Christopher Rohlicek

 crohlicek | 🌐 crohlicek.github.io | ✉ christopher_rohlicek@alumni.brown.edu | ▪ 617.686.3089

SUMMARY

Machine Learning and Data Scientist trained in pure and applied mathematics with research experience applying deep learning (AI/ML) at leading laboratories, including the Broad Institute of MIT and Harvard in Dr. Victoria Popic's lab. Collaborative team member with a demonstrated ability to take initiative and build interdisciplinary relationships. Skilled communicator with a special interest in applications of statistical machine learning to biological systems, and exploring different methods of model complexity reduction.

WORK EXPERIENCE

Broad Institute of MIT and Harvard   September 2021 - present
Computational Associate II – Popic Lab

- Researching applications of novel deep learning approaches to detecting genomic structural variation; Leading development of methods of complex variant detection and benchmarking
- Design and implement software infrastructure for scientific pipelines using Python and Bash, processing large-scale genomic/biological datasets for analyses used in publication
- Developed simulation engine for modeling context-aware complex genomic variation and creating synthetic genomes for method development and evaluation; Released as **open-source software package**
- Leading investigations of generative adversarial network (GAN) applications to histology imaging
- Initiated and led interdisciplinary reading group exploring machine learning and applications in computational biology

Brown University – Carney Institute for Brain Science   June 2020 - August 2021
Research Assistant to Prof. Jason Ritt

- Investigated dynamical dimensionality reduction of neurophysiological brain models
- Used recurrent neural networks to solve inverse problem posed by empirical neural spike data
- Implemented data ETL pipelines to process and prepare synthetic and real biological datasets for model training and evaluation
- Performed EDA with Python data analysis libraries to guide modeling strategies
- Implemented custom recurrent neural network architectures in PyTorch to test hypotheses of effective forms modeling the dynamical systems underlying the data

MIT Lincoln Laboratory   May 2019 - August 2019
Research Intern with Dr. Michael Brandstein

- Applied signal processing and speech recognition techniques to identify bird species from flight call data
- Designed and implemented deep learning speech recognition pipeline to train a classifier on a large dataset of audio recordings
- Combined a CNN image encoder and k-nearest neighbors classifier to process spectrogram image data; Improved on best published accuracy in species-identification benchmark task from 94% to 99%

O&R Patent Law   May 2018 – August 2018
Software Intern

- Designed and developed Python-based software tools for enterprise to automate manual tasks in patent law workflow, saving hundreds of person-hours
- Used Python and SQL to automate entity resolution and summarization of records for financial analysis

EDUCATION

2020 - 2021 **Brown University** – ScM in Data Science
Relevant Coursework: Probability, Statistics, and Machine Learning; Data Engineering;
Statistical Learning; Deep Learning and Special Topics in Data Science; Modern Applications of
Probability and Statistics

2016 - 2020 **Harvard University** – AB in Applied Mathematics
Relevant Coursework: Matrix Methods in Data Analysis, Signal Processing, and Machine Learning;
Decision Theory; Abstraction and Design in Computation; Computer Networks; Optimization;
Honors Linear Algebra and Multivariable Calculus and Real Analysis; Theory of Groups and
Vector Spaces; Vector Space Methods for Differential Equations

PUBLICATIONS AND PRESENTATIONS

Popic, V., **Rohlicek, C.**, Cunial, F., Garimella, K., Meleshko, D., Hajirasouliha, I. (2023) **"Cue: a deep-learning framework for structural variant discovery and genotyping."** *Nature Methods.*

Brendel, M. Honigsberg. R, Maharjan S., **Rohlicek, C.** "Weakly-supervised tumor purity prediction from H&E stained slides." Poster presented at 2022 Starr Cancer Consortium Retreat; April 2022; Cold Spring Harbor, NY.

"A deep learning approach to structural variant discovery" and "Cue: A framework for cross-platform structural variant calling and genotyping with deep learning," presented with V. Popic at Models, Inference& Algorithms Initiative interinstitutional seminar series, Broad Institute of MIT and Harvard, February 2022.

SELECT ACADEMIC PROJECTS

Molecular Toxicity Prediction
Deep Learning and Special Topics in Data Science, Brown University
Completed research on the use of CNNs and BERT-inspired models for the processing of SMILES molecule representations; implemented a novel strategy for predicting toxicity.

Neural Network Parameter Reduction Using Pruning and Matrix Decomposition
Matrix Methods, MIT (cross-registered)
Conducted experiments in PyTorch comparing effects of magnitude-based pruning techniques to matrix decomposition methods of network size reduction.

TECHNICAL SKILLS

Programming Languages	Python, Java, OCaml, SQL, R, MATLAB, Julia, Bash
Libraries	PyTorch, Tensorflow, Keras, Pysam, Numpy, Pandas, Scikit-learn