Dennis Johan Loevlie

Multimodal foundation models for tabular data · probabilistic deep learning

+1 724 841 8769 · dennis.loevlie@tufts.edu · dennisloevlie.com · Google Scholar · github.com/loevlie

EDUCATION

Centrum Wiskunde & Informatica & University of Amsterdam — Incoming, Aug 2026 –

PhD in Computer Science. **Advised by Madelon Hulsebos (TRLab, CWI) and Jan-Willem van de Meent (AMLab, UvA); ELLIS PhD Program international co-advisor TBD.** Focus: multimodal tabular foundation models; probabilistic deep learning.

Tufts University, Medford, MA — Sep 2024 – May 2026 (expected)

M.S. in Computer Science. GPA: 3.95 / 4.00. Transcript.

Carnegie Mellon University, Pittsburgh, PA — Sep 2019 – Dec 2020

M.S. in Chemical Engineering. GPA: 3.91 / 4.00.

West Virginia University, Morgantown, WV — Sep 2016 – Aug 2019

B.S. in Chemical Engineering, *cum laude*, with Honors.

PUBLICATIONS

Machine Learning

E. Harvey*, **D. J. Loevlie***, A. A. Satani, W. Chen, D. M. Kent, and M. C. Hughes. "A Multi-Dataset Benchmark of Multiple Instance Learning for 3D Neuroimage Classification". In: *Conference on Health, Inference, and Learning (CHIL) 2026*. *Equal contribution. [code]. 2026. Forthcoming.

E. Harvey, **D. J. Loevlie**, and M. C. Hughes. "Synthetic Data Reveals Generalization Gaps in Correlated Multiple Instance Learning". In: *ML4H 2025 Symposium (Findings Track)*. 2025.

Chemistry & Computational Science

D. J. Loevlie, B. Ferreira, and G. Mpourmpakis. "Demystifying the Chemical Ordering of Multimetallic Nanoparticles". In: *Accounts of Chemical Research* 56.3 (2023). [code], pp. 248–257.

M. Salem, **D. J. Loevlie**, and G. Mpourmpakis. "Single Atom Alloys Segregation in the Presence of Ligands". In: *Journal of Physical Chemistry C* 127.46 (2023), pp. 22790–22798.

R. Ding, I. M. Padilla Espinosa, **D. Loevlie**, S. Azadehranjbar, A. J. Baker, G. Mpourmpakis, A. Martini, and T. D. B. Jacobs. "Size-Dependent Shape Distributions of Platinum Nanoparticles". In: *Nanoscale Advances* 4 (2022), pp. 3978–3986.

A. V. Nagarajan, **D. J. Loevlie**, M. J. Cowan, and G. Mpourmpakis. "Resolving Electrocatalytic Imprecision in Atomically Precise Metal Nanoclusters". In: *Current Opinion in Chemical Engineering* 36 (2022), p. 100784.

RESEARCH EXPERIENCE

Tufts University — Hughes Lab — Aug 2024 – Present

Graduate Researcher with Prof. Michael Hughes

- Co-led (equal contribution) the CHIL 2026 multi-dataset benchmark of multiple instance learning (MIL) for 3D neuroimage classification across 7 MRI/CT datasets and 9 tasks; showed a simple mean-pooling MIL baseline matches or outperforms recent attention-based MIL and 3D CNN methods on 4 of 6 moderate-sized tasks while being **25× faster to train**.
- Contributed to the ML4H 2025 study on a *generalization gap* between correlated and non-correlated MIL architectures: implemented SmAP (the newest NeurIPS MIL method) into the lab codebase, reduced a training-example bottleneck from *quadratic to linear scaling* to unlock experiments on a substantially larger semi-synthetic dataset, and helped design the synthetic benchmark with a Bayes-optimal upper bound.
- Designed attention- and transformer-based MIL pipelines on 3D MRI/CT for early indicators of dementia and stroke; trained on Tufts HPC with multi-GPU distributed setups.
- Developing a regularization method to encourage interpretable attention scores, and methods for supervised learning under noisy labels.

Tufts University — Sinapov Lab

Jan 2025 – Present

Graduate Researcher with Prof. Jivko Sinapov

- Trained Qwen2.5-Coder-7B with Group Relative Policy Optimization (GRPO) and a custom 3-part reward (structure, aesthetics, semantic alignment) to generate SVGs from text; achieved an **18% improvement** on a benchmark covering aesthetics, alignment, and code validity.
- Investigating why vision-language models (Qwen3-VL) and open-vocabulary segmentation models (SAM 3) fail at object counting in dense scenes; designing tool-use and trajectory-based methods drawing on cognitive-science accounts of human counting.

University of Pittsburgh — CANELa Lab

Jun 2021 – Jan 2023

Graduate Researcher with Prof. Giannis Mpourmpakis

- Proposed a weight-initialization method that delivered a **71% RMSE reduction** for predicting nanoparticle stability across compositions and atomic orderings (*Loevlie et al., Acc. Chem. Res. 2023*).
- Ensured fair evaluation of ML architectures for material-property prediction; co-author on *Salem et al., JPCC 2023*.
- Wrote the ML methods and background section for the *Nagarajan et al.* review (Curr. Opin. Chem. Eng., 2022).

Carnegie Mellon University — Kitchin Group

Dec 2019 – Dec 2020

Graduate Researcher with Prof. John Kitchin

- Trained a CNN classifier to extract experimental information from microscopy images; rebuilt Mathematica analysis tooling as fast, interactive Python.
- Released `nb_search`, an open-source Python package for searching and indexing across Jupyter notebooks.

AWARDS & HONORS

2026 **1st place** (10 teams), *BrainStorm Neural Decoder Challenge* — real-time auditory decoding from 1024-channel ECoG; 95% accuracy, sub-ms inference, edge-deployable model.

2024 **Community Grant**, *Hugging Face* — for Depth Anything demonstration work on video.

2019 **1st place** (Advanced category), *AVEVA National Simulation Competition*.

2018 **2nd place**, *AIChE National Poster Competition*, Computing & Process Control Division.

INDUSTRY EXPERIENCE

Massachusetts General Hospital & Harvard Medical School

2025 – Present

Machine Learning Researcher — computational pathology

- **Systems:** building segmentation models for colon-cancer pathology slides at gigapixel resolution; pipeline integrates with clinical workflows for downstream evaluation.
- **Research:** dataset-construction methodology and collaboration with pathologists to ground model outputs in expert-validated labels.

KEF Robotics, Pittsburgh, PA

Jan 2023 – Aug 2024

Senior Computer Vision & ML Engineer (2024) · Computer Vision Engineer (2023)

- **Systems:** sole ML engineer on a one-year, $500K project; led a team of five shipping on-device object detection, monocular depth, and 3D map generation; deployed on flight hardware via TensorRT; **45% inference-speed gain** at $<1\%$ accuracy cost. Two in-person demos, selected for follow-on funding.
- **Research:** fine-tuned Mask2Former for power-line segmentation across RGB and IR modalities via transfer learning and cross-modal label transfer.

ContentsPal

May 2025 – Aug 2025

Multimodal AI Intern (insurance-tech startup, MIT-led)

- **Systems:** integrated and tested learning-based duplicate detection and open-vocabulary instance segmentation in React / React Native.
- **Research:** prototyped retrieval and matching components on multimodal product data.

AiThElite Dec 2020 – May 2021

Lead Data Scientist (collegiate-athletics AI startup)

- Built end-to-end data pipeline (Selenium scrapers → NumPy/Pandas feature engineering → scikit-learn models) and shipped a Django/AWS-hosted prediction product.

OPEN SOURCE

- `huggingface/transformers` — contributor; merged PRs on model implementations and bug fixes
- `Franblueee/torchmil` — core contributor; multiple merged PRs across the multiple-instance-learning library
- `loevlie/neuropt` — author; LLM-guided neural architecture search and hyperparameter tuning
- `JohanDL/Depth-Anything-Video` — author; Hugging Face Community Grant recipient

SELECTED TALKS

Synthetic Data Reveals Generalization Gaps in Correlated Multiple Instance Learning. Poster, ML4H 2025 Symposium, San Diego, CA. *Dec 2025.*

Computer Vision for Unmanned Aerial Vehicles. Invited talk, XChangeIdeas Pittsburgh. *2023.*

SERVICE & OUTREACH

Industry volunteer, youth robotics team building accessibility tools for blind soccer players 2024

TECHNICAL SKILLS

Languages Python · C++ · JavaScript · MATLAB
ML / AI PyTorch · JAX · Hugging Face (Transformers, Datasets, TRL, PEFT) · GRPO / DPO · multimodal learning · RL post-training · multiple instance learning · computer vision
Engineering FSDP · DeepSpeed · vLLM · FlashAttention · Triton · CUDA · TensorRT · SLURM / HPC · multi-GPU distributed training · Docker · AWS