Alexander Dukart

aydukart | Alexander Dukart | https://aydukart.github.io | ald349@pitt.edu

Summary

Fourth-year Ph.D. student in Statistics at the University of Pittsburgh with research interests in machine learning. Work focuses on explainable artificial intelligence and developing methods to understand black-box models, aiming to bridge the gap between predictive performance and interpretability. Research is supervised by Professor Lucas Mentch.

Education

2022 - Present	PhD Statistics at **University of Pittsburgh**	
2022 - 2025	MA Statistics at **University of Pittsburgh**	
2018 - 2022	B.S. Mathematics and Economics at **University of Minnesota**	GPA: 3.94

Research Projects

Random Forests
Investigating why random forests perform well across different signal-to-noise ratio (SNR) regimes to understand their adaptability and the role of regularization in machine learning.
Statistics Department Rankings (publishing 2025)
Developing advanced statistical methods and new theoretical frameworks for ranking academic statistics departments and understanding the statistics of tiers.
Semantic Similarity in Ensemble Tree Methods
Exploring methods for grouping decision trees within ensemble models based on semantic similarity to improve interpretability and model efficiency.

Consulting Experience

Extensive experience across diverse fields including Natural Language Processing, Computer Vision, Psychiatry, Nursing Research, Neuroscience, Education Research, and Biology.

Project Highlights

- **Evaluating LLM Reasoning:** Assisted in evaluating large language models' reasoning capacity by verifying statistical methodology and correctness of mathematical computations.
- **Physician Bias Research:** Contributed to a project on physician bias against children with type 2 diabetes by performing data cleaning and supporting statistical analysis.
- **COVID-19 Vaccine Resistance Survey:** Worked on evaluating a survey measuring COVID-19 vaccination resistance, with responsibilities including question design assessment, data cleaning, and analysis support.

Teaching Experience

- **Instructor** — *Data Modeling Using R (EE 5373)*
  Led an introductory data science course for graduate students in the electrical engineering department, integrating LLM usage and expanding content to include various machine learning concepts.
- **Teaching Assistant** — *Statistical Learning and Data Science (STAT 1361/2360)*
  Assisted in an advanced data science and machine learning course for senior undergraduates and

graduate students.

- **Instructor/TA** — *Applied Statistical Methods (STAT 1000)*
  Taught an introductory statistics course for undergraduates.

Skills

- **Programming Languages:** R, Python, Matlab
- **Python Libraries & Frameworks:** NumPy, Pandas, TensorFlow, Hugging Face, NLTK
- **R Libraries & Frameworks:** Tidyverse, ggplot2, Shiny, XGBoost
- **Research Areas:** Machine Learning, Explainable Artificial Intelligence, Model Interpretability, Random Forests, Large Language Models (LLMs)
- **Statistical Competencies:** Statistical Modeling, Data Analysis, Data Cleaning, Research Design