

Josh Kraus

Data Science Intern

joshtkraus@gmail.com

(919) 889-1180

Chapel Hill, NC

[LinkedIn](#) | [Portfolio](#)

Education

M.S. Analytics

Georgia Institute of Technology

August 2022 - Present

GPA: 4.00, Anticipated Graduation: May 2024

M.M. Education

University of Florida

January 2020 - December 2023

GPA: 4.00, Summa Cum Laude

B.M. Performance

University of North Carolina - Greensboro

August 2016 - December 2019

GPA: 4.00, Summa Cum Laude

Skills

Programming

Python: NumPy, Pandas, Scikit-Learn,

Beautiful Soup

R: Tidyverse, Caret, Forecast, Knitr

SQL: MySQL

SAS

Supervised Learning

Support Vector Machine, Artificial Neural

Network, General Linear Model, Decision

Tree, Naïve Bayes Classifier, K-Nearest

Neighbor, Ensemble Methods, Factor

Analysis

Unsupervised Learning

K-Means, Principal Component Analysis,

Gaussian Mixture Model, Kernel Density

Estimation, Anomaly Detection

Analytics Modeling

Time Series Forecasting, Simulation, Linear

Programming

Data Visualization

Python: Matplotlib, Seaborn

R: ggplot2

Microsoft Power BI

Essential Skills

Communication, Collaboration, Project

Management, Strategic Planning, Attention

to Detail, Critical Thinking, Creativity,

Adaptability

Work Experience

Global Product Analytics Intern

Liberty Mutual Insurance

June 2023 - August 2023 | Boston, MA

- Developed an innovative actuarial model for four homeowners' perils, utilizing GLMs and incorporating property and auto factors to predict loss costs.
- Successfully handled a dataset of over 10 million rows, utilizing SAS, SQL, and Python for efficient data manipulation, extraction, and analysis.
- Optimized modeling efficiency by implementing sorted variables and grouped factors, utilizing statistical methods to reduce task time and dimensionality.
- Streamlined the analysis of model outputs by creating interactive dashboards with Microsoft Power BI to enable faster decision making.

Founder & Analytics Consultant

Mindful Modeling

January 2022 - Present | Raleigh, NC | [Website](#)

- Build data-driven solutions for educational institutions to better utilize resources.
- Project(s): Enable high volume, high contact student services organizations to efficiently provide staffing to meet student demands by visualizing time series forecasts and historical trends of event attendance by students. The solution utilizes an interactive dashboard built in Microsoft Power BI and analysis performed in Python and R.

Music Educator

Green Level High School, Cary High School

August 2016 - Present | Cary, NC

- Develop curricula, and teach the fundamentals of music performance.
- Collaborate to establish goals and best practices for educating students.
- Address the individual needs of students by adapting teaching methods.

Notable Projects

Machine Learning for Medical Image Analysis

- Led a collaborative research initiative to develop seven machine learning models using Python to classify pathology findings in chest X-ray images.
- Employed advanced techniques, including data balancing through undersampling and data augmentation in Python, to enhance model performance and address class imbalance challenges.
- Conducted rigorous evaluation of model performance using industry-standard metrics such as ROC curves and AUC scores.
- Produced comprehensive reports, effectively communicating research findings and contributing to the advancement of medical image analysis.

Optimal March Madness Bracket using Simulation

- Utilized Python for web scraping, data aggregation, and manipulation.
- Developed seven bracket completion strategies integrating supervised learning, game theory, and domain knowledge.
- Conducted 1000 Monte Carlo simulations on the seven bracket strategies.
- Compared and evaluated their performance using confidence intervals, hypothesis tests, and cumulative distribution functions.