Alex Lavaee

alexlavaee.me | github.com/lavaman131 | alexlavaee | alavaee@bu.edu

EDUCATION

Boston University | B.S. Data Science   *May 2025*

- **Cumulative GPA: 3.95**; *summa cum laude*

RESEARCH EXPERIENCE

Kirchhausen Lab   Aug 2024 - June 2025
Advisor: Tomas Kirchhausen (Harvard Medical School)   *Boston, MA*

- Researched and developed 3D computer vision self-supervised representation learning approach for understanding the molecular and cellular world with downstream tasks of object detection and tracking for automatic processing of large scale molecular movies.

- Managed and utilized multi-node Deep Learning cluster for distributed training (DDP, FSDP) in PyTorch. Pinpointed I/O bottlenecks in data loading and devised methods to speed up data loading by 20x.

- Created scalable data pipeline for terabytes of dense 3D volumes of molecular movies and implemented web dataset for efficient data streaming and preprocessing.

Image and Video Computing (IVC) Group   Sept 2023 - June 2025
Advisor: Margrit Betke (Boston University)   *Boston, MA*

- Researched 3D pose estimation through vision and language models to represent human motion and improve our understanding of motion for many things including physical therapy treatment methods.

- Engineered MotionGPTv1.5, using LLMs to effectively represent motion-aware descriptions of exercises and represent exercises as a collection of smaller actions and by devising an LSTM architecture for motion duration prediction (regression task).

- Developed a 3D human-joint-based self-supervised representation learning approach, Motion-JEPA, building on top of V-JEPA (Video Joint Embedding Predictive Architecture) to improve human motion understanding.

Responsible AI, Law, Ethics & Society   Jan 2024 - Apr 2024
Advisor: Shlomi Hod (Boston University), Talia Schwartz (UC Berkeley)   *Boston, MA*

- Developed case-study to analyze the impact of annotator demographics on the performance of machine learning models.

- Fine-tuned pretrained Distil-Bert model to predict if text is offensive for, measuring train and validation performance for different annotator demographics.

- Dataset bias impact: Models trained on white male data demonstrate superior performance compared to those trained on white female or black female data, explaining the poor performance of models trained on minority racial groups, as the dataset is skewed towards white individuals.

Themis AI   Aug 2023 - Feb 2024
Advisor: Alexander Amini (Themis AI, MIT CSAIL)   *Cambridge, MA*

- Developed a transformer-based Deep Learning algorithm for 3D object detection and instance segmentation of point clouds from 3D LIDAR scans of autonomous vehicle terrains and buildings.

- Devised method to measure uncertainty in model predictions including aleatoric, epistemic, and vacuity uncertainty.

- Created scalable Python pipeline for collecting, preprocessing, and compressing large databases (2TB+) of 3D LIDAR point cloud scans of autonomous vehicle terrains and buildings leading to 100x reduction in storage space (20GB) of data.

Przytycki Lab Dec 2022 - May 2023
Advisor: Pawel Przytycki (Boston University) *Boston, MA*

- Analyzed extensive multimodal datasets, uncovering key insights into the progression and functioning of human diseases.
- Engineered Python algorithms and scripts for large-scale genomic data wrangling, enhancing data interpretability and processing speed.

Edwards Lifesciences June 2022 - Aug 2022
Advisor: Cristhian Potes *Irvine, CA*

- Pioneered the development of a Deep Learning architecture that predicts the presence of valvular heart disease from raw arterial waveform data built on top of TensorFlow, achieving a 15% improvement in precision and recall compared to traditional machine learning and statistical models.
- First Deep Learning algorithm in the Applied Machine Learning Group at Edwards Lifesciences to be submitted to the FDA for clearance.
- Algorithm is currently integrated as part of Edwards Lifesciences' medical devices in clinical trials.

TEACHING

Introduction to Deep Learning (6.S191) Jan 2024 - Jan 2025
∼ 200 students, MIT *Cambridge, MA*

- Helped teach concepts including multi-layer perceptrons, RNNs, transformers, auto-encoders, VAEs, GANs, deep reinforcement learning, and stable diffusion.
- For a detailed syllabus see: http://introtodeeplearning.com

Programming for Data Science (DS210) Jan 2023 - Feb 2024
∼ 100 students, Boston University *Boston, MA*

- Helped teach students machine learning with Python and data structures and algorithms with Rust.
- For a detailed syllabus see: https://onak.pl/teaching/ds210-spring_2022.php

WORK EXPERIENCE

Microsoft July 2025 - Present
Software Engineer *Redmond, WA*

- Building AI for Windows.

Microsoft May 2024 - Aug 2024
Software Engineering Intern *Redmond, WA*

- Developed AI solution for Windows to improve feature error identification from 40 days to 40 minutes on average and reducing possible error feature candidates by 95%, boosting developer productivity.
- Created scalable distributed data pipeline for ingesting and processing petabytes built on top of Azure.
- Facilitated team's migration to Azure Data Factory for running ETL jobs.

Edwards Lifesciences June 2023 - Aug 2023
Software Engineering Intern *Irvine, CA*

- Developed a full-stack medically-aware transcription application with React in the frontend, FastAPI in the backend, and Docker for DevOps, combining OpenAI's Whisper speech-to-text model with prompt engineering, improving transcription accuracy by 10% and processing speed by 2x compared to Microsoft Teams' proprietary transcription service.

- Engineered an in-house private full-stack medical chatbot Generative AI (LLM) solution with Python's Streamlit library in the frontend and a fine-tuned version of Meta AI's Llama 2 model from Hugging Face for MLOps, with conversation history support and document-based question answering using a custom local vector database, providing scalable infrastructure for 50+ users.

Boeing   June 2021 - Aug 2021
Software Engineering Intern   *Huntington Beach, CA*

- Created a risk assessment systems engineering pipeline in PyTorch for natural disaster relief, utilizing geospatial data to assess building damage, which reduced estimated search and rescue time of natural disasters by 50%.

- Led an integrated product team of 8 members for DonkeyCar, an autonomous car (Computer Vision) simulation project, built on top of TensorFlow.

PUBLICATIONS

- **Lavaee, A.**, 2024. Sketch 'n Solve: An Efficient Python Package for Large-Scale Least Squares Using Randomized Numerical Linear Algebra. arXiv preprint arXiv:2409.14309

- Yang, Q., Ravikumar, S., Schmitt-Ulms, F., Lolla, S., Demir, E., Elistratov, I., **Lavaee, A.**, Lolla, S., Ahmadi, E., Rus, D. and Amini, A., 2023. Uncertainty-aware Language Modeling for Selective Question Answering. In 2024 Responsible Language Models (ReLM).

RECENT AWARDS

- **Academic Excellence Award**   2025
  One of three recipients of academic excellence for undergraduate career at Boston University.

- **Barry Goldwater Scholarship** – Honorable Mention   2024
  First second-year nominee from Boston University in four years.

- **MinneMUDAC Data Science Competition** – Second Place   2023
  Developed predictive algorithm for baseball attendance prediction and presented to Minnesota Twins MLB team.

CLUBS/ORGANIZATIONS

BU Spark!   Oct 2022 - May 2024
Boston University   *Boston, MA*

- Doubled the reach of the BU Spark! Learning Ambassador "Micro Challenges" project, a learning by doing initiative that allows students to follow technical CS mini-projects that can be completed within a relatively short time frame, designed to be innovative and foster learning, by facilitating 10+ student-led projects.

- Migrated project infrastructure to GitHub and introduced a Notion-based workflow, streamlining collaboration and increasing team productivity by 30%, while ensuring seamless access to resources and documentation for all team members.

- Collaborated in a team to develop "BullyProof," a full-stack React and Python Chrome extension, utilizing Natural Language Understanding technology with BERT model on Hugging Face to filter out hate speech and discriminatory content on Twitter, resulting in a 50% reduction of such content.

- Utilized core principles in project management and development including user-centered design, agile methodologies, user interviews, and prototyping techniques, resulting in a fully functional prototype in 4 months.

Relevant Coursework

Randomized Algorithms	Machine Learning	Linear Algebra
Natural Language Processing	Data Engineering	Multivariable Calculus
Computer Vision	Applied Statistics	Data Structures and Algorithms
Cloud Computing	Bayesian Statistics	

Skills

- Programming Languages: Python, JavaScript, HTML, CSS, Rust, C++, Java, Bash
- Frontend/Backend Frameworks: React, Tailwind, FastAPI, Flask, Node.js
- Tools: PyTorch, TensorFlow, CUDA, Keras, Transformers, Agentic AI (LangChain, AutoGen), Scikit-learn, SciPy, Pandas, NumPy, Matplotlib, Spark, Hadoop, MapReduce
- Databases: SQL (Azure SQL, Oracle SQL), NoSQL (CosmosDB, MongoDB), Vector Database (Pinecone, Chroma)
- DevOps Pipelines: Docker, Azure, AWS, Google Cloud, Git, CI/CD, Linux, Bash scripting