Arjun Sarao
https://arjunsarao.dev/
arjun.sarao@uwaterloo.ca
https://github.com/DarkHawk727

Work Experience

Generative AI Engineer | *City of Hamilton ITS* Sept. 2024 – Dec. 2024

- Developed internal RAG chatbot web app using `embedchain`, `AzureOpenAI`, `streamlit`, and `chromadb`; Deployed to 240 users.
- Researched and presented business use-cases and capabilities of cutting-edge generative AI tools to stakeholders.

Jan. 2024 – Apr. 2024

- Developed RAG chatbot web app using `embedchain`, `AzureOpenAI`, `streamlit`, and `chromadb`.
- Researched and presented business use-cases and capabilities of cutting-edge generative AI tools to stakeholders.

Full-Stack Web Developer | *Starai Tutoring* May. 2022 – Mar. 2023

- Developed website using `JavaScript (React.js)` in conjunction with `Bulma.css`, hosted it on `GitHub Pages`.
- Implemented sign-up form submission via `Firebase` cloud-function and SMS notifications with `Twilio`.

Projects

ViT-tinygrad | *Vision Transformer Paper Implementation* Jan. 2025

-
-
-

TabTransformer-tinygrad | *Tabular Transformer Paper Implementation* Oct. 2024

- Reimplemented *TabTransformer: Tabular Data Modeling Using Contextual Embeddings* paper from Amazon in `tinygrad`.
- Used Hyperparameter Optimization for number of MLP hidden layers; Added custom SELU activation function; Implemented early stopping.
- Compared performance of TabTransformer against MLP on `insurance_co` dataset; Achieved a *74.4%* AUC performance (4.7% improvement, paper result).

SeeFood | *Image Recognition Java App with TensorFlow CNN* Jan. 2022

- Used `BeautifulSoup4` and `bing_image_downloader` `Python` packages to scrape ~3.5k images from Kaggle and Bing to create image dataset.
- Developed `CNN`s using `TensorFlow`; Achieved validation accuracy of approximately *90%*.

Education

Honours Bachelor of Computer Science (CO-OP) | *University of Waterloo* Sept. 2022 – present

- **Relevant Coursework**: Introduction to Database Management (**SQL**), Data Structures and Algorithms (**C++**), Compilers (**C++**), Linear Models, Statistics, Linear Algebra II, Calculus III, Object-Oriented Software Development (**C++**), Elementary Algorithm Design and Data Abstraction (**C**), Introduction to Database Management (**SQL**), Designing Functional Programs (**Racket**)

Self Learning

Machine Learning Specialization | *Stanford Online + Coursera* Aug. 2022

- Supervised Machine Learning: Regression and Classification (**Linear Regression**, **Logistic Regression**)
- Advanced Learning Algorithms (**NN**s, **Decision Trees**)
- Unsupervised Learning, Recommenders, Reinforcement Learning (**Clustering**, **Anomaly Detection**, **Collaborative Filtering**, **Deep Q-Learning**)

Deep Learning Specialization | *Deeplearning.ai + Coursera* Jul. 2022

- Neural Networks and Deep Learning (**NN**s)
- Improving Deep Neural Networks: Hyperparameter Tuning, Regularization, and Optimization
- Convolutional Neural Networks (**CNN**s)
- Sequence Models (**RNN**s, **GRU**s, **LSTM**s, **Transformers**)

Technical Skills

Developer Tools: `Linux (Ubuntu)`, `Git`, `Github Actions`, `Google Cloud Platform (Firebase)`, `VSCode`
Libraries: `tinygrad`, `PyTorch`, `NumPy`, `pandas`, `Matplotlib`, `DSPy`, `langchain`, `streamlit`, `sympy`, `supabase`