Joshua Chang

joshuachang.me | j79chang@uwaterloo.ca | joshjchang | JoshChang8

SKILLS

Languages: Python, SQL, R, Java, C++, JavaScript, BASH, Git
Tools and Technologies: Open AI, Langfuse, Scikit-Learn, PyTorch, Stable Diffusion, Pandas, NumPy, AWS, Git
Technical Skills: Natural Language Processing, Prompt Engineering, Generative AI, Statistical Analysis, Data Collection

EXPERIENCE

Machine Learning Engineer | Keplar.io
January 2024 – April 2024

- Developed a **Constitutional AI Evaluator LLM** by implementing **Prompt Engineering** methodologies, using a set of guiding principles for data quality assessment, emulating Reinforcement Learning from Human Feedback (**RLHF)**
- Leveraged **Optimization by Prompting** (OPRO) techniques to enhance LLM performance, increasing **precision** by **~10%**
- Constructed a **benchmarking suite** of quantitative and qualitative tests to assess the performance of LLMs
- Utilized **image generation models** with **ControlNet** to generate product design variations, providing canny and depth image conditioning inputs to refine diffusion models

Data Scientist | Quantolio Financial Technologies
May 2023 – Aug 2023

- Led end-to-end feature development, conceptualizing, and implementing an interactive **UI** using **Streamlit** to showcase diverse **portfolio optimization algorithms** and visualizing cumulative returns through dynamic graph visualizations
- Containerized the application using **Docker** and orchestrated deployment on **Heroku**, ensuring accessibility for clients
- Communicated with clients and investors, effectively conveying **product features** and **value propositions** to foster strong relationships and drive investment interest

Data Engineer and Computer Vision Research Assistant | University of Waterloo
Jan 2023 – May 2023

- Utilize **YOLOv5 object detection** to detect the pitcher and analyze the **3D Pose Estimation** of their biomechanics to improve performance and aid in injury prevention
- Aggregate and clean Hawkeye baseball data from **SQL databases** and **AWS S3** storages to train various models

Site Reliability Engineer | The Globe and Mail
Sept 2022 – Dec 2022

- Created an **AWS Lambda function** using various AWS services, such as CloudTrail, EventBridge, and IAM to **scale Kinesis data streams** to **eliminate under and over utilization**, saving costs on resources
- Migrated **10+ AWS Aurora dashboards** from AWS Cloudwatch to Datadog, collecting **50+** metrics of key health and performance indicators to reduce costs and resources
- Created monitoring and alerting for **250+ Airflow DAGs** to catch **30+** anomalies by creating a Datadog metrics monitor provisioned using **Terraform**

PROJECTS

Spotify Song Recommender
Python, Pandas, NumPy, Jupyter

- Designed and developed a Spotify recommendation application conducting comprehensive analyses on user's playlists utilizing the **Spotify API** and delivering personalized recommendations based on **mathematical similarity algorithms**
- Collected, processed, and refined dataset of **1M+ songs** using **Pandas** and **NumPy** to construct **item-feature matrices** for mathematical similarity measures
- Implemented **multiprocessing** techniques to enhance efficiency of CPU-bound operations, resulting in **35%** reduction in computation time
- Orchestrated containerization of web application using **Docker** and enabled **automatic deployment** on Google Cloud Run with Google Build triggers for **continuous integration**

EDUCATION

University of Waterloo
Waterloo, ON

Candidate of Bachelor of Applied Science in Systems Design Engineering, Artificial Intelligence Option
- Relevant Coursework: Data Structures and Algorithms, Human Factors in Design