Kirin P. Danek

✉ kirin@useunix.com 📞 612-481-9784 ⌂ Kirin2D2.github.io in linkedin.com/in/kirindanek

Education

Princeton University, Princeton, NJ   *Expected Graduation: May 2026*
- B.S.E. in Computer Science   **GPA**: 3.75 — **Departmental**: 3.91
- Minor in Statistics and Machine Learning
- Relevant Coursework: Machine Learning (Course Assistant), Statistics, Data Analysis, Data Structures and Algorithms, Linear Algebra, Multivariable Calculus, Discrete Mathematics, Programming Systems

Danish Institute for Study Abroad, Copenhagen, DK   *Expected: Jan 2025 - May 2025*
- Relevant Coursework: Artificial Intelligence, Complex Networks, Analysis of Big Data

Skills

Programming Languages: Java, Python (Pandas, PyTorch, TensorFlow), C/C++, Golang, Rust, R, SQL
Tools: Git/GitHub, Unix/Linux (bash scripting), LaTeX, Stata, VScode, JetBrains platforms, Emacs
Interests: Causal Inference, Missing Data, Machine Learning Theory, Model Compression, Explainable AI, Algorithms, Probability, Statistics

Research

Independent Work: XAI for Compression, Princeton University, Princeton, NJ   *Dec 2024 - Present*
- Research modifying Disentangled Relevant Subspace Analysis to identify irrelevant subspaces in Deep Learning models. Prune model nodes corresponding to irrelevant subspaces, substantially reducing model complexity (i.e., reduce memory, energy cost, time complexity) without losing valuable information.
- Advised by professor Pedro Paredes.

Student Researcher: LLM Performance, Princeton University, Princeton, NJ   *Winter 2024-25*
- Assist professor Pramod Viswanath to create AI teaching assistant for Princeton's ECE470: Principles of Blockchains. Allow students to receive instant expert-level responses to coding and conceptual course queries.
- Curate blockchain-related text dataset + fine-tune LLM, adjust hyperparameters, optimizers, templates adapted from Llama-3 (Python, PyTorch, LLaMA-Factory).
- Demonstrate superior relevance and accuracy for specific blockchain theoretical and practical queries versus existing models (GPT-4o, LLaMA-3).

Author: Proxy Selection in Divergent Data Sets, Princeton University, Princeton, NJ *Jul 2024 - Present*
- Co-author forthcoming data science paper with professor Howard Lavine (UMN).
- Create and develop novel Proxy Finder algorithm. Introduce "correlation reduction" to existing imputation methods to empirically identify a proxy for an essential missing variable.
- Rigorously test performance on synthetic and naturally occurring political science data.
- Eliminate intense human involvement from previous proxy identification methods, allowing political scientists to empirically judge proxy strength, with substantially reduced uncertainty.
- Work accepted at Midwest Political Science Association 2025 conference, intended presentation: April 2025.

Work Experience

ML Undergraduate Course Assistant, Princeton University, Princeton, NJ   *Sep - Dec 2024*
- Grade students' written work and programming assignments (Python) for COS324: Introduction to Machine Learning.
- Hold weekly office hours to assist students with machine learning and statistics concepts, written work, and programming.

Data Science Research Intern, University of Minnesota, Minneapolis, MN   *May - Aug 2024*
- Performed imputation analyses on missing data in survey dataset (Python, Stata); tested low-rank representations; built and fine-tuned neural network (PyTorch) to approximate Bayesian sampling.
- Used clustering analyses to identify and visualize distinct high-dimensional classes of survey respondents, creating empirical cutoffs to meaningfully determine who should be considered "status threatened."
- Project: *Status Threat: The Core of Reactionary Politics*, Drs. Howard Lavine and Christopher Parker.

Featured Projects (see GitHub for more)

LoRMA - Low Rank Matrix Approximation algorithm.   *Spring 2024*
- Implement algorithmic LoRMA (Python, Pandas) via gradient descent to approximate complex datasets as two low-rank matrices.
- Build and test movie recommendation engine, predicting user ratings from MovieLens dataset. Maintain accurate and meaningful low-rank (40) matrix representations of large dataset (600x10000).

multiClassCnn - Computer vision image classifier.   *Spring 2024*
- Use convolutional neural network to build image classifier (Python, PyTorch).
- Trained model achieved 98% test accuracy in classifying images of hand-drawn numbers from the MNIST dataset.

Fraud Detection - ML model for classifying credit card fraud.   *Fall 2023*
- Develop a fraud detection model (Java) for Princeton University's COS226.
- Use clustering, decision stumps, and boosting (AdaBoost) to achieve > 80% test accuracy for detecting fraudulent transactions.
- Model physical space as graph, use graph algorithms for dimensionality reduction of real-world data.