SEAN MCOSKER

Fredericksburg, VA · 5402735339 · 6mcoskers@gmail.com
LinkedIn: https://www.linkedin.com/in/sean-mcosker/
Website: https://seanmcosker.github.io/

EDUCATION

Geisel School of Medicine at Dartmouth College, Hanover, NH   **Sep 2022**
Master of Science in Quantitative Biological Sciences, 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
Relevant Coursework: *Genomics, Macroevolution, Epidemiology I/II*

WORK EXPERIENCE

Fluence, Arlington, VA   **Jan 2024-Present**
AI and Data Science, Digital Offers
- Designed and implemented retrieval-augmented generation (RAG) application powered by a large language model (LLM) for internal documentation and external market operating information, reducing time for information retrieval by 80%. Led team of three to create this application
- Responsible for evaluating and implementing novel AI/ML offerings in the energy storage space, including state-of-health estimation using classical ML, inverter anomaly recognition using deep neural networks, and optimizing controls systems using reinforcement learning.
- Created end-to-end data engineering workflow for cell balance data, taking data from cloud database, performing data transformation with Python, then visualizing in PowerBI for end-users, automated to run monthly with Cron.

Technical Skills Used: Python, AWS, LLMs, Deep Learning, Pandas

Fluence, Arlington, VA   **June 2023-Sep 2023**
Applied Machine Learning 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 – helped design AWS architecture for this approach
- Created detailed taxonomy of AI/ML methods in the renewable energy space with strategic vision for capturing a percentage of market share equivalent to 200 million USD
- Created a 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-Sep 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