Charlie Godfrey

godfrey-cw.github.io | godfrey.cw@gmail.com | www.linkedin.com/in/godfrey-cw

Machine learning scientist with research results in explainability, interpretability, robustness, novel architectures and foundation models.

Skills

- **Machine learning research**: empirical science of foundation models, novel deep learning architecture components, evaluating and understanding trustworthiness of machine learning systems.
- **Data modalities**: computer vision, reinforcement learning, natural language processing.
- **Tools**: Python (`pytorch, numpy, scipy, pandas`), Git, Bash, Linux, cluster computing (SLURM), AWS (EC2).

Experience

Postdoctoral Research Associate, Pacific Northwest National Laboratory — October 2022 - Present

- Evaluated distribution-shift robustness of a nueral networks after application of recently introduced *model editing algorithms*, finding that *all editing strategies studied resulted in decreased robustness*, providing a cautionary tale to those deploying edited models and enabling ML practitioners to make a better-informed selection of editing algorithms.
- Found that symmetries of neural network architectures give rise to symmetries of hidden feature representations. These results can be applied to determine what interpretability metrics make sense for a given neural network, *potentially dramatically narrowing the scope of model debugging efforts*.
- Explained how convolutional neural networks to perturbations along spatial Fourier frequencies, demonstrating that *increasing depth and/or weight decay can result in models with sensitivity more closely reflecting frequency statistics of training data*.

Program Associate, Mathematical Sciences Research Institute — March-May 2019

- Participated in the Birational Geometry and Moduli Spaces semester program.
- Presented research on logarithmic Chow-to-Hodge cycle maps at the institute's graduate student seminar.

Education

PhD in Mathematics, The University of Washington-Seattle — June 2021

- Extended results on singularities using inductive construction algorithms for semi-simplicial schemes. Built a new Fourier-type transform on differential forms with poles using duality theory. Defined generalizations of ordinary elliptic curves over finite fields and studied their deformations.
- Completed the eScience Institute's Advanced Graduate Data Science Option
  - PhD-level courses in machine learning, data visualization and statistical inference
  - Implemented machine learning methods like LASSO, kernel regression and k-means clustering in raw `numpy` and `scipy`
- Department of Mathematics Graduate Fellowship (2018-2019)

Student, MSRI Mathematics of Machine Learning Summer Graduate School — July 29-August 9 2019

- Attended mini-courses and problem sessions on statistical learning, convex optimization, bandits, deep learning and reinforcement learning.

- Presented an expository account of recent work on linear stochastic bandits.

Graduate Mentor, Washington Experimental Math Laboratory January 2019-December 2020

- Mentored undergraduate research projects on foundations of quantum mechanics and mathematical epidemiology.

Master's of Science in Mathematics, The University of Washington-Seattle June 2018
Bachelor's of Science in Mathematics and Physics, The University of Wisconsin-Madison May 2014

Publications

1. **Charles Godfrey**, Michael Rawson, Henry Kvinge and Davis Brown. Fast computation of permutation equivariant layers with the partition algebra. In *ICLR 2023 Workshop on Physics for Machine Learning*.
2. Davis Brown, **Charles Godfrey** (equal contribution), Cody Nizinski, Jonathan Tu, Henry Kvinge. Robustness of edited neural networks. In *ICLR 2023 Workshop on Mathematical and Empirical Understanding of Foundation Models*.
3. Henry Kvinge, Davis Brown and **Charles Godfrey**. Exploring the Representation Manifolds of Stable Diffusion Through the Lens of Intrinsic Dimension. In *ICLR 2023 Workshop on Mathematical and Empirical Understanding of Foundation Models*.
4. **Charles Godfrey**, Davis Brown (equal contribution), Tegan Emerson and Henry Kvinge. On the Symmetries of Deep Learning Models and their Internal Representations. In *NeurIPS 2022*. Code available at github.com/pnnl/modelsym.
5. Elizabeth Coda, Nico Courts, Colby Wight, Loc Truong, WoongJo Choi, **Charles Godfrey**, Tegan Emerson, Keerti Kappagantula and Henry Kvinge. Fiber bundle morphisms as a framework for modeling many-to-many maps. In *ICLR 2022 workshop on geometrical and topological representation learning*.
6. Higher Direct Images of Ideal Sheaves, Correspondences in Log Hodge Cohomology and Globally F-Full Varieties. PhD thesis, University of Washington 2021.

Preprints

1. Correspondences in log Hodge cohomology (2023).
2. Henry Kvinge, Grayson Jorgenson, Davis Brown, **Charles Godfrey** and Tegan Emerson. Neural frames: A Tool for Studying the Tangent Bundles Underlying Image Datasets and How Deep Learning Models Process Them (2022).
3. **Charles Godfrey**, Elise Bishoff, Myles Mckay, Davis Brown, Grayson Jorgenson, Henry Kvinge and Eleanor Byler. Convolutional networks inherit frequency sensitivity from image statistics (2022). Code coming soon to github.com/pnnl/frequency_sensitivity
4. Takumi Murayama and **Charles Godfrey**. Pure subrings of du bois singularities are du bois singularities (2022).
5. Higher direct images of ideal sheaves (2022).

Invited Talks

1. May 2023 SPIE Defense + Commercial Sensing 2023.
2. May 2023 ICLR 2023 Workshop on Physics for Machine Learning.
3. February 2023 Boston College Math and Machine Learning Seminar.

4. January 2023 Joint Mathematics Meetings (Boston, MA).
5. November 2022 Purdue Algebraic Geometry Seminar.
6. October 2020 AMS Fall Eastern Sectional Special Session on Algebraic Singularities in Arbitrary Characteristic.
7. October 2020 University of Washington Algebra and Algebraic Geometry Seminar.
8. April 2019 Mathematical Sciences Research Institute Graduate Student Seminar.

Organizing

1. Co-organizer, Pacific Northwest Seminar on Topology, Algebra, and Geometry in Data Science.