Charlie Godfrey

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

Skills

- **Research**: empirical science of foundation models, novel deep learning architecture components, interpretable and explainable AI, evaluating and understanding trustworthiness of machine learning systems.
- **Data modalities**: computer vision, reinforcement learning, natural language processing.
- **Tools**: Python (PyTorch, Huggingface, NumPy, SciPy and Pandas), Git, Bash, Linux, cluster computing (SLURM), AWS (EC2).

Experience

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

- Showed that recently introduced *model editing algorithms* (model updates with low data and compute requirements) *decrease distribution-shift robustness*, enabling ML practitioners to make better-informed decisions about deploying edited models.
- Found that symmetries of neural network architectures give rise to symmetries of hidden feature representations, *narrowing the scope of model debugging efforts*.
- Explained how convolutional neural networks develop sensitivity to Fourier frequencies, providing a case study of how training data, network architecture and optimization hyperparameters impact model robustness.
- Implement training/evaluation of large deep learning models (including transformers). Experience with distributed training, parallel hyperparameter sweeps, unsupervised training, fine-tuning of pretrained models and transfer learning.

Select Publications

1. **Charles Godfrey**, Henry Kvinge, Elise Bishoff, Myles Mckay, Davis Brown, Tim Doster and Eleanor Byler. How many dimensions are required to find an adversarial example?. In *The 3rd Workshop of Adversarial Machine Learning on Computer Vision at CVPR 2023*.
2. **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*.
3. 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*.
4. 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*.
5. **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.
6. Higher Direct Images of Ideal Sheaves, Correspondences in Log Hodge Cohomology and Globally F-Full Varieties. PhD thesis, University of Washington 2021.

Select Invited Talks

1. February 2023 Boston College Math and Machine Learning Seminar.
2. January 2023 Joint Mathematics Meetings (Boston, MA).
3. November 2022 Purdue Algebraic Geometry Seminar.
4. October 2020 AMS Fall Eastern Sectional Special Session on Algebraic Singularities in Arbitrary Characteristic.

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)

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