David Hu

571-551-9994 | azg9cb@virginia.edu | linkedin.com/in/davidhu426 | github.com/dhu2022-dev

Education

University of Virginia Charlottesville, VA

Bachelor's in Computer Science and Statistics, Minor in Business *Sep. 2022 – June 2026*

- Relevant Coursework: Data Structures & Algorithms, Machine Learning, Probability Theory, Financial Analytics

Experience

Software Engineering Intern - Technology Internship Program June 2025 – Aug. 2025

Capital One – OptiCloud Team, Cloud Platforms (CORE) *Richmond, VA*

- Developed and deployed a production-grade anomaly detection system that identified faulty cloud optimization recommendations, boosting AWS spend efficiency by 35% and saving $20,000+ annually.
- Architected scalable ML pipelines using Databricks and AWS Glue, processing 10M+ training records across S3, PostgreSQL, and DynamoDB; deployed STL and Autoencoder models with SHAP-based explanations.
- Built and integrated Go services powering real-time health dashboards for FinOps and Cloud Governance teams, using config-driven APIs and dynamic visualization components (Panel Definitions and DataViews).

AI Developer Intern May 2024 – Aug. 2024

American Chemical Society (acs.org) *Washington, D.C.*

- Developed a machine learning pipeline for journal classification. Implemented text preprocessing, LDA topic modeling, and visualization, leading to a 30% improvement in document topic insights. Trained BERTopic LLM to categorize manuscript enhancing categorization efficiency by 25%.
- Created an AI script in Python to extract research funder information from publications. Used the spaCy LLM with regex to isolate funder and grant information. The script reduced manual data entry time by 40%, enhanced data management accuracy in classifying future funders, and helped connect researchers to 1000+ new funders.
- Automated API data extraction to improve article search relevance. Processed and managed data with Docker, SQL (PostgreSQL), Microsoft Azure, and Doccano. This new method cut researcher data retrieval time by 50% and streamlined data processing workflows.

Lead Developer Jan 2024 – Present

Red Light Management *Charlottesville, VA*

- Led a team of 5 in a consulting project with Red Light Management (RLM), a mid-sized concert venue. Developed a predictive model for concert ticket sales, coordinated efforts with RLM to address their objectives, and designed the AI application architecture.
- Developed a machine learning model with Azure ML Studio and trained the model on Ticketmaster and Spotify API data. The model has an 85% accuracy in predicting ticket sales.
- Engineered a user-friendly web application to manage the services for RLM's non-technical employees. Built on Django, Python, HTML/CSS/JS, and hosted on AWS DynamoDB for additional hands-free back-end support.
- Improved event planning accuracy and strategic artist selection, leading to a 20% increase in concert revenue. The app also enhanced forecast precision and reduced ticket sale overestimation errors by 50%.

Projects

UVAHacks Options Automation Tool (1st Place) | *Python, OAuth 2.0, Ngrok, Requests, JSON, Streamlit* April 2024

- Best Finance Hack at UVA's Hoohacks 2024, the top financial project among over 2000 competitors.
- Built a financial analytics web app with Streamlit, reducing data visualization and portfolio management time by 50%.
- Integrated Black-Scholes and Fourier models for real-time options pricing, improving pricing accuracy by 40%.

HackMIT AI Notetaker (Y-Combinator Winner) | *Python, TensorFlow, OpenCV, Modal* September 2023

- Developed an AI-powered application at HackMIT that processes handwritten notes or lecture recordings to generate structured notes and diagrams based on user feedback.
- Implemented Optical Character Recognition (OCR) using TensorFlow and OpenCV for text extraction from notes images.
- Leveraged Modal's stable diffusion functionality to generate high-quality images and diagrams from user input, integrating it with the application's note-taking features.
- Awarded a lifetime opportunity to pitch to HackMIT sponsor YCombinator for our idea.

Skills and Certifications

Skills: Python, Java, SQL, Machine Learning (LDA, BERTopic, Scikit-learn), Cloud (AWS, Microsoft Azure), Data Analysis, API Integration (Ticketmaster, Spotify), Flask, Django, TensorFlow, Streamlit, OpenCV, Pandas, Docker, PostgreSQL

Virtual Job Simulations: JP Morgan Chase Agile, Quant, Advanced SWE Virtual Experience; BCG Intro to Strategy Consulting

Other: Portrait Photography Patent (No. 11086197), ICPC Regionals Honorable Mention