Current as of 2024/02/26
mattrmd@umich.edu
mraymond.info

Matt Raymond

ML, Nanochemistry, Computer Science

Education

University of Michigan — Ann Arbor, MI
Ph.D. Signal/Image Processing and Machine Learning — 08/22–04/26
GPA: 3.92
- Co-advised by Drs. Angela Violi and Clayton Scott
- President of the student signal processing seminar: `websites.umich.edu/~speecsseminar/`
- Thesis Topic: *Generative Modeling of Nanoparticles Via Transfer Learning*

University of Michigan — Ann Arbor, MI
M.Sc. Computer Science — 08/20–04/22
GPA: 3.91
- Member of the VioliGroup computational biochemistry lab (3 semesters, 2 summers)
- President of the Machine Learning Theory Reading Group, 1 semester

Chapman University — Orange, CA
B.Sc. Computer Science, Music Minor — 08/16–05/20
GPA: 3.86
- Member of the Provost List, 8 semesters
- Recipient of the Chancellor's Scholarship, 8 semesters
- Tutor and Supplemental Instructor for Computer Science and Math, 4 semesters

Papers

Joint Optimization Significantly Improves Gradient Boosting — Ongoing
TBD
- Implement GPU-accelerated proximal Nesterov method for joint optimization of piecewise linear ensembles (JOPLEn) with nonsmooth convex losses
- Benchmarked against existing ensemble methods on 65 regression datasets
- JOPLEn significantly improves performance of gradient boosted ensembles ($p \approx 10^{-8}$)
- JOPLEn improves feature sparsity for ensemble feature selection
- In the multitask setting, JOPLEn shares regularization across tasks

Machine Learning Models for Nanoparticle Growth in Nonthermal Plasma — Ongoing
TBD
- Develop surrogate ML models for estimating sticking coefficients of silane nanoparticles in nonthermal plasma
- Supervised undergraduate students
  - Jonathan Lin and Zewei Yu
- Currently cleaning results and preparing paper for publication

Multitask Learning of Universal Features for Chemistry Datasets — Ongoing
TBD
- Develop novel multitask impurity function for gradient boosted trees
- Extend Scikit-Learn to include our method (using Python and Cython)
- Track experiments using Data Version Control and test using PyTest
- Outperform MultiBoost and Dirty LASSO on 7 diverse chemistry datasets, including proteins, nanoparticles, and small molecules
- Show that multiple chemical scales can be represented using a few universal features

Domain-Agnostic Predictions of Nanoscale Interactions in Proteins and Nanoparticles — 04/23
Nature Computational Science (cover Article)
- Developed a Deep Learning framework for predicting generalized nanoscale interactions
- Implemented permutation-invariant Neural Network using Tensorflow (TF)
- Migrated competitors code from TensorFlow (TF) 1 to TF 2 for testing
- Paper DOI: 10.1038/s43588-023-00438-x, Code DOI: 10.24433/CO.8157811.v1

Conferences

Joint Optimization of Piecewise Linear Ensembles 28/03/24
Michigan Student Symposium for Interdisciplinary Statistical Sciences

Hybrid MD-ML for Efficient Modeling of Particle Growth in Non-Thermal Plasma 2023
APS Annual Gaseous Electronics Meeting

Posters

A Taste of Your Own Medicine: Tracing Butyrate Production in The Gut 09/23
University of Michigan EHAIL Symposium

- Bacterial butyrate production is associated with improved gut health, but the mechanisms are not well understood
- Analyzed proprietary gut microbiome data from Michigan Medicine using Python
- Utilized Fused Graphical LASSO to identify microbial interactions
- Recovered known interactions and identified novel interactions for *in vitro* testing

Work Experience

Directed Study & Research 01/21–present
Dr. Scott and Dr. Violi

- Perform novel research in sparse structured multitask feature selection
- Advise computational biochemists on machine learning methodology and literature
- Supervise student researchers; Geometric Deep Learning and Deep Gaussian Processes

Instrument Programmer Long Beach, CA
Lotus Instruments 09/19–11/19

- Developed controls for government-contracted, custom gas chromatography instruments
- Analyzed documentation and created custom libraries for serial data transfer

Software Engineering Intern Troy, MI
Toyoda Gosei 05/19–08/19

- Saved 2,000 man-hours and $60,000 per year through automated purchase order tracking
- Implemented a web-based asset tracking software using full-stack ASP.NET
- Collaborated with Cost Management to solidify requirements and return on investment

Books

Linear Algebra for Data Science, Machine Learning, and Signal Processing Ann Arbor, MI
Cambridge University Press 05/23–09/23

- Proofread and edited textbook draft for Dr. Jeffery Fessler
- Independently verified proofs and suggested improvements for clarity and correctness
- Caught LaTeX typesetting errors
- Available 2024 from Cambridge University Press

Projects

The Implicit Bias of Gradient Descent on Separable Multiclass Data Ann Arbor, MI
U-M Course: EECS 598, 559 12/22, 05/23

- Developed a conjecture and proof sketch for extending *The Implicit Bias of Gradient Descent on Separable Data* to include multiclass PERM losses
- Showed numerically that our conjecture holds for certain well-known loss functions

Real-Time Distributed Learning in Connected & Autonomous Vehicles (CAVs) Ann Arbor, MI
U-M Course: EECS 571 12/21

- Designed distributed learning protocol for sparse gradient propagation
- Implemented simulated learning environment in Tensorflow
- Demonstrated superior generalization, with fewer assumptions than Federated Learning

Domain Exploration Through Artificial Curiosity Ann Arbor, MI
U-M Course: EECS 545 12/20

- Developed simulated Martian terrain for training and evaluation
- Beginning with Shmidhuber's theoretical basis for artificial curiosity, developed an implementation using convolutional auto-encoders
- Defined heuristic "Explorational Value" for evaluating path explored by model
- Performed evaluation against naive models to illustrate effectiveness of artificial curiosity

Needlecast: On-the-Fly Reconfiguration of Spacecraft Flight Software Ann Arbor, MI
U-M Course: EECS 587 12/20

- Collaborated with NASA staff to draft specifications for protocols
- Designed a library for booting NASA core Flight System (cFS) applications on-the-fly

- Implemented Needlecast as a plug-and-play header file for NASA core cFE
- Developed a simulated network switch and web interface for straightforward debugging

AI-Driven Contemporary Archaeology for The International Space Station Orange, CA
U-M Course: EECS 587 01/20

- Analyzed project requirements with Dr. Walsh (co-PI of ISS Archeology)
- Compiled facial training dataset for 240 ISS astronauts
- Utilized convolutional neural networks to label astronauts' faces in NASA photo archives