Thomas Docog

thomasdocog@gmail.com | 713-235-0072 | github.com/tdocog

Education

Texas A&M University, BS in Computer Engineering, Minor in Mathematics, GPA: 4.0/4.0   Expected May 2027

Experience

Quantitative Analyst Intern, Samuel A. Ramirez & Co. – New York, NY   June 2025 - Aug. 2025
- Engineered features from over 2.9 billion rows of loan records, transforming and cleaning data to be fit for modeling
- Achieved 430% speed-up using Polars in large-scale feature generation through vectorized computation
- Synthesized individual loan payment histories from historical delinquency data for predictive modeling
- Built a machine learning pipeline to predict loan outcomes from delinquency and payment history data, incorporating dimensionality reduction and feature engineering to improve model performance

Undergraduate NLP Researcher, Texas A&M University – College Station, TX   Sept. 2024 – May 2025
- Integrated benepar constituency parser into the spaCy pipeline to generate high-quality parse trees of podcast transcripts, enabling structured syntactic analysis of disfluent speech.
- Developed and implemented a transcription tree parsing algorithm to extract relevant tokens following disfluent speech patterns from 82,600+ podcast transcriptions
- Co-authored: Maria Teleki et al. "Evaluating Large Language Models for Disfluency Removal." Transactions of the Association for Computational Linguistics (under review), 2025

Teaching Assistant, Texas A&M University – College Station, TX   Aug. 2024 – Dec. 2024
- Taught introductory computing concepts, including programming fundamentals and data structures
- Conducted office hours to support students one-on-one, helping them with course material and problem sets
- Guided students through debugging and taught effective coding practices to improve code quality and efficiency

Web Development Intern, Summit Medical Staffing – Fremont, NE   June 2023 – July 2023
- Overhauled site interface, optimizing website navigation and enhancing overall user experience
- Diagnosed and resolved website bugs, optimizing functionality and minimizing user disruptions
- Revamped the mobile website, streamlining access to job listings improving user engagement and accessibility connecting nurses to 4,500+ healthcare opportunities

Projects

Cashlyn Finance, AI Financial Literacy Agent   Link To Repo
- Deployed an AWS EC2 Instance, configured Nginx as a web server to host Cashlyn Finance
- Configured the integration of a React frontend and Express.js backend using Nginx as a reverse proxy
- Awarded Second Place in the USAA Challenge at TAMUhack 2025 - Video Demo Link

Live ASL Translator   Link To Repo
- Utilized WebSockets for real-time streaming of webcam frames, capturing and broadcasting video data from the React frontend to the Flask backend for immediate processing
- Processed live camera frames using Python Imaging Library (PIL) to decode and preprocess image data, converting frames into tensor representations for inference
- Placed 3rd in Prelims, and 5th in Finals at TidalHack Fall 2024 - Video Demo Link

Awards, Skills & Interests

Languages: Python, C, C++, SQL, R, JavaScript, TypeScript, Scheme, Lua, Verilog

Technologies: PyTorch, scikit-learn, Polars, Pandas, Flask, React, BeautifulSoup

Awards: Eagle Scout, Hispanic Scholarship Fund Scholar, SC25 Scholar

Interests: Traveling, Language Learning, Fashion, Climbing, Japanese & Spanish Language, Scuba Diving, Soccer