Jonathan Kerby-White

Data Scientist | ML Engineer | Statistics and Data Science MS @ UW-Madison | Applied ML & Research Projects
GitHub Profile: https://github.com/jkerbywh
Personal Portfolio: https://jkerbywh.github.io/portfolio/
LinkedIn Profile: https://www.linkedin.com/in/jonathan-kerby-white-7617a2118
Madison, Wisconsin 53703 (219) 771-5528 kerbywhite@wisc.edu

EDUCATION

University of Wisconsin–Madison, College of Letters & Sciences, Madison, WI
Master of Science in Statistics and Data Science *Expected Graduation: May 2026*

Indiana University Bloomington, College of Arts and Sciences, Bloomington, IN
Bachelor of Science in Pure Mathematics *Graduated in May 2024*

EMPLOYMENT EXPERIENCE

Data Science Graduate Project Assistantship (PA) *September 2024 – Current*
World-class Instructional Design and Assessment (WIDA) Graduate PA through the UW – Madison Wisconsin Center for Education Research (WCER), Madison, WI

- Analyzed 12M+ K-12 English Learner assessment records using Python, SQL, Stata, R, and Excel
- Performed time-series and regression modeling to support education policy decision making for 30+ states
- Delivered insights to stakeholders and government agencies across the U.S.

Statistics and Mathematics Graduate TA/Tutor *August 2021 – Current*
Graduate and Undergraduate TA for Multivariate Calculus and Statistics Tutor at UW Madison and Indiana University

Information Technology (IT) Help Desk Technician *August 2024 – May 2025*
Library IT Help Desk Technician at UW – Madison dealing with Windows, iOS, and Linux operating systems

PROJECTS

NBA Game Prediction Model (Python, scikit-learn) *In Progress*

- Built ML pipeline to predict NBA game outcomes using team stats and rolling means
- Engineered features like team win streaks, plus-minus differential, and back-to-back games
- Trained XGBoost and logistic regression models; tuned with GridSearchCV

Machine Learning Marathon (MLM) *September – December 2024*
12-week hackathon with real-world ML projects on Kaggle at UW – Madison

- Performed regression analysis using Python to predict house prices on a high-dimensional dataset.
- Cleaned, transformed, and analyzed the data using the Python packages Pandas and PySpark.

Edge Detection and Image Smoothing *December 2024*
Edge Detection and Image Smoothing of .png file in R

- *Created smoothing and sharpening filters with matrix convolution*
- *Wrote functions to apply custom filters to image matrices for cosmetic and edge effects*
- *Visualized before/after using grayscale transformation and edge overlays*

Web Scraping in R *November 2024*
Exploratory Web Scraping in R

- Wrote code to scrape NFL data and IMDB Top 250 Movie List for exploratory data analysis.

Unbeatable AI – Connect Four / Tic-Tac-Toe *October 2024*

- Developed recursive Minimax algorithm with alpha-beta pruning for optimal move selection.
- Designed full UI using base R plotting and identify() for human-AI interaction
- Implemented win detection logic for rows, columns, and diagonals

Laplacian Eigenmaps and Orthogonal Polynomials *May 2023 – August 2023*
Research done during the 10-week REU at the University of Connecticut (UCONN)

- Analyzed the k = 1 case to determine that the eigen-coordinates are Chebyshev Polynomials.
- Used Python packages such as NumPy and Matplotlib to better visualize the eigenmaps.
- Published *Convergence, optimization and stability of singular eigenmaps* in arxiv.org.

TECHNICAL SKILLS

Languages & Tools: Python, R, SQL, Bash, HTML/CSS, Stata, SAS, JavaScript
Libraries & Frameworks: scikit-learn, XGBoost, TensorFlow, NumPy, Pandas, Matplotlib, PySpark
Technologies: Git/GitHub, Linux, Conda, Google Colab, APIs, Web Scraping, Web Development
Domains: Machine Learning, Deep Learning, Statistical Inference, Regression, EDA, Time Series