Do Chung

+1 (647) 624 1623
dchung@uwaterloo.ca
github.com/dchung1209
linkedin.com/in/dchung1209

Technical Skills

Programming Languages: Python, SQL, R, C, C++, JavaScript, Bash/Shell
Frameworks: PyTorch, TensorFlow, Scikit-Learn, LlamaIndex, Langchain, LangGraph, Flask
Libraries: Hugging Face, Pandas, NumPy, Matplotlib, Seaborn
Tools: Git, Tableau, Docker, AWS

Experience

Assistant Chief May 2025 – Present
Republic of Korea Army *South Korea*

Machine Learning Engineer (Co-op) Jan 2025 – April 2025
BoardRoom *Toronto, Canada*

- Engineered an end-to-end data enrichment pipeline using **Python** and **MongoDB** to identify potential customers and verify 1,000+ executive profiles via **REST APIs**
- Built an AI Agent workflow that automates lead discovery and personalized mailing, cutting the lead-to-outreach cycle by 50%

Data Scientist (Co-op) Sep 2023 – Dec 2023
Intellicode *Suwon, South Korea*

- Developed state-of-the-art deep learning models **(CNN, Attention, LSTM)** for speech emotion recognition, incorporating techniques from recent research papers
- Implemented a **Generative Adversarial Network (GAN)** for data augmentation, improving model stability and boosting performance metrics in emotion recognition tasks
- Conducted analyses of benchmark techniques in speech emotion recognition, leading to the integration of **multimodal models** in speech emotion recognition systems

Projects

PDF Chatbot | *OpenAI API, LlamaIndex, LangChain, MySQL, Docker, AWS EC2, Streamlit*

- Developed an **Agentic Retrieval-Augmented Generation (RAG)** chatbot supporting multiple **LLMs** for complex PDF analysis; containerized with **Docker** and deployed on **AWS EC2**
- Implemented a **RAG** pipeline for document search using **LlamaIndex** and integrated a **reranker**, improving query result relevance

Webcomic Analytics | *Beautiful Soup, OpenAI API, Scikit-Learn, Hugging Face, Tableau, Pandas*

- Scraped and refined metadata, visualizing genre and series trends in **Tableau** to highlight popular categories
- Developed a hashtag generator for webcomics using the **OpenAI API**, improving hashtag quality and semantic relevance through **prompt engineering** and **CLIP** integration
- Designed a recommendation system using **K-Nearest Neighbors (KNN)** and **knowledge graph** techniques for collaborative filtering

Electrocardiogram (ECG) Signal Detection | *PyTorch, AWS EC2, Scikit-Learn, XGBoost, Flask, Seaborn, Pandas*

- Processed ECG data and deployed models for anomaly detection via a **REST API** using **Flask** on **AWS EC2**
- Implemented and optimized deep learning models, including **Bi-LSTM** and **1D CNN-LSTM**

Bird Species Recognition | *PyTorch, Matplotlib, Pandas*

- Developed and Find-tuned a **Vision Transformer (ViT)** model from scratch for classifying 20 bird species
- Optimized model training time on a large dataset through **parameter tuning** and **mixed precision**

Education

University of Waterloo Sep 2022 – May 2027
Bachelor of Mathematics, Honours Statistics (Co-op), Minor in Computing *Waterloo, ON*