# Wenhao PAN

wenhaop.github.io | linkedin.com/in/wenhao-pan-uw | wenhaopanwork@gmail.com | (510) 646-3809

## EDUCATION

**University of Washington, Seattle** — Seattle, WA
Ph.D. in **Statistics** (GPA: 3.94/4.00) — 09/23 - Present
- Coursework: Discrete & Continuous Stochastic Processes, Multiple Testing, High-dimensional Statistics, Machine Learning for Big Data, Bandits, Measure Theory, Stochastic Calculus (In Progress), Financial Markets (Coursera).

**University of California, Berkeley** — Berkeley, CA
B.A. in **Statistics** and B.A. in **Computer Science** (*Summa cum laude*) — 08/19 - 05/23
- Coursework: Machine Learning, Deep Learning, Convex Optimization, Data Structure, Causal Inference, Linear Modelling, Time Series.

## RESEARCH EXPERIENCE

**University of Washington, Seattle | Advisor: Tyler McCormick, Zaid Harchaoui** — Seattle, WA
**Project:** Incorporating Network Interference into Performative Prediction — 08/24 - Present
- Replaced the Stable Unit Treatment Values Assumption with network interference in a causal experiment about search engines under performative prediction framework.

**University of Washington, Seattle | Advisor: Daniela Witten** — Seattle, WA
**Project:** Improving the Data Thinning Algorithm through the Rank-transformed Subsampling Algorithm — 08/23 - 02/24
- Improved the statistical power of Data Thinning algorithm while maintaining type-I error control through Rank-transformed subsampling.

**Berkeley Artificial Intelligence Research Lab | Advisor: Anil Aswani** — Berkeley, CA
**Project:** Accelerated Nonnegative Tensor Completion via Integer Programming — 05/22 - 05/23
- Accelerated a nonnegative tensor completion algorithm to increase its applicability to healthcare, computer vision, and other domains.
- Developed 10 variants by optimizing data structures, accelerating gradient descent, and applying blended pairwise conditional gradients.
- Reduced the computation time of using aforementioned approach to demosaic a $90 \times 60$ pixel image from 11,300 seconds to 1,500 seconds.

**Lawrence Berkeley National Laboratory | Advisor: Haichen Wang** — Berkeley, CA
**Project:** Graph Transformer Neural Network for Regression in Particle Physics — 01/22 - 01/23
- Improved the deep-learning model proposed in the paper "A Holistic Approach to Predicting Top Quark Kinematic Properties with the Covariant Particle Transformer" for predicting the kinematic quantities of Higgs bosons in the $t\bar{t}H$ production samples.
- Fit a PyTorch Graph Transformer model with 13 million parameters onto 7 million samples using the supercomputer Cori.
- Addressed the model's underprediction of the transverse momentum (pT) of Higgs bosons by reweighting its loss function.
- Increased the percentage of Higgs bosons with both true and predicted pT in a high pT interval from 26% to 44%.

**Oski Lab | Advisor: Cyrus Dioun** — Berkeley, CA
**Project:** Natural Language Processing and Deep Learning for Product Classification — 02/21 - 10/22
- Hand-coded 3,100 cannabis products to create a labeled dataset to build multi-label classification deep learning models.
- Optimized a Keras TextCNN model through hyperparameter tuning to achieve a 93.7% average F1 score for five labels in the testing set.
- Fine-tuned a PyTorch BERT model with Hugging Face to achieve a 95.3% average F1 score for five labels in the testing set.
- Recognized as the most significant student contributor to this project in the acknowledgment of its working paper.

## PUBLICATIONS

- Kura, K., **Pan, W.** and Allen, G (2024). Classifying and Interpreting Moral Judgment Using Reddit Data from r/AmItheAsshole. Submitted to Pluralistic Alignment Workshop at NeurIPS 2024.
- **Pan, W.**, Aswani, A. and Chen, C. (2023), Accelerated Nonnegative Tensor Completion via Integer Programming. Frontiers in Applied Mathematics and Statistics, 9, p.1153184.

## PROJECTS

**Sequential Investment and Universal Portfolio Algorithms** — 03/24 - 06/24
- A class project about experiments on portfolio selection algorithms on real stock data. The full report is on the personal home page.

**Constructing Priors that Penalize the Complexity of Gaussian Random Fields** — 01/24 - 03/24
- A class project about simplifying and reproducing the paper in the project title. The full report is on the personal home page.

## TEACHING EXPERIENCE

**University of Washington, Seattle | Teaching Assistant** — Seattle, WA
- STAT 516, Stochastic Modeling of Scientific Data — 09/24 - Present
- STAT 390, Statistical Methods in Engineering and Science — 06/24 - 08/24
- CSE 416, Introduction to Machine Learning — 03/24 - 06/24
- STAT 180, Introduction to Data Science — 01/24 - 03/24

## TALKS & PRESENTATIONS

**Why Does Transformer Not Work For The Higgs Boson?** — Berkeley, CA
ATLAS Lawrence Berkeley National Laboratory 2022 Annual Meeting — 01/23

## SKILLS

- Languages: Python, R, Java, C, MATLAB, SQL.
- Packages & Tools: PyTorch, PySpark, Keras, HuggingFace, NumPy, Pandas, Matplotlib, Scikit-Learn, Scipy, Jupyter, Conda.