Arnav Grover

(203)-554-9901 | arnavgrover@gmail.com | linkedin.com/in/arnav-gr0ver | github.com/Arnav-Gr0ver

Education

Purdue University West Lafayette, IN
Bachelor's in Computer Science, Concentration in AI/ML *July 2023 - Present*

Technical Skills

Languages: Python, R, SQL, Java, C++, C, HTML, CSS, JavaScript
Frameworks: PyTorch, TensorFlow, FastAPI, Flask, Streamlit, Selenium, JUnit
Libraries: scikit-learn, Keras, pandas, NumPy, Matplotlib, Pydantic, Unstructured, SQLite

Experience

Machine Learning Engineering Intern February 2024 – May 2024
HawkenIO *Remote*

- Developed a sophisticated whitepaper generation system using FastAPI, providing SaaS clients with customized AI-driven insights on business development, significantly enhancing client acquisition and engagement.
- Engineered and implemented a technical framework for client data extraction, preprocessing, embeddings, and storage, integrating vector similarity algorithms for Retrieval-Augmented Generation (RAG) and interfacing with OpenAI's GPT-4 API. Utilized data validation techniques and Postman for API testing, resulting in robust and reliable client-specific proposal creation.

Open Source Machine Learning Developer September 2023 – Present
Amazon Web Services (AWS) *Open-Source Contributor Initiative*

- Selected amongst Top 4% of global applicants to join a program aimed at building my open-source portfolio, gaining industry-level software development experience, and receiving mentorship from experienced engineers at AWS.
- Improved test coverage of client handling, schema, and port validation in the ML-Commons plugin of the OpenSearch Data Analytics suite from 66.66% to 80.51% by meticulously addressing edge cases. This effort significantly bolstered the system's resilience and reliability, enhancing user experience and confidence.
- Introduced Java Microbenchmarking Harness (JMH) in the GitHub Actions workflow, enabling automatic benchmarking of models for each release, facilitating systematic performance evaluation and tracking, and driving continuous improvement of the ML Commons plugin.

Financial Data Science Intern July 2022 – August 2022
Sandbrook Capital *Stamford, CT*

- Initiated data orchestration project under guidance of the Chief Data Science Officer. Wrote a Python script to extract, manipulate, and stage 15+ million lines of nodal data monthly from S&P Capital IQ Market Intelligence Platform to the firm's Snowflake data lake allowing for efficient analysis of ISOs & regional power grids.
- Derived meaningful insights for the firm's ESG investment strategy by producing a comprehensive policy outlook on the German wind/renewable market and offshore wind market through the analysis of 25+ government publications and consultancy reports.
- Practiced principles of financial modeling, fund management, and current emerging trends within the development of climate infrastructure through the completion of case studies with firm associates.

Projects

IEEE | *2024 BigData Cup Challenge* July 2024 – Present

- Developing a supervised classification model to distinguish bursty event cascades from normal events in social networks, utilizing natural language processing techniques and rigorous research methodologies in the *Characterizing User Behavior in Social Networks: Propagation, Prediction, and Sensemaking Challenge.*

Awards

Network For Teaching Entrepreneurship, World Series of Innovation Challenge, International 1st Place
Intuit Social Innovation Challenge, National Finalist
Purdue Catapult Hacks, AI x Business Hackathon, 1st Place
Milton Fisher Scholarship for Innovation and Creativity, Awardee