Fadi Atieh

520 N Kingsbury St, Chicago, IL 60654
fadi.atieh98@gmail.com
(413) 801-2922
github.com/fadi987 huggingface.co/fadi77

EDUCATION

Ph.D. in Machine Learning Starting Fall 2026
Carnegie Mellon University, Pittsburgh, PA
Advisor: Max Simchowitz & Nicholas M. Boffi
Funded by U.S. Department of Energy Computational Science Graduate Fellowship (DOE CSGF)

Masters of Engineering (GPA: 5.0/5.0) 09/2020 - 02/2022
Massachusetts Institute of Technology, Cambridge, MA
Thesis: A Novel Statistical Procedure Toward the Discovery of the Higgs Boson

Mathematics & Computer Science (GPA: 5.0/5.0) 09/2016 - 06/2020
Massachusetts Institute of Technology, Cambridge, MA

ADVANCED COURSEWORK

- Measure-Theoretic Probability • Non-Asymptotic Statistics • Information Theory
- Probabilistic Graphical Models • Mathematical Optimization • Theory of Computation
- Statistical Mechanics • Quantum Physics • Chaos Theory • Real Analysis • Distributed Systems

RESEARCH EXPERIENCE

ML Research under Prof. Aditi Krishnapriyan, UC Berkeley 06/2025 – 09/2025
Independent Researcher

- Co-authored work showing Transformers predict molecular energies and forces without graph, symmetry, or physical priors, with errors competitive with SOTA GNNs.
- Showed attention scores decay with inter-atomic distances; defined an "effective attention radius" and showed it adapts to local atomic density, revealing a learned molecular topology.
- Analyzed attention dynamics across layers, revealing a progression from local to global information flow and head-level specialization for short-, medium-, and long-range molecular interactions.
- Integrated the Muon optimizer into distributed training pipeline, improving training stability.

Master's Thesis under Prof. Yury Polyanskiy 09/2020 - 02/2022
Graduate Researcher at Prof. Yury Polyanskiy's Group

- Discovered a novel hypothesis testing methodology toward the discovery of the Higgs boson addressing limitations of Kaggles HiggsML challenge to improve analysis of the data produced by the CMS and ATLAS experiments.
- Trained a neural network classifier to approximate the log likelihood ratio of decay observations coming from Higgs and plugged the approximation into a non-parametric Wilcoxon rank-sum test.
- Backed the methodology by theoretical rigor and experimental verification where I analyzed data from the LHC (Large Hadron Collider) and Monte Carlo simulations.

Photonics & Modern Electro-Magnetics Group under Prof. Marin Soljačić 02/2018 - 12/2018
Undergraduate Researcher at Research Laboratory of Electronics at MIT

- Demonstrated through numerical simulations and mathematical treatment the stochastic convergence of a probabilistic algorithm toward solving the Ising problem leading to a peer-reviewed publication in Nature Communications.
- Implemented fast matrix multiplication - to be used in MCMC algorithms - by programming a FPGA chip using VHDL taking advantage of parallel computation.

Synthetic Neurobiology Group under Prof. Edward Boyden 02/2017 - 09/2017
Undergraduate Researcher at MIT Media Lab

- Analyzed data and constructed testing plasmids for building a molecular biology system that causes neurons to write a history of their activation into nucleic acids in a project called TickerTape.
- Designed and debugged experiments independently to find the right combination of proteins/RNAs with the correct concentrations for the molecular mechanism to work.

PUBLICATIONS

Kreiman, T., Bai, Y., **Atieh, F.**, Weaver, E., Qu, E., Krishnapriyan, A. S. (2025). Transformers Discover Molecular Structure Without Graph Priors. *arXiv:2510.02259*.

Roques-Carmes, C., Shen, Y., Zanoci, C., Prabhu, M., **Atieh, F.**, Jing, L., Dubček, T., Mao, C., R. Johnson, M., Čeprić, V., D. Joannopoulos, J., Englund, D., Soljačić, M. (2020). Heuristic Recurrent Algorithms for Photonic Ising Machines. *Nature Communications*.

WORK EXPERIENCE

Citadel LLC 02/2022 - 09/2024

Quantitative Research Engineer at GQS (Global Quantitative Strategies)

- Developed and deployed quantitative models predicting various market variables working closely with quantitative researchers.
- Built a scalable sampling/fitting/inference pipeline for niche models with unique mathematical formulations that are not amenable to preexisting infrastructure.
- Built a scalable model evaluation framework for analyzing research models with plug-and-play metrics of a user's choice.
- Built infrastructure for serving live predictions of various market variables to be used in live trading.

Citadel LLC 06/2021 - 08/2021

Research Engineering Intern at GQS (Global Quantitative Strategies)

- Optimized the main research file system through changes in file representation and C++ API reducing I/O hits, latency, and storage by an order of a magnitude.
- Analyzed external-source market data for price and volume accuracy reconciling with in-house data.
- Built an automated testing/linting Bazel framework for Python scripts within a large code base.

Citadel LLC 07/2020 - 08/2020

Software Engineering Intern at Commodities

- Built a Flask server used by traders to mark futures curves in commodities and achieved a 100× speed up from the legacy server.
- Exposed C++ instrument valuation functionality into Python using Pybind11 and wrote documentation of the exposed interface.

TEACHING EXPERIENCE

Theory of Computation with Professor Michael Sipser (18.404) Fall 2020 & Fall 2021

Teaching Assistant

- Taught weekly recitations where I prepared 3 to 4 practice problems to collaboratively solve with students and elaborate on intricate concepts presented during lectures.
- Co-piloted lectures on Zoom with professor and other TAs answering live questions for students.
- Gave two sessions of office hours per week to answer questions students had on lecture content or help with homework.
- Received outstanding ratings from students averaging 6.7 and 6.8, in both years respectively, out of 7.

Introduction to Statistical Data Analysis with Professor Yury Polyanskiy (6.401) Spring 2021

Teaching Assistant

- Designed and built a high-end final project drawing from my master's thesis which was included in the class curriculum.
- Gave weekly office hours to supplement material taught that week and answer questions.
- Wrote and graded problem sets, projects, and exams under supervision of professor. Received outstanding ratings from students averaging 6.8 out of 7.

HONORS AND AWARDS

U.S. Department of Energy 04/2026

Computational Science Graduate Fellowship (DOE CSGF)

- Highly selective national fellowship (∼30 awarded annually) supporting doctoral research at the intersection of high-performance computing and science; includes full tuition, a $45,000 annual stipend for up to 4 years, and a research practicum at a DOE national laboratory.

Phi Beta Kappa Honor Society 05/2020

Phi Beta Kappa Member

- Election is conducted by a faculty committee, which reviews the academic records of seniors to find students with superlative records and clear evidence of breadth in the liberal arts and the science fields.

Invited Speaker, MacVicar Day 2019 03/2019

The Educated Student: Thinking and Doing for the 21st Century

- Selected as one of eight speakers for MIT's annual MacVicar Day symposium; talk available on YouTube.

PROJECTS

Arabic PL-BERT: Phoneme-level BERT Training Pipeline 02/2025 – 04/2025

Personal Project

- Trained a phoneme-level BERT on Arabic Wikipedia (1.2M pretrain / 200K finetune articles) with normalization, diacritization (Catt), and phonemization.
- Released weights + code on Hugging Face for downstream ASR/TTS.

Arabic TTS: Dataset Construction & StyleTTS2 Fine-tuning 02/2025 – 04/2025

Personal Project

- Curated a 100+ hr Arabic TTS corpus (denoising, diarization, Whisper transcripts); fine-tuned StyleTTS2 on 22 hr audiobook subset using PL-BERT features.
- Adapted training for Arabic (e.g., removed English-specific WavLM loss); open-sourced model + demo on Hugging Face.

Neural Network Library in Rust 04/2024

Personal Project

- Built a neural network library in Rust with a Python interface for training and inference.
- Designed a clean API and demonstrated functionality on synthetic data in Jupyter.

Tic-Tac-Toe MCTS (Monte Carlo Tree Search) Game Agent in Rust 03/2024

Personal Project

- Implemented a tic-tac-toe agent using Monte Carlo Tree Search (MCTS) in Rust.
- Learned Rust during the project and built a CLI for interactive gameplay.

MATHEMATICS OLYMPIAD

- Honorable Mention in the Asian Pacific Mathematics Olympiad XXV & XXVI. 2013 & 2014
- Bronze Medal in the Syrian National Math Olympiad Finals. 01/2013