Cameron Erdman

cmerdman.dev

Experience

JP Morgan Chase Sep. 2024 – Present

AI R&D Data Scientist

- Proposed and developing an Agentic AI solution to ingest business needs from Jira's MCP, propose a plan, orchestrate agents and tools, write playwright test scripts, self heal errors, and report outcomes in order to automate UI integration testing, a monthly 20-hour task.
- Researched and implemented Autoencoder, LOF, and IF anomaly detection models; explainable AI concepts including SHAP and LIME analysis; and Pareto optimality for multi-objective optimization to support an ensemble learning approach on multiple financial data sets, the current solution is a 12% improvement on existing systems.
- Built an internal Metric Store which converts proprietary business logic into SQL based metrics; currently supporting daily refreshes on 130 metrics to 12 downstream controllers, leading to a 70% user compute use reduction and 50% query speed increase.
- Converted high complexity data investigation tasks into learnable logic to be fed into Databricks Genie spaces, providing a tailored LLM based data analytics tool to ease business user workload
- Presented multiple internal webinars explaining topics such as Databricks, anomaly detection, and data engineering to groups of 20 to 100 people.

AI & ML Data Science Intern May 2023 – Aug. 2023

- Developed a 6 model ensemble anomaly detection workflow to check millions of daily trade contracts for upstream errors in the data engineering pipeline.
- Built a data quality process to check the accuracy of incoming and outgoing data streams to support the transition of data pipelines, automating a daily 30 minute task.

Dr Christian Blanco Feb. 2022 – Aug. 2023

Research Assistant

- Contributed to published research showcasing the financial impacts of illegal and unsustainable business practices.
- Utilized OpenNLP for named entity extraction, topic modeling, and sentiment analysis on 100k documents over 40 years.

Amazon May. 2022 – Sep. 2022

Data Analytics Intern

- Discovered 500 unresponsive or poor response medical prompts which when routed to our product leads to a boost of 250 to 500 thousand monthly active users.
- Performed customer segmentation analysis on the Alexa Health user base and presented my findings to management.
- Took over a root cause analysis task after our Business Intelligence Engineer left the team, identified the cause of a metric discrepancy and reworked the SQL query in the ETL pipeline to prevent future errors.

Education

Georgia Institute of Technology Atlanta, GA

Masters of Science in Computer Science *Jan. 2025 – May 2027*

Ohio State University Columbus, OH

Bachelor of Science in Data Analytics; Minors in Economics, Computational Analytics *Aug. 2020 – May 2024*

Projects

Mechanistic Interpretability Research | *Transformers, Sparse-AutoEncoders* Present

- Currently I am using toy models to better understand concepts such as Superposition and Branch Specialization.

National Ecological Observatory Network Time Series Classification | *keras* September 2021

- Developed Time Series Classification models using Convolutional and Recurrent Neural Networks to single out and determine the status of airplanes at the KOSU airport. Completed under Dr. Tanya Berger-Wolf.