YOHAN V PANDYA

503-268-2756 | yvpandya@wisc.edu | https://yohanpandya.github.io/ | https://www.linkedin.com/in/yohan-pandya-4401b4255/

EDUCATION

University of Wisconsin-Madison **Expected May 2025**
Bachelor of Science in Computer Science GPA: 3.81/4.00, Dean's List
Related Coursework: Intro to Algorithms, Data Structures, Discrete Mathematics, Computer Systems, Intro to Big Data Systems, Intro to Machine Learning, Linear Algebra, Multivariable Calculus

TECHNICAL SKILLS

Programming Languages: Java, Python, C, Linux, SQL, HTML, CSS, React.js
Frameworks/Tools: Git, PyTorch, Numpy, TensorFlow, Docker, Pandas, gRPC, Kafka, CQL, Apache Spark, BigQuery

WORK EXPERIENCE

***Intern*, ShenZhen JiangYun Intelligence** **January 2023 - May 2023**

- Actively optimized performance and memory usage of a stable diffusion model through weekly fine-tuning of critical hyperparameters, such as sampling method, batch count, and batch size.
- Utilized Google Colab, Jupyter Notebooks, and a local host to execute the model, while achieving expertise in diffusion model tools like DreamBooth and LORA (tools that are geared towards optimizing diffusion model performance while minimizing processing power).

***Topocoder,* Traffic Technology Services** **June 2021- August 2022**

- Utilized Google Earth Pro to gather and analyze traffic intersection data, contributing significantly to the integration of traffic signal timers in Audi vehicles.
- Conducted in-depth data analysis on government documents to extract signal phases, effectively incorporating this information into the company's databases.
- Engaged with cloud-based technologies while ensuring the quality and accuracy of topographical coding completed by other team members.

PROJECTS

***Research Assistant*, UW Madison School of Materials Science & Engineering** **June 2023 - Present**

- Studied MIT paper on Crystal Diffusion Variational Autoencoder (CDVAE) which identified stable materials with constrained chemical combinations.
- Effectively implemented CDVAE model on Center for High Throughput Computing's GPUs and performed repeated testing and iteration to resolve compatibility issues.
- Improved model's accuracy from 68% to 75% by adjusting batch size. Generated over 3500 stable materials, and currently am aiming to continue optimizing models by increasing time and space efficiency.

Personal Projects **September 2022 - Present**

- AI Covid Predictor
    - Used PyTorch to create a regression model that can predict how many deaths there will be for a Wisconsin census tract, given the number of people who have tested positive, broken down by age.
- ModelServer
    - Developed a full-stack machine learning model serving infrastructure using gRPC, Python, and Docker.
    - Created a gRPC-based server for real-time predictions, improved efficiency by using a multi-threaded client as well as an LRU cache. Built a Docker container for easy deployment.
- Kafka Weather Data
    - Implemented Kafka-based data transfer system using protobuf serialization for weather data transmission, configuring 4 partitions and a single topic with replication factor 1 and exactly-once semantics.
    - Created consumer that generates visual data representation reflecting the temperature trends across the specified months, enabling easy comprehension and analysis of the gathered information.