customers were onboarded.

- **Computer Vision Algorithm Developer** Tel-Aviv
  Extreme Reality 3D *Apr. 2012 - Mar. 2015*
  - Developed algorithms for real-time human pose estimation and gesture recognition in video. Our technology had been licensed by the likes of Samsung and SEGA games.

PUBLICATIONS

- FreshBrew: A Benchmark for Evaluating AI Agents on Java Code Migration, Submitted (2025)

- GitChameleon: Evaluating AI Code Generation Against Python Library Version Incompatibilities, Preprint (2025)

- Aurora-M: The First Open Source Multilingual Language Model Red-teamed according to the U.S. Executive Order, Accepted to COLING 2025 Industry Track.

- An algorithm for improving Non-Local Means operators via low-rank approximation, Accepted to IEEE Transactions on Image Processing (Volume:25 , Issue: 3).

EDUCATION

- **M.Sc., Applied Mathematics, Magna Cum Laude**
  Tel-Aviv University

- **B.Sc., Computer Sciences and Mathematics**
  Bar-Ilan University

COMMUNITY INVOLVEMENT

- Contributed to developing the LAION OpenAssistant model.

- 2x Kaggle Expert, winner of 3 medals in NLP competitions.