alexloftus2004@gmail.com
# Alexander Loftus
alex-loftus loftusa alex-loftus.com

I strive to develop new machine learning and algorithmic techniques that help humans understand and control LLMs in more meaningful, powerful ways. In the past, I have worked on representation learning, scalable systems, volumetric segmentation, and network statistics. Some highlights:

**Textbook author**: Publishing contract with Cambridge University Press.

**1st place ranking, $100,000 Machine Learning competition**: Work featured on cover of Scientific American. Competed against 1249 teams. Vesuvius scroll ink detection.

**Publications in top conferences**: Best poster award at NeurIPS 2023 LatinX workshop, first author work in ICLR 2024

**Open-source contributions**: Primary contributor to microsoft network statistics package graspologic.

**Teaching and Leadership**: Led a team of three to develop an object detection augmentation algorithm; a team of five to contribute to a brain network estimation pipeline; assistant director managing 8-12 instructors.

## Education

**Northeastern University** — Boston, MA
*PhD* Computer Science — 2024-
*Advisor*: Dr. David Bau

Interpretability, evaluations, and training dynamics in Code LLMs.

**Johns Hopkins University** — Baltimore, MD
*MSE* Biomedical Engineering: Machine Learning & Data Science Focus — 2020-2022
*Advisor*: Dr. Joshua Vogelstein
*Thesis*: Hands-On Network Machine Learning

dean's list, highest honors, GPA 4.0/4.0.

**Western Washington University** — Bellingham, WA
*BS* Behavioral Neuroscience — *Minors*: Chemistry, Philosophy — 2014-2018
*Founder & President*, Computational Neuroscience Club
*Vice President*, Neuroscience Club

Built computational neuroscience club from scratch, taught weekly seminars.

## Textbook

**Hands-on Network Machine Learning**: *Eric Bridgeford,* ***Alexander R. Loftus****, Joshua Vogelstein.* Cambridge University Press, in copy-editing phase. 2025.

Spectral representation theory on networks. 530 pages, 147 figures.

## Publications

* indicates equal contribution.

**NNsight and NDIF: Democratizing Access to Open-Weight Foundation Model Internals**: ***A.R. Loftus****, *J.Fiotto-Kaufman**, et al.* ICLR 2025.

Easily explore and manipulate foundation model internals with no engineering overhead.

**A Saliency-based Clustering Framework for Identifying Aberrant Predictions**: *A. Tersol Montserrat,* ***A.R. Loftus****, Y. Daihes.* Paper, NeurIPS LatinX AI Workshop, 2023. **Won best poster.**

Use embeddings of saliency map crops to identify predictions caused by spurious features.

**A low-resource reliable pipeline to democratize multi-modal connectome estimation and analysis**: *J. Chung, R. Lawrence,* ***A.R. Loftus****, et al.* Paper, in review at Nature Methods, 2024

Turn diffusion MRI scans into adjacency matrices. Code on github.

**Role of CAMKII in Associative Conditioning and GLR-1 Expression in C. Elegans**: *M. Pribic,* ***A.R. Loftus****, et al.* Poster, Society for Neuroscience, 2017.

Removing a protein involved in learning blocks associative conditioning in worms.

## Talks

**State of the Art in Knowledge Editing**: ***A.R. Loftus***, 2023
Current techniques in knowledge localization and editing in LLMs and diffusion models.

**1st Place Solution - Vesuvius Ink Competition**: *R. Chesler,* ***A.R. Loftus****, A. Tersol Montserrat, T. Kyi*, 2023
Presenting on our winning solution to a $100,000 Kaggle competition, part of the $1,000,000 Vesuvius competition.

**ICML Conference Highlights**: ***A.R. Loftus***, 2023
Machine learning techniques in drug discovery and medicine at ICML 2023.

**Working with LLMs**: ***A.R. Loftus***, 2023.
Introduction to LLM engineering. Talk given to 100 people at the AI/ML San Diego meetup.

**Linear Algebra, from Dot Products to Neural Networks**: ***A.R. Loftus***, 2023.
Created a YouTube tutorial series on the fundamentals of linear algebra for machine learning.

**Effects of an unc-43 (CaMKII) Gene Deletion on Short-Term Memory for Associative Conditioning in C. elegans**: ***A.R. Loftus***, Psychfest 2017.
Mechanistic understanding of roundworm neural circuitry.

## Fellowships & Awards

**First Place Winner** 2023
Kaggle Vesuvius Competition, $100,000.

**Khoury Distinguished Fellowship** 2024
Northeastern University PhD fellowship.

**Best Poster Award** 2023
NeurIPS 2023 LatinX AI Workshop.

**MIT EECS GAAP** 2023
MIT mentorship program.

**AWS Research Grant** 2019
$10,000 grant for computational research on cloud services.

## Experience

**Data Scientist** San Diego, CA
Creyon Bio 2022-2024

*ESP Embeddings*: Developed constrastive feature representation learning approach for electrostatic potential data. Resulted in AUC improvement in downstream classification accuracy from 78 to 89.
*Neuron Toxicity Detection*: Built deconvolution and segmentation pipeline to detect toxicity in neurons.

**Machine Learning Research Engineer** Rockville, MD
Blue Halo 2021-2022

*Conditional Image Generation with Generative Adversarial Networks*: Synthetic data augmentation. Led the switch to diffusion-based methods over GANs.
*Detecting Objects with Enhanced Yolo and Knowledge Graphs*: Used knowledge graphs to enhance object detection on videos.
*Geometric Multi-Resolution Analysis*: Built out infrastructure for a hierarchical clustering method that used wavelets and signal processing techniques.

**Research Assistant** Baltimore, MD
Johns Hopkins University — Dr. Joshua Vogelstein 2018-2021

*Network Machine Learning*: Publishing contracts offered by both Springer Publishing and Cambridge University Press.
*Open-Source Contributor to Microsoft network ML package Graspologic*: Built dimensionality reduction models on networks.
*Primary maintainer & Developer of brain network estimation pipeline*: Diffusion MRI to graphs pipeline. AWS cloud-computing integration with pytest CI/CD infrastructure. Eliminated 1000 lines of code and halved computation time.

**Assistant Director** Seattle, WA
iD Tech Camps — University of Washington 2014-2018 summers

*Leader and Manager*: Administrator for a STEM education camp which taught C++, Python, Java, game design, and robotics at the University of Washington. Managed 8-12 instructors with 80-120 students per week.

**Research Assistant** Bellingham, WA
Western Washington University 2015-2018

*Associative learning in C. elegans*: Python automation pipeline cut 5 days of work down to minutes. Resulted in research presented at the Society for Neuroscience, 2017.

## Advisory

**Consultant** — Krnel.ai
Advisory role for cybersecurity-based interpretability startup. — Spring 2025

**Advisor** — Harvard University
Interpretability in Large Language Models with the Harvard Student AI Safety Team. — Spring 2025

## Teaching

**Head Teaching Assistant** — Johns Hopkins University
Foundations of Computational Biology and Bioinformatics, *EN.BME.410/634* — Spring 2021

**Teaching Assistant** — Johns Hopkins University
NeuroData Design II, *EN.BME.438/638* — Spring 2020

**Teaching Assistant** — Johns Hopkins University
NeuroData Design I, *EN.BME.437/637* — Fall 2019

**Teaching Assistant** — Western Washington University
Introduction to Behavioral Neuroscience, *PSY.220* — Winter 2017

**Curriculum Designer** — iD Tech Camps
Built curriculum used across 50 locations in the United States by tens of thousands of students. — Spring 2017

**Instructor** — iD Tech Camps
Taught programming and game design to high school students. — 2014-2018 summers

## Skills Summary

**Languages**: Python, R, Rust, Bash, CSS, Mojo, English, Broken Spanish

**Tools & Frameworks**: pytorch, pytorch-lightning, tensorflow, jax, numpy, scipy, pandas, polars, sklearn, seaborn, matplotlib, docker, AWS, google cloud (GCP), photoshop, SQL, weights & biases, mlflow, kubernetes, linux

**Areas of Expertise**: Linear algebra, probability & statistics, deep learning, information theory, transformers, diffusion models, convolutional autoencoders, GPUs and CUDA, public speaking, leadership & management, teaching, natural language processing, computer vision

## Fun

**Gaming**: Starcraft 2 grandmaster in high school, competed and won in Seattle-area tournaments.

**Music**: Fingerstyle guitarist. Played at open mic nights.

**Dancing**: Partner dance instructor and competition winner. Fusion, West Coast Swing, Zouk, Salsa, Bachatta.