

David Zhan

301-323-5579 | dazhan@seas.upenn.edu | [LinkedIn](#) | [Github](#) | [Website](#)

EDUCATION

University of Pennsylvania

May 2026

Bachelors of Science in Computer Science, Masters of Science in Computer Science

GPA: 4.0/4.0

- **Minors:** Mathematics, Data Science
- **Coursework:** Machine Learning, Cloud Computing, Computer Architecture, Distributed Systems, Deep Learning
- **Involvements** Penn Aerial Robotics (Software Lead), Machine Learning Research, Penn Club Soccer, Teaching Assistant

PROFESSIONAL EXPERIENCE

Software Engineer Intern

May 2026 – August 2026

Google

Waterloo, CA

- Summer 2026 (GCP: Applied AI)

Machine Learning Engineer Intern

May 2025 – August 2025

Amazon — *PyTorch, AWS Bedrock/SageMaker/EKS/EC2, Kubernetes*

Sunnyvale, CA

- Designed and deployed a fully automated distributed data-distillation pipeline on AWS SageMaker and EKS for large-scale pretraining augmentation, increasing augmentation throughput by **30x**
- Developed a reasoning-centric data-augmentation methodology leveraging teacher-student distillation via AWS Bedrock, yielding **+2.6 pp** BBH-CoT and **+1.2 pp** MMLU-CoT improvements on Nova, Amazon's flagship LLM, using only **10 K** curated samples.

Software Team Lead

August 2024 – Present

Penn Aerial Robotics — *Python, ROS2, PX4, Gazebo*

Philadelphia, PA

- Led development of autonomous UAV vision systems, including a payload detection pipeline leveraging binary thresholding and contour analysis for real-time onboard inference.
- Directed simulation and hardware-in-the-loop testing in Gazebo and PX4 SITL to validate autonomous flight performance.
- Competed at SAE Aero Design West 2025; placed **2nd place nationally** in advanced class with autonomous flight routines contributing to mission success.

Software Engineer Intern

June 2024 – August 2024

Ventosity — *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 10 usability improvements, resulting in a **15%** reduction in app crash rates.

PROJECTS

Cloud Search Engine

 — *Java, AWS EC2, Distributed Systems*

Dec 2025

- Implemented a **web server**, **distributed key-value store**, and a “mini-Spark” engine using only the `java.net` library
- Crawled and indexed **100k+** **webpages** on an AWS EC2 instance, storing HTML content in a fault-tolerant KVS
- Developed end-to-end processing pipeline for **TF-IDF**, **PageRank**, inverted indexing scoring to support keyword search
- Achieved **sub-second query latency** for terms via table partitioning, caching, and optimized distributed job scheduling

News Source Classification Model

BeautifulSoup, PyTorch, scikit-learn

Spring 2025

- Collected and cleaned **3,800+** headlines from Fox and NBC via BeautifulSoup-powered web scraping
- Built TF-IDF representations and word-embedding inputs for both single + multi-channel TextCNN architectures
- Conducted hyperparameter optimization to reach **82%** accuracy and **79%** F1-score—eq. to a fine-tuned BERT baseline

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**

PUBLICATIONS AND AWARDS

Biosensors and Bioelectronics: X (Second 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

SKILLS

Languages: Python, Java, C++, Swift, Dart, JavaScript, SQL, NoSQL

Frameworks: React, NodeJS, NextJS, Tailwind, TypeScript, ThreeJS, Bootstrap Flutter, FastAPI, Flask, Django

Tools: Supabase, DynamoDB, AWS, Convex, Neo4j

Machine Learning: pandas, numpy, scikit-learn, TensorFlow, HuggingFace, PyTorch, OpenCV