DAVID CUI

cui.da@northeastern.edu | github.com/dcui02 | linkedin.com/in/david-cui | davidcui.net

EDUCATION

Northeastern University, Khoury College of Computer Sciences   Boston, MA
Candidate for BS in Computer Science and Economics   *Expected May 2025*
- **GPA**: 3.72/4.0, Dean's List
- **Selected Coursework**: Computer Systems, Object Oriented Design, Algorithms & Data Structures, Database Design, Programming in C++, Applied Econometrics, Linear Algebra

SKILLS & TECHNICAL TOOLS

- **Programming Languages**: Python, Java, SQL, R, C++, C, LaTex, Racket, Lean, HTML, CSS, Javascript
- **Technologies & Frameworks**: AWS , Linux, IntelliJ, Visual Studio, Snowflake, RStudio, DataGrip, Eclipse, Pycharm, Overleaf, DrRacket, MySQL, Appsmith, Octoparse, Flask, Docker, React
- **Speaking Languages**: Conversational Mandarin Chinese and Spanish

PROFESSIONAL EXPERIENCE

Starday Foods (CPG Startup)   Remote
Junior Data Science Software Developer   *Feb 2023 - December 2023*

Fully Automated TikTok Post Ingestion *| AWS (Athena, Glue, S3, Lake Formation), Python (Pandas, JSON)*
- Developed an automated data retrieval system that gathered 100K+ TikTok videos and details per day and efficiently structured JSON data into organized data tables for seamless analysis
- Established and managed a 12 table scalable data platform to handle inflow and storage of TikTok data using AWS tools including S3, IAM, Lake Formation, Glue Notebook, Glue Crawlers, and Athena

Similar Hashtag Generator *| AWS (Lambda, API Gateway), Python (Pandas, JSON), Dataiku*
- Created an applet that ingests a set of terms and generates a list of up to 250 similar hashtags per term by developing an AWS Lambda function that uses an external API to gather data
- Increased Starday's food and beverage ontology keyword database by 25% in 1 month by deploying the applet via AWS API Gateway and Dataiku DSS

Product Database & GUI *| Octoparse, Appsmith, Snowflake, SQL, Python (Pandas)*
- Employed Octoparse to web scrape over 100,000+ products from major retailers like Target and Sprouts, and organized and structured the collected data into distinct database tables
- Integrated the datatables into Snowflake and implemented an intuitive and user-friendly GUI, enabling seamless interaction and data querying for stakeholders

Open-end Survey Response Analyzer *| AWS Comprehend API, Python (Pandas, NLTK)*
- Utilized AWS Comprehend API to extract keywords from open-end survey responses and identify matches between extracted keywords and custom food ontologies
- Enabled processing of thousands of text-based survey responses to discover patterns and insights for researchers and product designers

Fake Survey Response Detector *| Python (Pandas, Numpy, Scikit-learn)*
- Developed a data preprocessing module that converts categorical survey responses into numerical values, enabling the utilization of machine learning algorithms for classification
- Trained and deployed several models on the transformed numerical data to predict the authenticity of survey responses, saving researchers hours of manual inspection for fake entries

ADDITIONAL

- **Sales Associate, Home Depot** (Dec 2021): Assisted 30-40 customers per day to find merchandise, collaborating with members of different departments
- **Interests**: Fintech, AI, Fantasy Football, Economics Podcasts, Cooking, Travel, Electronic Music