

# Konstantinos Kastritis

Data Scientist | ML Engineer

(905) 960-1970 · [konst.kastritis@gmail.com](mailto:konst.kastritis@gmail.com)

<https://kastritis.dev> · <https://github.com/kkastr> · <https://linkedin.com/in/konstantinos-kastritis/>

## PROFILE

Data scientist and AI engineer building LLM-based applications, predictive models, and data-intensive simulations. Strong in Python, scientific computing, and machine learning with practical experience designing end-to-end data pipelines, model development, and deployment.

## SKILLS

Programming: Python, TypeScript, C, C++, CUDA

ML & Data: NumPy, Pandas, SciPy, SQL, scikit-learn, PyTorch, XGBoost, Model evaluation, Matplotlib, Plotly

LLM & GenAI: Context and prompt engineering (system instructions, JSON-structured outputs), multi-model orchestration, agentic workflows, OpenAI/Gemini/HuggingFace/Groq APIs

Data Engineering: FastAPI, ETL pipelines, Data visualization, APIs, Websockets

DevOps: Git, Linux, Docker, CI/CD

Frontend: React, Tailwind

## EXPERIENCE

### Creator / Lead Developer

[Synecho](#)

Jun 2025 – Present

- Launched a SaaS that analyzes variation in LLM outputs so users can track shifts in brand/topic perception across repeated prompts.
- Built a React/Tailwind UI with a prompt workflow and dashboard for statistics, judge summaries, and per-request prompt–response transcripts.
- Implemented a Python/FastAPI backend with SQLite persistence, coordinating async calls to Gemini, Groq, and OpenAI and aggregating outputs into quantitative metrics (similarity, sentiment, term/phrase statistics).
- Orchestrated a multi-step “LLM-as-judge” workflow (candidate models + judge) with JSON-structured outputs to score, summarize, and constrain hallucinations in responses.
- Reduced perceived latency with concurrent LLM calls and WebSocket streaming; containerized with Docker and deployed on a VPS behind a reverse proxy with CI/CD via GitHub Actions.

### Researcher

Ontario Tech University

Jun 2024 – Dec 2024

- Extended HOOMD-blue physics simulation software (Python/C++/CUDA) with custom particle interactions and tests to validate GPU-accelerated simulations.
- Analyzed simulation and experimental datasets with Python (NumPy/Pandas), generating plots and metrics for publication.
- Documented the HOOMD-blue modifications and build process so collaborators could reproduce changes.
- Contributed to research discussions, proposing data-driven approaches to modeling and analysis.

## Data Scientist

Self-employed

August 2022 – May 2024

- Shipped independent ML projects, including polymer conformation generation and Reddit sentiment analysis.
- Explored ML and data science techniques on real datasets beyond formal roles to deepen practical skills

## Researcher

Ontario Tech University

Sep 2021 – Jun 2022

- Built data pipelines in Python to ingest, clean, and store simulation outputs, improving storage efficiency.
- Defined metrics to quantify experimental/simulation behavior, reducing data collection overhead.
- Enhanced simulation tooling to automate batch runs and post-processing, enabling more complex studies.
- Provided technical guidance to graduate students on data analysis and scripting.

## Graduate Research & Teaching Assistant

Ontario Tech University

Jan 2015 – Dec 2020

- Developed C++/CUDA APIs for GPU-capable simulation software, increasing performance by ~50% and enabling faster simulation studies.
- Provided data-driven and theoretical insights to experimental collaborators and co-authored two peer-reviewed publications based on combined simulation and experimental results.
- Supported group migration to GPU computing; wrote documentation and trained six students in custom simulation tools.
- Led labs and tutorials for physics and computer science courses teaching ~50 students per term core skills in Physics, Linux, Python, and data visualization.

## SELECTED PROJECTS

**Polymer Conformation Generation** (<https://alysides.kastritis.dev>) | PyTorch, Transformers, FastAPI, React

- Trained a transformer model on MD simulation data of coarse-grained semiflexible polymers to learn the equilibrium distribution of bond vectors and chain conformations.
- Used the model to sample near-equilibrated chains as MD starting states, reducing the need for long thermalization runs whose relaxation time grows at least quadratically with chain length.
- Built an interactive FastAPI + React web app where users select chain length (within limits for performance), generate sample conformations on demand, and inspect basic configuration statistics.

**Social Media Sentiment Analysis** | Python, SQL, scikit-learn, Plotly

- Scraped 100+ posts and 25000+ comments from five subreddits.
- Trained a scikit-learn classifier for sentiment analysis, achieving 86% accuracy on held-out data.
- Built a dashboard with six key metrics (sentiment by subreddit, comment volume over time, etc.) to explore Reddit user behavior.

## EDUCATION

**Ontario Tech University – M.Sc., Materials Science** | Aug 2019

Thesis: Computational Studies of Semiflexible Polymer Dynamics Under Confinement

**Ontario Tech University – B.Sc., Physics** | Dec 2016

Thesis: N-Body Simulations of Dark Matter Halos

## AWARDS

NSERC Undergraduate Research Award

## LANGUAGES

English (fluent), Greek (fluent)