David Zhan

301-323-5579 | dazhan@seas.upenn.edu | LinkedIn | Github | Website

Education

University of Pennsylvania 2026

Masters of Science in Computer Science, Bachelors of Science in Computer Science *GPA: 4.0/4.0*

- **Minors**: Mathematics, Data Science
- **Coursework:** Data Structures and Algorithms, Data Systems, Operating Systems, Machine Learning/AI, Natural Language Processing, Computer Vision, Cloud Computing, Deep Learning, Network Systems
- **Involvements** Penn Aerial Robotics (Software Lead), Machine Learning Research @ Penn, Penn Club Soccer, Penn Club Badminton, Math 1410 Teaching Assistant

Professional Experience

Machine Learning Engineer Intern May 2025– Present

Amazon *— PyTorch, AWS Bedrock/Sagemaker/EKS/EC2, Kubernetes* *Sunnyvale, CA*

- Core AGI team
- Implement new distributed training pipelines for mid-training using data augmentation for Nova base models

Software Team Lead August 2024 – Present

Penn Aerial Robotics — Python, ROS2, PX4, OpenCV, C++ *Philadelphia, PA*

- Spearheaded the development of a computer vision payload detection algorithm for UAVs utilizing binary thresholding
- Developed internal position representation from drone camera using RANSAC and ROS2 integration
- This high-accuracy system will compete in SAE Aero Design 2025.

Software Engineer Intern June – August 2024

Ventoscity — Flutter, NodeJS, Typescript, SQLite *College Park, MD*

- Redesigned cross-platform mobile app using Flutter, achieving 20% faster load times and 30% smaller app size.
- Integrated AI-powered food scanning with 90% accuracy for portion sizes and fiber content analysis.
- Conducted extensive user testing with 100+ external participants implementing over 25 usability improvements, resulting in a 15% reduction in app crash rates.

Research Intern June – August 2023

National Science Foundation — Swift, SwiftUI, C++ *College Park, MD*

- Developed a full-stack mobile app for emerging gut-health start up at UMD; pivotal in securing over $1M in VC funding.
- Submitted Publication to **Nature Medicine** (Third Author)

Mobile Developer Jan – Dec 2023

The Daily Pennsylvanian — React Native *Philadelphia, PA*

- Monetized DP+ mobile app for the first time in history, increasing total revenue of the Daily Pennsylvanian by 23%

Projects

News Source Classification Model — *BeautifulSoup, PyTorch, scikit-learn* May 2025

- Collected and cleaned **3,800+** headlines from Fox and NBC via BeautifulSoup-powered web scraping
- Built TF–IDF and word-embedding inputs for both single-channel and multi-channel TextCNN architectures
- Optimized hyperparameters to reach **82%** accuracy and **79%** F1-score—on par with a fine-tuned BERT baseline

Data Labeler — *React, Tailwind, FastAPI, Supabase, MEGA, AWS, Render* December 2024

- Dense captioning service to be used for training custom image models
- Implemented custom quality control, aggregation metrics; Amazon MechTurk for crowdsourcing data and crowd payment
- Captioned **400+** images, with a **210%** increase in words and **50%** in relevance compared to LLM generated captions

Blip — *Next.js, Tailwind, TypeScript, ConvexDB, Clerk, OpenAI, Cerebras* November 2024

- Short-form audio social media platform encouraging micro-learning during transition periods of the day
- Out of **500+** participants competing, winner of **PennApps XXV: Best Entertainment Hack**

Part-Of-Speech Tagger — *Python, numpy* September 2024

- Hidden Markov Model implementation using **Viterbi, Beam Search and Greedy** inference methods for POS tagging
- Using English Penn Treebank dataset for training; achieved a **96.3%** accuracy on randomized test set
- Model ranked top 2 in a 400+ graduate student class

Awards

Submission to Nature Medicine (Third Author) 2025

Smart Underwear: A Novel Wearable for Long-Term Monitoring Of Gut Microbial Gas Production Via Flatus

PennApps XXV: Best Entertainment Hack 2024

AIME Qualifier 2022

National Merit Scholar 2021

Skills

Languages: Python, Java, C, Swift, Dart, JavaScript, Typescript, SQL

Frameworks: React, NodeJS, NextJS, Tailwind, BootStrap, Flutter, FastAPI, Flask, Django

Tools: Docker, Kubernetes, AWS

Machine Learning: SageMaker, Bedrock, TensorFlow, HuggingFace, PyTorch