

Matt Zhang

set.stun@gmail.com
github.com/MattUnderscoreZhang
linkedin.com/in/matt-zh

ML researcher with projects in physics, game-playing agents, cybersecurity, and text analysis

Education

PhD, Physics | University of Illinois at Urbana Champaign | 2013 – 2020

B.S., Physics & Mech. Eng. | University of Texas at Austin | 2008 – 2013

Overview

Programming languages – Python (13 years experience), C++, Java, SQL/PostgreSQL, Dart/Flutter, Verilog, VHDL, HTML + CSS, Javascript (SolidJS, TypeScript), MATLAB, bash/zsh scripting

Machine Learning – experience working with time-sequence data, text and sentiment analysis, image analysis, reinforcement learning, image generation, statistical analysis, unsupervised clustering, classification and regression, genetic algorithms, point cloud data, data generation and augmentation, and common libraries like PyTorch, TensorFlow, Jax, XGBoost, Scikit-Learn, OpenAI Gym, and StableBaselines

DevOps – Docker, CI/CD, GitHub workflows & actions, API design, code architecture, unit testing / hypothesis testing / test-driven development, UX design, AWS services (RDS, Aurora, IAM, Lambda, API Gateway, etc.), cybersecurity

Researcher Profile – scholar.google.com/citations?user=100Jo84AAAAJ

h-index = 14

Machine Learning Projects

Senior Data Scientist, SimSpace (Sep 2020 - Present)

Project manager, autonomous cyberattack agent

Sep 2020 – Feb 2023

- Led seven people to create an AI penetration testing agent that could launch autonomous cyber attacks
- Used RL techniques such as DQN, DRQN, ReF-ER, and hierarchical RL, with cybersecurity-specific modifications
- Also explored approaches such as behavior trees, Bayesian techniques, and heuristics based on environment state
- ML work included simulated environment creation, observation and action space design, creation of both simulated and live action libraries, neural net design, experiment tracking, and attack analysis and visualization
- Other work included VM network management, attack actuator and malware creation, building data formatting and storage tools, unit and integration testing, Dockerization and deployment, API development, message passing and logging, documentation, and integration with other services
- Management approach minimized meetings and emphasized pair programming among team members

Graduate Researcher, UIUC (Aug 2013 – Sep 2020)

Lepton identification with tracks

Jan 2019 – Sep 2020

- Trained RNN-based tool for identification of low-energy leptons based on reconstructed tracks
- Tool identified particle decay origin, energy and trajectory of particle, and particle type
- Improved accuracy and background rejection over previous best algorithm, which relied on boosted decision trees
- Tool has subsequently been used in physics searches involving leptons from low-mass-difference decays
- Supervised one undergraduate researcher

Identification and generation of calorimeter decay showers

Feb 2017 – Nov 2019

- Created particle reconstruction and identification software for CLIC, a proposed linear collider
- Trained 3D-CNN algorithm on data from decay showers in calorimeters
- Tool could determine particle decay process, estimate particle energy, and reconstruct flight trajectory
- Improved classification accuracies over existing techniques by over 40x in certain energy regimes

- Extended tool for use in detectors with non-uniform geometries, like ATLAS and CMS
- Contributed to second GAN-based tool that could generate simulated calorimeter showers
- Used batch computing techniques to train models on the Blue Waters supercomputing cluster
- Applied the following techniques: data augmentation, data cleaning, variable-weight loss functions (to deal with imbalanced input classes), hyperparameter optimization
- Lead a team of nine researchers, and presented results at international workshops and seminars

Vertex reconstruction in particle collisions

Apr 2015 – Apr 2016

- Applied unsupervised clustering and FFT-based techniques on 3D voxelated data to improve vertex finding
- Performed statistical comparisons between techniques, and performed hyperparameter optimization
- Improved reconstruction efficiency and computing time over previous best algorithm

ML Course Instructor, Education Justice Project (EJP) (Jan 2019 – June 2019)

ML intro course for prisoners at Danville Correctional Center

Jan 2019 – June 2019

- Program was geared towards prisoners with upcoming release dates, to help gain applicable job skills
- Helped students develop simple applications in image recognition, text parsing, game playing, GAN, etc.
- Course counted for CS credit in the U Illinois system

Non-ML Projects

Senior Data Scientist, SimSpace (Sep 2020 - Present)

Penetration testing tools for cyber ranges

Feb 2023 – Present

- Developed tools for VM ranges used by militaries, national governments, and large companies for cyber training
- Work involved Docker swarm deployment using custom tools, integration testing, setting up proxy-tunnel communications, building attack actuators, synchronizing logging across multiple ranges, coordinating messages between services, setting up CI/CD, etc.

Graduate Researcher, UIUC (Aug 2013 – Sep 2020)

Search for evidence of Supersymmetry

May 2019 – Sep 2020

- Analyzed data from the ATLAS detector at CERN for evidence of theories beyond the Standard Model of Physics
- Developed statistical techniques for estimating collision backgrounds, based on correlated particle decay products
- Contributed to legacy C++ codebase with an extremely large international team
- Processed terabytes of data on the largest grid computing network in the world

Track-finding firmware for ATLAS detector

Jan 2018 – June 2020

- Developed FPGA algorithms to analyze ~100k collision events per second
- Designed algorithms at clock-edge limits, to hit nanosecond-level processing speed requirements

Photon detector development for South Pole Telescope

May 2014 – Dec 2014

- Worked on focusing and detection planes for the third-generation photon detectors in the South Pole Telescope
- Designed, simulated, manufactured, and tested sub-millimeter wavelength optical components

Undergraduate Researcher, UT Austin (Aug 2008 – July 2013)

Plasmon physics measurements on metamaterials

Jan 2012 – July 2013

- Developed methods of launching and measuring plasmonic resonances on metamaterials
- Investigated optical responses of quantum emitters near hyperbolic metamaterials
- Performed data analysis and theoretical calculations on ellipsometry of thin films
- Worked with femtosecond laser and optics systems, vapor deposition, and AFM microscopy

Graphene synthesis

May 2011 – Jan 2012

- Investigated copper-catalytic growth of monolayer graphene using RF heating in a zero-flow gas environment
- Helped assemble vapor deposition chamber and maintain vacuum, heating, and cooling systems
- Ran growth tests and performed code-based analysis using Raman spectroscopy and SEM imaging