David Pitt

Seeking roles in applied science and MLOps

Email: david.h.pitt@gmail.com
Website: dhpitt.github.io
Github: github.com/dhpitt

Education

- **Harvey Mudd College** — Claremont, CA
  B.S., Math and Computer Science - in-major GPA 3.9 — *Aug 2019 - May 2023 Ex.*

  Relevant coursework: *Stochastic Processes (graduate-level), Advanced Linear Algebra, Geometry of Big Data, Mathematics of Big Data, Intermediate Probability & Stats, Applied Algebraic Topology, PDEs, Dynamics, Artificial Intelligence, Computability and Logic, Data Structures & Algorithms, Object-Oriented Programming*

  - **Honors and awards**: Harvey S. Mudd Scholarship, Dean's list every semester.
  - **Math 157 head grader**: Graded all assignments for upper-division probability theory course.
  - **CS5 grader/tutor**: Graded assignments and held weekly TA sessions for advanced intro computer science course.
  - **Community involvement**: Co-founder of HMC Chess, member of Engineers for a Sustainable World

Experience

- **Groundlight AI** — Seattle
  ML Engineer — *Sept 2022 onward*
  - **ML codebase**: Building out features and improving usability of Groundlight's internal machine learning library.
- **Proofpoint** — Claremont
  Consulting mathematician — *Aug 2022 onward*
  - **Adversarial attacks**: Modularizing a library for testing adversarial attacks on large language models.
- **Groundlight AI** — Seattle
  Applied Science Intern — *May 2022 - Aug 2022*
  - **Model tuning**: Implemented search strategies from literature for optimizing hyperparameters of ML models in vision pipeline. Designed and developed a model-agnostic cost-sensitive calibration framework. **Decreased calibration error in production by 21%**.
  - **Science**: Developed and trained new domain-specific feature extractors using self-supervised learning. Explored a novel method for anomaly detection in video. Refreshed internal techwiki, added mathematical proofs. Led paper-reading journal club meetings, presented NIPS/CVPR papers to an audience of engineers and scientists.
  - **DevOps**: Automated end-to-end ML model testing workflows in production API.
- **Inspirit AI** — Remote
  Instructor — *May 2022 - present*
  - **Teaching AI**: Led nightly sessions of a Python-based AI course for high school students.
- **Granite Point Capital** — Boston
  Equity Research Analyst — *May 2021 - Aug 2021*
  - **Equity research**: Led a team of interns to cover 20+ blockchain/digital currency companies and use fundamental tools to deliver investment theses to management
  - **Trade analytics package**: Independently designed and built a lightweight framework to allow PMs to analyze trade flow using prime broker data logs. Used to investigate quality of deal stock sold by smaller brokers.
- **Nanotronics Imaging** — New York
  Junior R&D Associate — *Jun 2020 - Aug 2020*
  - **Image preprocessing**: Implemented a U-net in Tensorflow to correct an artifact of high-res microscopy for use downstream in computer vision pipeline.
  - **Classical computer vision**: Implemented image segmentation/stitching to preprocess another imaging artifact.

Projects

- **Graph neural network for EEG data processing**: (Work in progress) Researching and developing a graph-attention-based approach that structures EEG signals as dynamic networks. Implementing on a series of public EEG datasets.
- **Fractal animator**: Designed tools to visualize and animate ultra-high-resolution animations of dynamical system end behavior. Uses parallelization to render and save to disk orders of magnitude faster than previous tools.
- **Chaotic Encryption**: Designed my own encryption scheme based on chaotic behavior to encode text and image data.
- **N-body Simulator**: Designed my own numerical integrator to simulate solutions to arbitrary N-body problems with intial conditions.

Tools & Technologies

- **Languages**: Python, C++, SQL, Bash, Java, Maple
- **Frameworks**: NumPy, Pandas, scikit-learn, OpenCV, NLTK, Pytorch, Pytorch-Geometric, Tensorflow, Keras, WandB
- **Tools**: Linux, Git, Docker, GitHub Actions, DVC, MySQL, Neo4j
- **Communication**: English (native), Mandarin Chinese (business proficient)