

SEAN MCOSKER

Fredericksburg, VA · 5402735339 · 6mcoskers@gmail.com

LinkedIn: <https://www.linkedin.com/in/sean-mcosker/>

Website: <https://seanmcosker.github.io/>

EDUCATION

Guarini School at Dartmouth College, Hanover, NH

Sep 2022

Master of Science in Data Science, Concentrations in Health Data Science, Epidemiology

Relevant Coursework: *Applied Machine Learning, Algorithms in Big Data, Biostatistics I/II*

Dartmouth College, Hanover, NH

June 2021

Bachelor of Arts, Major in Biology, Minor in Religion

GPA 3.6/4.0

Relevant Coursework: *Genomics, Macroevolution, Epidemiology I/II*

WORK EXPERIENCE

Fluence, Arlington, VA

June 2023–Sept 2023

Applied Machine Learning Graduate Intern

- Theorized and implemented anomaly prediction system for inverter components using an unsupervised learning approach spanning 200 billion data points per year and 4000 features
- Created detailed taxonomy of AI/ML methods and companies in the renewable energy space and created an associated knowledge base to better inform stakeholder decisions and strategic vision
- Designed plan for searchable SQL database for Fluence industry connections with visualization in PowerBI

Technical Skills Used: Python, Github, Pytorch, SQL, Pandas, AWS, PowerBI, Big Data, Data Analysis, Unsupervised Learning

Form Bio, Seattle, WA

Dec 2022–Sept 2023

Consultant

- Performed analysis of workflow run data and wrote scripts with Python to implement a documentation changelog
- Completed total overhaul of client-facing documentation for bioinformatics and genomics suite
- Reworked documentation files in HTML to be compatible with company command-line interface tools
- Designed language-agnostic company style sheet in Google Colab for use in Python and R workflow runs

Technical Skills Used: Python, R, Github, Command-Line Interface, Google Colab, Jira, Agile methodologies, HTML, YAML

Dartmouth-Hitchcock Medical Center, Hanover, NH

Mar 2022–Present

Machine Learning Research Scientist

- Developed Graph Neural Network workflows specifically tailored to use data valuation metrics such as the Data Shapley for use in interpretable and efficient machine learning
- Used GNN workflows for use in node and graph-level image classification on squamous cell carcinoma work
- Worked through Discovery Linux cluster for submitting and evaluating workflows

Technical Skills Used: Python, Pytorch, Tensorflow, SQL, Pandas, Numpy, Linux, Github, Distributed Computing, Machine Learning

TECHNICAL PROJECTS

Granum Health, Principal Analyst

- Helped found startup (now incorporated with % equity) focused on data-driven diagnostics for chronic fatigue/long-COVID patients. I am currently the chief statistical programmer/analyst where I am focused on classification and clustering based on proteomic and patient-reported symptom data. Recently accepted to ICorps Buffalo accelerator

BJJ Webscraping + ELO-Style Ranking

https://github.com/seanmcosker/BJJ_analysis

- Built custom web-scraper with Python to scrape BJJ tournament data. Cleaned data with Excel, SQL, Python, and used custom ELO-style ranking system to assign ranking to competitors. Automated with Cron and Databricks.

SKILLS & INTERESTS

Technical: Python, SQL, R, Pytorch, Tensorflow, Keras, Data Analysis, Machine Learning, Deep Learning, Github, Jira, Agile Methodologies