Rohan Bafna

Software Engineer

rbafna.com
rbafna@umich.edu
864.908.2002

Education

University of Michigan   Class of 2025

- B.S.E in Computer Engineering: Data Structures and Algorithms • Web Systems • Computer Architecture • Discrete Math

Experience

Software Engineer Intern, Bold Metrics   May-Aug 2023

- Rewrote backend using FastAPI with Python to generate auto-docs and improve response time, while implementing custom validations and logging, resulting in a 10% decrease in latency & ~15% reduction in failed calls for all clients.
- Developed an automated testing script using Python, Selenium, and AWS to validate API calls and check if Smart Size Chart was returning correct recommendations for thousands of product pages across 12 clients and 5 unique APIs.

Software Engineer Intern, Bold Metrics   May-Aug 2022

- Created a CLI with Python that enabled all 12 clients to upload hundreds of new garment size charts every month to PostgreSQL DB to automate training, increasing training frequency by 10% and model accuracy by 5%.
- Implemented a new prediction-testing script that checks the predicted 'best size' against training data, ensuring the model returns correct predictions relative to the inputs, resulting in a 20% reduction in API calls to QA.
- Built new size charts using past body measurements and sorted garments into unique categories based on sizing, increasing the number of supported clothing types and garments to deliver 2.5x more recommendations for our 3 largest clients.

Machine Learning Research Intern, University of South Carolina   Jan - Sep 2021

- Designed image ML models using Tensorflow and applied 2, 4, 8, and 16 bit quantization to each layer. Distributed them across microcontrollers and GPUs, and pinpointed the distribution that yields maximum accuracy - deployed onto glasses for the blind.
- Deployed models that utilized quantization-aware training and post-training quantization, and created a hybrid model that effectively reduced model size by ~70% while retaining 75% of the accuracy of the full 64-bit model.
- Implemented a novel Bayesian Optimization algorithm that optimizes ML accuracy around unique bit-quantization for each layer.

Machine Learning Intern, Clemson University - ICAR   May-Aug 2020

- Created a simple MNIST solver from scratch with four hidden layers and the ReLu activation function, yielding a 94% accuracy.
- Experimented with popular machine learning models in both MatLab and TensorFlow in Python to conceptually understand the strengths and weaknesses of each model, while digging into unique model structures to implement in our own custom designs.
- Worked together with teammates to present our learnings of each new generation of machine learning models - from handwriting recognition to object and video tracking, covering neural networks, CNNs, ResNets, and RNNs.

Projects

Algorithmic LoFi Music Generator   Present

- Designing an algorithm-based online lofi music generator, leveraging Tone.js for sound synthesis and Node.js for backend processing, enabling dynamic creation of 2.5 to 3-minute unique tracks accessible to hundreds of people at the same time.
- Comparing this algorithmic approach to AI tools to see what advantages and drawbacks exist in each method for future use.

Informate   Summer 2021

- Built a cronjob (used Flask, deployed on Heroku) that regularly scrapes popular news websites with Python and BeautifulSoup4 for their top 20 articles each day and stores them in a Mongo database.
- Developed a landing page with React that pings an API to access the database and display metadata from the articles including headlines, summaries, and URLs. Displays the latest 10 articles from each website.

Tools • Python, JavaScript, C++, Java, JupyterLab, Flask, Django, NumPy, Node.js, React, Tensorflow, AWS, MongoDB, SQL