

Connor Robertson

Data Scientist

@ cnr.robertson@gmail.com C> cnrrobertson.github.io @cnrrobertson cnrrobertson Bentonville, Arkansas

Postdoctoral researcher specializing in computational data science. Strong foundation in mathematics, high performance computing, and software development for a variety of physical applications. Proficient in Python, Julia, R, and Matlab, and comfortable working with remote/cloud Unix systems and with reporting and visualization for non-technical stakeholders.

Experience

Postdoctoral Researcher

Sandia National Laboratories - *Computational Data Science*

2023 Aug. – Present Remote

- Led project to calibrate epidemiological agent-based models of COVID-19 to match historical time-series data with Bayesian inference.
- Improved computational efficiency of simulations using random forest, gaussian process, and neural ordinary differential equation modeling.
- Learned new technologies and tools in uncertainty quantification during project.

PhD Researcher

New Jersey Institute of Technology - *Department of Mathematical Sciences*

2019 Sep. – 2023 May Newark, NJ

- Independently executed project using symbolic sparse regression for interpretable partial differential equation modeling of active nematic fluid systems directly from video experiments.
- Developed video and image processing to extract orientation and velocity data features from video data for physics-guided feature engineering.

Graduate Student Researcher

Oak Ridge National Laboratory - *Center for Nanophase and Material Science*

2021 May – 2022 Jan. Remote

- Project lead to forecast growth and interaction of mutant bacterial strains with spatially-modified LSTM recurrent neural networks.
- Video and image processing to validate model performance against experimental data and to deploy on local experimental equipment.

Data Scientist

Coventina LLC - *Cofounder*

2018 Jan. – 2018 Dec. Provo, UT

- Piloted beta project to automatically generate reports for city utilities providing forecasting of water main breaks via text and visualization to non-technical stakeholders.
- Fused internal and external data sources including soil conditions, usage, and weather to develop physics-guided features for pipe conditions.
- Developed code toolkit to balanced datasets, ensure data quality, and to provide regression and tree-based models for prediction.
- Balanced feedback and technical limitations to achieve business-aligned outcomes.

Curriculum Developer

Brigham Young University - *Department of Mathematics*

2016 Sep. – 2018 Apr. Provo, UT

- Collaborated with group of professors and students to develop and write Python programming assignments teaching data science and numerical computing.
- Research involved the use of network and graph theory to model optimal locations for new water infrastructure in developing countries.

Education

New Jersey Institute of Technology

2018 - 2023 Newark, NJ

PhD in Applied Mathematics

Brigham Young University

2011 - 2018 Provo, UT

BS in Computational Mathematics

Programming Languages



Skills/Exposure

- numpy • scipy • pandas • scikit-learn • pytorch • jax • SLURM • SQL

OSS Contributions

- TidierPlots.jl - R packages ggplot2, patchwork in Julia
- TidierData.jl - R package dplyr in Julia

Spoken Languages

- English • Spanish

Achievements/Certifications

Q Clearance (TS equivalent)

- Department of Energy

Outstanding Graduate Student Award

- NJIT College of Science and Liberal Arts

Machine Learning Seminar Chair

- NJIT DMS

Data Science Education Community of Practice Fellow

- American Physical Society

Ahluwalia Doctoral Fellowship

- NJIT DMS

Graduate Research Fellowship Program

- (Honorable mention) NSF

Publications

- Symbolic diagnostics to interpret and analyze neural network models
OSTI 2024
- Bayesian calibration of stochastic agent based model via random forest
ArXiv 2024
- Performing Video Frame Prediction of Microbial Growth with a Recurrent Neural Network
Frontiers in Microbiology: Systems Microbiology 2023
- Investigating the growth of an engineered strain of Cyanobacteria with an Agent-Based Model and a Recurrent Neural Network
bioRxiv 2021
- Using Survey Data and Mathematical Modeling to Prioritize Water Interventions in Developing Countries
Water Resource Management 2021

Conferences

Talks

- Bayesian Calibration of Stochastic Agent Based Model via PCA Based Surrogate Modeling
SIAM Conference on Uncertainty Quantification – Trieste, Italy 2024
- Data-driven continuum modeling of active nematics via sparse identification of nonlinear dynamics
SIAM Conference on Computational Science and Engineering – Amsterdam, Netherlands 2023
- Data-driven continuum modeling of active nematics via sparse identification of nonlinear dynamics
Annual Meeting of the APS Division of Fluid Dynamics (APS DFD) – Indianapolis, Indiana 2022
- Data-driven continuum modeling of active nematics via sparse identification of nonlinear dynamics
Annual Meeting of the American Physical Society (APS March) – Chicago, Illinois 2022
- Neural networks for function approximation and data-driven modeling
Machine Learning and Optimization Seminar - Department of Mathematical Sciences NJIT – Newark, New Jersey 2021
- Facility location using Markov chains
CPMS Student Research Conference - Brigham Young University – Provo, Utah 2018
- Efficiency of Water Distribution in Water Poor Areas of the World
Student Days - SIAM Annual Meeting – Pittsburgh, Pennsylvania 2017

Posters

- Data-driven discovery of PDEs for active nematic systems
National Academy of Inventors NJIT – Newark, New Jersey 2022
- Discovering governing equations of an active nematic system using PDE-Find
GAMM Juniors' Summer School – (virtual) Magdeburg, Germany 2020
- Aligning Self-Propelling Particles in Non-trivial Domains
Frontiers in Applied and Computational Mathematics – Newark, New Jersey 2019

Organization

- Department of Mathematical Sciences - NJIT
Machine Learning and Optimization Seminar Chair – Newark, New Jersey 2023

Professional Associations

- Society for Industrial and Applied Mathematics (SIAM) 2017 – Present
- American Physical Society 2022 – 2024