

DAVID CUI

cuidavid02@gmail.com | [Github](#) | [LinkedIn](#) | davidcui.net | +1(617) 866-9522

EDUCATION

Northeastern University, Khoury College of Computer Sciences

Boston, MA

B.S. in Computer Science and Economics

Sep 2021 - Dec 2024

- **GPA:** 3.7/4.0, Dean's List
- **Selected Coursework:** Algorithms & Data Structures, Computer Systems, Object Oriented Design, Database Design, Logic and Computation, Programming in C++, Applied Econometrics, Linear Algebra

PROFESSIONAL EXPERIENCE

Starday Foods (CPG Startup)

Remote

Junior Data Science Software Developer (Part-time)

Sep 2024 - Present

- Built an automated data pipeline with AWS Lambda, SQS, and Glue to process TikTok hashtag data, storing results in S3 and delivering weekly summaries to Snowflake

Junior Data Science Software Developer (Internship)

July 2024 - Aug 2024

- Optimized access to Meta Llama 3.1 and Anthropic Claude 3.5 models using AWS Bedrock and SageMaker, processing over 100,000 queries efficiently through chatbot integrations and developing chatbot wrappers for seamless user interactions
- Developed custom APIs and Lambda functions in AWS to analyze over 100,000 product images and descriptions, extracting attributes and automating analysis to deliver deeper insights and reduce manual effort by over 500%
- Engineered an automated data pipeline on AWS S3 with Lambda functions, connecting over 10,000 consumer purchase records and survey insights to CDP profiles, driving seamless data integration and boosting analytics precision

Junior Data Science Software Developer (Co-op)

Feb 2023 - Dec 2023

- Designed and managed a scalable data platform on AWS to ingest and process over 1,000,000 TikTok videos daily, transforming JSON data into structured tables for seamless analysis using S3, Glue, Athena and Lake Formation
- Built a product database by web scraping data from over 100,000 products across major retailers using Octoparse, integrating structured tables into Snowflake
- Developed a machine learning pipeline in Python to detect fake survey responses, preprocessing categorical data and training models to classify the authenticity of over 3,000 responses, significantly reducing manual review time for stakeholders

SKILLS & TECHNICAL TOOLS

- **Programming Languages:** Python, Java, SQL, R, C++, C, LaTeX, Racket, HTML, Javascript, ACL2, Lean
- **Technologies:** AWS, Linux, Snowflake, MySQL, Docker, Octoparse
- **Frameworks:** Flask, React.js, PyTorch, JUnit, Pytest

PROJECTS

Founder & Developer

Automated Sports Betting Service

July 2024 - Present

- Built automated Discord bots to deliver profitable long-term expected value (EV) sports wagers to over 100 paying clients, while also providing personalized insights and guidance to help them understand and apply EV betting strategies
- Programmed Discord bots using Python and the Discord.js library to post wagers by web scraping odds from multiple sportsbooks and comparing them to de-vigged fair value odds from globally recognized sources (e.g., Circa Sportsbook).