Mason Lee

5662 West 63rd Street Los Angeles California, 90056 | Phone: (310) 266 6598 | E-Mail: mason_lee@brown.edu

EDUCATION

Brown University, *B.A. Machine Learning (Computer Science)* Providence, RI | **Expected Graduation May 2025**
Brown University, *B.A. Applied Math* **Brown University,** *B.A. Earth and Planetary Sciences* (G == graduate course)

Relevant Courses: Deep Learning for Scientists and Engineers (G), Algorithm Aspects of Machine Learning, Remote Sensing, Tackling Climate Change With Machine Learning, Reinforcement Learning (G), Sequential Decision Making (G) Machine Learning for Earth Systems, Environmental Modeling, Deep Learning, Statistical Inference, Linear Algebra, Independent Research Study (2 year) , Computational Earth Science (G), Software Engineering. Dynamical Systems, Parallel Computing in HPC Systems (G)
Buckley School 4.85 GPA, *Summa Cum Laude* **Los Angeles**, CA | Class of 2020

AWARDS

International Science and Engineering Fair (ISEF) Finalist, *Applied Science Research.*
Stockholm Junior Water Prize, *Applied Science Research.*
Journal of Emerging Investigators Published Author, *Applied Science Research.*
Larry Daughtry Math and Science Award, *Single recipient of California level award for outstanding research by high schooler.*
Chevron Design and Engineering Challenge State Finalist, Autodesk Inventor CAD and pitch competition.
Hack for Humanity @ Brown Finalist, Tally Unified Disaster Database
Keynote Speaker @ Youth Climate Conference Los Angeles, Lead climate conference bring together 10 Los Angeles high schools
Alumni: UCLA Computing Circle, UCLA Math Circle, COSMOS
Organizations: Brown Machine Intelligence Group, BigAI Robotics Lab, OpenAI Forum Community Contributor

WORK EXPERIENCE

CRUNCH Lab: *Machine Learning Researcher* *Brown University September* 2023 – Future
- Created novel physics-informed DeepOnet architecture to learn latent space dynamics for general differential operators
- Formulation of lyapunov stable network bounds, evaluated network on high dimensional coupled stiff-kinetics PDE problem.

Cloud Fire AI: *Machine Learning Engineer* *San Fransisco June* 2024 – Future
- Conditional Normalizing Flow for probablistic wildfire severity and monte-carlo Burn Predictions. Worked with Planet labs to super resolve LANDFIRE vegetation maps 700x with Sentiel + Super Dove constellation. Reinforcement Learning wildfire suppression module to operationalize ELMFIRE. Building with transformer Geospatial Foundational Models (Clay)

Northern Change Lab / Brown University, *Machine Learning Researcher* Providence, RI | Jan 2022 – Jan 2024
- Solved problem of discontinuous river connection using CNNs and super resolution techniques. Implemented dual VGG16-UNET-based image segmentation models, custom loss, data augmentation, tuning, and physics constraints.

NASA Langley Fluid Dynamics Research , *Machine Learning Engineer Intern* *Langley, VA*| Sept 2023 – June 2024
- Sample efficient active-flow-control over airfoils using fluidicic oscillators. Exploration of PINNs for chaotic control.
- Distributed RL training with OpenFoam and Tianshou on multiple HPC env. C++, Python, Singularity, Bash Scripting

NASA Ames Research Center ODIN-FIRE, *Machine Learning Engineer Intern* San Fransisco, CA| August 2023 – Present
- Machine learning driven natural language pipeline to geolocate fires and provide situational awareness in initial attack
- Integration into distributed akka actor bases system, cesium web-dev, HPC computing, LLM fine-tuning + API creation.

NASA GISS, *Software Engineer / Wildfire Researcher Intern* New York, New York| June 2023 – August 2023
- Sensitivity of simulated wildfire spread to parameter uncertainty and spatial resolution of terrain and fuel
- HPC Parallel computing, data analysis, QGIS large data processing. Python, Fortran, Shell Scripting.

Initial Attack Wildland Firefighter / CALFIRE, *Type 1 Initial Attack Hand Crew* Sacramento, CA| June 2020 – May 2021
- Worked 24-48 shifts for 5 continuous months during California's 2020 worst fire season in history.
- Deployed to 15+ wildfires across the state: initial attack on the frontline of the fire, structure protection, evacuation.

Pluto Los Angeles / Experimental Event Space and Nightclub , *Co-Founder* Los Angeles, CA | August 2021 – Present
- Generated 1.5+ billion impression and $400,000 in revenue as head of marketing and data analytics.
- Filmed 4 feature length movies, planned and coordinated 100+ events, worked with A-List artists like Lil Baby, The Weekend, Kylie Jenner, Rihanna, etc. Operated and marketed events with large production budgets, project lead.
- Hosted LA Tech Week's largest event, partnered with several VC funds, deal flow for startups, helped sign artists

Biomass Supply Chain Optimization, Down To Materials, Data Scientist Los Angeles, CA | March 2022 – Aug 2022
Syska-Hennessy Engineering, Mechanical Engineering Intern Los Angeles, CA | June 2019 – Aug 2019
MLDevOps + Systems Consulting, Freelance (variable) Los Angeles, CA | June 2022 – Bespoke

RESEARCH ORIENTED EXPERIENCE

Disaster Migration Dynamics @ Brown Population Studies *ML Researcher, Providence, RI* | January 2023 – Current
- Applying graph neural networks + PINNs to create pop dynamic models for people displaced by wildfires across the US
- Leading team of 5 Brown undergrads supervised by Prof. Elizabeth Fussell and working alongside IRS disaster officials.

Methanotrophic Bioremediation, *Principal Investigator* Los Angeles, CA | January 2016 – January 2021
- 4 year applied microbiology research advised by CalTech PhD. Aidyl Gonzalez-Serrichio
- Developed a novel biotrickling filter inoculated with Methanotrophs capable of degrading chlorinated hydrocarbons.
- Winner of Los Angeles Science Fair (others above). Skills: Microbiology, electrical engineering, physics based modeling

Post-Wildfire Mycological Remediation, Co-Renewal Santa Cruz, CA | Jan 2019 – May 2022

Mason Lee

5662 West 63rd Street Los Angeles California, 90056 | Phone: (310) 266 6598 | E-Mail: mason_lee@brown.edu

- Founded the post-fire remediation branch of Co-Renewal and spearheaded the development of a novel biofilter using oyster mushrooms to immobilize heavy metal in burned structures. Deployed filters to 40+ decimated houses in Santa Cruz .
- Leader of meta-data analysis (50 page lit review), fundraised, grant writing to WSEF, experimental design and execution.

Environmental Toxicology / UCSC Peter Weiss *Research Assistant* Summer 2018

- Worked 5-days a week for 3 months studying the effects of eutrophication on methyl-mercury levels in SC bay

SKILLS & INTERESTS

Technical Skills: Tensorflow, Pytorch, Jax, SQL, Google Earth Engine, QGIS, ArcGIS, HPC Computing (Docker, Singularity, MPI, OpenMP, CUDA). **Strengths:** Physics-Informed Neural Networks, Computer Vision, Research **Language:** Python, Javascript, C++, Java, Fortran, **Interests:** Big wall climbing, Jiu Jitsu, DJing, Soccer

Personal Projects (On-Going as of 10/4/23)

PhyDream: Enhancing DreamerV3 with Physics-Informed Latent Space Adjustments and Differentiable Physics-Engines
JAX, Python, HPC, CUDA

- Creating physically conistent generalized world models across RL enviroments using differentiable kinematics engines to backprop analytical gradients. Including this in end-to-end training will hopefully show increased learning rates and rollout convergence.
- I have a hypothesis that physics will the syntax of information for RL agents, just as language is the syntax for LLMs.

SINDy-PETs: A Flexible Physics Informed Model Based Reinforcement Learning Regime: *RL, PINNs, Applied Math*

- We introduce two methods for incorporating physics-informed bias into Probabilistic Ensembles with Trajectory Sampling (PETS): Residual Correction (RC) and Opinion Input (OI). Our approach uses the Sparse Identification of Nonlinear Dynamics (SINDy) algorithm to learn the physics model with a robust backend to inject useful properties such as conservation of mass, energy and momuentum. SINDy-PETS is most useful in data-sparse environments, especially rigid body systems governmed by ODEs / PDEs.

Improving Probabilistic Model Based Reinforcement Learning Control in Chaotic Environments with Conditional Normalizing Flows: *RL, PINNs, Applied Math*

- In this paper, we present a Model-Based Reinforcement Learning (MBRL) algorithm named Conditional Normalizing Flows Monte Carlo Probabalistic Inference for Learning COntrol (CNF-MC-PILCO). The algorithm relies on Gaussian Processes (GPs) and Conditional Normalizing FLows (CNF) to model the system dynamics and on a Monte Carlo approach to estimate the policy gradient. We show improvement in policy convergence using a multi-modal dynamics model in environments with a bifurcating state-space.

Parallel Computing Theory
MPI, OpenMP, Cuda

- In implementation of parallel computing techniques across homogenous and heterogenous HPC environments. Parallelising the training of a machine learning model that uses a fourier basis decomposition as a dynamics function.

Physics-informed Dyna-Style Model-Based Deep Reinforcement Learning for Dynamic Wildfire Modeling,
RL, PINNs, Wildfire Science, Modelling

- In many real-life examples, data-collection is expensive (learning how to control a forklift in RL will lead to many crashes), thus there exists a need for DRL methods that are highly sample efficient and accurate. Assuming we know a general governing equation (ex: conservation laws), we can apply prior physics knowledge into our data-driven model of the environment. I propose that championing such methods built on top of legacy fire physics models could enable the generation of comprehensive initial attack plans on wildfires.
- Skills: Transformers, Actor-Critic Methods, Model-Based RL.

Multi-Fidelity Bayesian PINNs for Sea-Surface Temperature Observation: *RL, PINNs, Applied Math*

- Remotely sensed data over specific regions may occur at irregular spatiotemporal resolutions, highlighting the need for algorithms that can easily intergrating multiple sources while providing a roboust frame work for UQ. I implemented a Bayesian PINN over the Boston Bay to track sea-surface temperature.

Mapping Supra-Glacial Rivers with a VGG16-Based UNET *Hamiltonian NN (HNN)*

- The Henon-Heiles is a benchmark non-linear system due its well understood transition from order to chaos. I implemented the HNN based method described by (Choudhary et al 2020) in Tensorflow. HNNs are unique because they exploit the structure of conservative systems implicity and learn the gradient of the gradient of the loss function.
- Skills: PINNs, Tensorflow, Pytorch, Hamiltonian physics, Chaos Theory, Nonlinear Dynamics

Implementation of PINNs for Chaotic Systems, *Hamiltonian NN (HNN)*

Mason Lee

5662 West 63rd Street Los Angeles California, 90056 | Phone: (310) 266 6598 | E-Mail: mason_lee@brown.edu

- The Henon-Heiles is a benchmark non-linear system due its well understood transition from order to chaos. I implemented the HNN based method described by (Choudhary et al 2020) in Tensorflow. HNNs are unique because they exploit the structure of conservative systems implicity and learn the gradient of the gradient of the loss function.
- Skills: PINNs, Tensorflow, Pytorch, Hamiltonian physics, Chaos Theory, Nonlinear Dynamics

Deep Reinforcement Learning with OpenAI Gym (cartpole), *DQN, Rainbow DQN, Monte Carlo Tree Search (MCTS), TD Learning, Value / Policy Iteration.*

- Exploration of various reinforcement learning techniques on cartpole (balancing a pole on top of a moving cart) with OpenAI Gym. Based on methods from Reinforcement Learning (Sutton and Barto).
- Skills: Reinforcement Learning, Model-Free RL, Model-Based Rl, Markov Decision Processes

Wildfire Prediction via Remotely Sensed Environmental Indicators, *Time Series via Recurrent Neural Networks: LSTMs*

- Traditionally, fire spread is modeled by simple physics-based modeling. While such models are ubiquitous, spread prediction is improved by a comprehensive set of environmental covariates. I implemented 3 techniques, LSTMs, MLPs, and XGBoost to help identify fire prone areas across California. Data was open-source remotely sensing data related to vegetation, meteorological conditions, as well as fire indicators.
- Skills: Fire modeling, RNNs, anomaly detection, class imbalance, synthetic data techniques.

Physics Informed Neural Networks (PINNs) Solving Burger's Viscous Equation *Environmental Modeling, Nonlinear systems*

- Burger's Viscous equation is a fundamental PDE to describe many turbulent fluid systems. I employed PINNs to simulate Burger's equation under varying conditions. Analysis and comparison with traditional FDEs.
- Skills: shock waves, nonlinear systems, custom loss functions, TensorFlow, model building, DL

Conditional Variational Auto-Encoders (CVAE), *Generative Modeling,*

- CVAEs are subclasses of auto-encoders that allow for more precise mapping of the latent space. I built a CVAE model capable of capturing the latent space of the MNIST data set, allowing me to generate new numbers based on conditional class probabilities. Also used on CIFAR.
- Skills: Custom NN, Tensorflow, Keras, Embeddings, Auto-encoding, data compression

Self-Attention: Transformer Based Image Captioning *Generative Modeling, MNIST,*

- Self-attention is a popular DL technique used in a variety of tasks, here I apply it to create a model capable of accurately captioning images. I used Word2Vec embeddings as a latent encoding for images. Also applied GRUs and LSTMs (a form of recurrent neural network) to the same task. All models were built from scratch.
- Skills: Machine Translation, NLPs, Tensorflow, Transformers, Linear Algebra, Architecture Modifications

Predicting Glacial Runoff Volume with XGBoost, *Remote Sensing, Machine Learning, Glacial Hydrology*

- Part of my work with the Northern Change Lab @ Brown. Lab members hypothesized that snow cover frequency affected runoff patterns in adjacent water basins. I examined historical runoff and satellite data to build an XGBoost model capable of predicting runoff from snow-cover. I created a comprehensive data-pipeline from scratch to process large amounts of sat data and filter it through google colab.
- Skills: Google Earth Engine, MODIS, Remote Sensing Data Processing (ortho-rectification, zenith angle filtering, raw band masks), scikit-learn based regression, sparse data, javascript, python.

Tally: A Unified Disaster Logistics Database, *Hack For Humanity Finalist @ Brown U*

- During my time as a wildland firefighter, I witnessed the chaos of a disaster situation, which is especially exacerbated when multiple agencies are bringing resources to the same region. Tally seeks to solve the inter-agency logistics issue by creating a unified database specific to each incident. Our team of 4 created a working website, UI-UX for both NGOs and disaster victims. We also built a bare-bones SQL database attached to a bar-code based tracking technique.
- Skills: Front-End Development (React, CSS, Node.JS), Natural Language Processing, Tensorflow, Pitching

Building Knowledge Graphs with Graph Neural Networks, *Recommender*

- A graph neural network (GNN) is a neural network that captures the dependence of a graph via message passing between the nodes of the graph. Created a Transductive Learning based GNN using the TensorFlow GNN lib capable of making simple movie recommendations to a user.
- Skills: Graph theory, convolutional GNNs, collaborative filtering, matrix factorization.

Estimating Soil Age Using Cosmogenetic Nucleotides, *Environmental Modeling*

- Cosmic rays are continuously produced by the sun and supernovae and consist of electrically charged particles. When these cosmic rays interact with the upper atmosphere, collisions and reactions lead to the formation of secondary particles, called

Mason Lee

5662 West 63rd Street Los Angeles California, 90056 | Phone: (310) 266 6598 | E-Mail: mason_lee@brown.edu

cosmic showers, which can generate radioactive atoms in rocks at and near the surface of the earth. I modeled this process under sedimentation and erosion using an upwind FDE which gave a concentration used to estimate the age of the soil.

- Skills: Environmental Modeling, Finite Difference Equations, Python, Matlab

Fractional Vegetation Abundance Mapping using Spectral Mixture Analysis, *Remote Sensing using QGIS / LANDSAT*

- SMA is a mathematical inversion in which the "mixed" spectral properties of a pixel are used to calculate the abundance of each portion of the mixture. SMA is a method coined by Prof Jack Mustard at Brown, I used this method to analyze scenes from Owens Valley, California to identify the vegetation abundance, this data was used in an on-going study on the nearby water basin.
- Skills: Remote Sensing, Processing Satellite Data (ortho rectification / radiometrically corrections)

Land Use Cover in the Amazon, *Remote Sensing using QGIS / LANDSAT*

- Used a variety of remote sensing methods such as NDVI and NDWI to identify and quantify the scale of human development in a remote Brazilian village over the past 8 years. Part of a remote sensing class I took.
- Skills: Remote Sensing, Processing Satellite Data, QGIS

Mapping Land Cover using Traditional Remote Sensing and Deep Learning, *Remote Sensing using QGIS / LANDSAT*

- Classification of multispectral images is a common technique for mapping different types of land cover on the basis of spectral properties. This approach exploits the fact that one land cover exhibits relatively uniform spectral properties, which differ from other land cover types. Images over Narragansett Bay, Providence, Rhode Island were analyzed using the statistical Maximum Likelihood Classification and a pre-trained VGG16 based CNN.
- Skills: Remote Sensing, Processing Satellite Data, QGIS, Deep Learning