

# Zeyu (Dennis) Wang

## PROFESSIONAL SUMMARY

Master of Information Technology (AI) student at the University of Melbourne (graduated) with a strong foundation in machine learning, software development and AI infrastructure. Solid experience in building and evaluating machine learning models, full stack software development with competence spanning all phases of the development life cycle. Passionate about Large Language Models (LLMs), cloud platforms and leveraging AI in practical solutions

**Phone:**

0414035766

**Email:**

denniswang0722@gmail.com

**Address:** U2015, 65

Dudley Street, West  
Melbourne, 3003

**Linkedin:**

www.linkedin.com/in/zeyu-wang-a18206268 [🔗](#)

## RESEARCH / INTERN EXPERIENCE

### Sound Event Tagging

MicroCode Electronics Co, Ltd, Nanjing, Jiangsu

Jan 2025 - Present

- Provided with real-time recorded raw audio files, conducted **audio analysis** with Librosa, built and trained convolutional neural network for audio tagging using **Keras** and audio specific features (MFCCs)
- Complied with clients' requirements, converted keras models to RKNN models for **deployment** on NPUs, validating **inference consistency** on embedded Linux system (Ubuntu) with scripted test cases; **automated** the conversion and testing workflow and documented full steps for client's further testing
- Currently experimenting with more complex models in processing more datasets

### Advanced RAG and AI Agent (Independent Research) [🔗](#)

Self Employed

May 2025 - Present

- Adapted both **document-based retriever** (BM25) and **vector retriever** (sentence embedding + vector database) for document retrieval; evaluated the pros and cons for **different vector databases** among FAISS, ChromaDB and Weaviate
- Utilised **huggingface** model for **speech recognition** that allowed RAG to incorporate video or audio inputs
- Implemented **query optimization** techniques for more accurate retrieval, as well as **contextual compression** techniques to improve the quality of retrieval documents
- Used **ColQwen2** as document retriever to filter image-like documents and **SmolVLM** for answer queries
- Built a plan-and-execute **AI agent** from scratch with **LangGraph** and **OpenAI API**, with the ability to use customized **tools** for **function calling** (MCP) and generate analysis report; referenced from **OpenManus**

### Conversational AI Research Assistant [🔗](#)

Melbourne Data Analytics Platform (MDAP)

Jul 2024 - Nov 2024

- Worked in a **cross-functional team**, understood the existing codebase through self-developed a similar legal document Q&A web chatbot with **RAG** using **Streamlit** and **Langchain**
- **Basic Pipeline:** Connected **OpenAI APIs** to the application; stored documents in **Chroma vector database** with vector embedding similarity search for document extraction; extended the chatbot with user able to switch to **Llama model**, connection enabled with llama.cpp

## RELEVANT PROJECTS

### Text Generation Classification [🔗](#)

Statistical Machine Learning Course

First Semester of 2025

- Given generated texts as lists of tokenised indices, applied BoW and bi/tri-grams BoW for processing; with **data cleaning**, **feature engineering** and data augmentation steps such as masked indices, low frequency words removal, combined with calculated numeric features
- Adapted out-of-fold cross validation; SMOTE and **test time augmentation** as techniques for dealing with **class imbalance**; **multi-task learning** for 2 datasets from separate domains
- Ensemble SGD models (Logistic regression, SVM, Naive Bayes) and implemented **auto encoder** in **PyTorch** for dimension reduction; achieved **2<sup>th</sup> position** on the public leaderboard

### LLM Training and Fine-tuning [🔗](#)

Kaggle Mining Misconceptions in Maths + WSDM Chatbot Arena Competitions

Nov 2024 - Dec 2024

- Fine-tuned Gemma-2-9b model and Qwen-2.5-14b model with **QLoRA** using **Huggingface transformers** and **training** libraries; Enhanced understanding in **PEFT training** and **quantization** of LLMs
- Mastered skills in training **embedding models** and **reranker models** such as Qwen3 and BGE models; **synthetic data generation** with LLM and prompt engineering to boost performance

## Image Classification

Computer Vision Course

Second Semester of 2024

- Performed **image segmentation** and background removal on ShapeStack dataset, applied **transfer learning** with pre-trained models to make **classification**
- Enhanced understanding in leveraging **pre-trained** VGG, Inception, ResNet from Keras; utilised Pytorch for building InceptionV4 model, and **YOLO-v9** for shape segmentation and classifying the stacking height

## Natural Language Inference

Natural Language Processing Course

First Semester of 2024

- Applied data preprocessing using **NLP techniques** (NLTK, sklearn) on 1 million evidences and thousands of claims
- Tried several **vectorizing** techniques such as Tf-IDF, BM25, BoW etc; eventually constructed a bi-encoder model with padding, embedding and **LSTM** layers using **Keras and TensorFlow**; combined with top-k embedding similarities; achieved **7th position** in Codalab competition out of 120 teams

## Full Stack Web Development

IT Project Course

Second Semester of 2023

- Built a Student Request Portal (allow students send queries easier) in a cross-functional team following the **Agile** methodologies; portal website integrated seamlessly with the university learning system
- **Backend Development:** Implemented **REST APIs** with **Spring Boot (Java)**, automated deployment and code changes via **CI/CD pipelines (Git)**
- **Database Management:** Managed **relational databases** through MySQL workbench and **MyBatisPlus** framework, which simplifying database access and take control of the SQL queries
- **Testing:** Wrote **unit tests** under Spring boot framework; conducted **integration tests** through Postman for APIs and features implemented
- **Frontend development:** Developed staff interface with **Vue3.js** and **AntDesignVue** component library

# EDUCATION

---

## Master of Information Technology | Artificial Intelligence

University of Melbourne, Melbourne

Mar 2024 - Jun 2025

- **Relevant Coursework:** Natural Language Processing, Computer Vision, Statistical Machine Learning, Software Processes and Management, Software Project
- **Weighted Average Mark:** 80 | First Class Honours

## Bachelor of Science | Computing and Software System


University of Melbourne, Melbourne

Mar 2021 - Dec 2023

- **Relevant Coursework:** Machine Learning, Computer System, Design of Algorithms, IT Project, Software Modelling and Design, Object-Oriented Software Development
- **Weighted Average Mark:** 83 | First Class Honours

# SKILLS

---

- **Languages** (in order of proficiency): Python | Java | SQL | R | JavaScript | C
- **Software Project Tools:** Git, Confluence, Trello, Agile
- **Coding Tools:** Cursor, Copilot, VsCode
- **Software Design Frameworks:** Spring Boot, Flask, React, Vue3, Node.js
- **Machine Learning / AI Libraries:** TensorFlow, PyTorch, Keras, HuggingFace, Scikit-learn
- Competence in **cloud platforms** (GCP, Amazon Bedrock, Cloudflare); investigated deployment of LLM-based application through **Cloudflare Workers and AI bindings** 
- Complete fluency in both **English** and **Mandarin**

# VOLUNTEER

---

## VCE Maths Tutor

Step Education + QS School

Jan 2021 - Nov 2024

- More than 3 years of experience teaching students across diverse age group, high school students been taught all achieved a VCE of 90+
- Coupled with exceptional ability to communicate with different age group of people, as well as strong skill to **deliver presentations** in multiple occasions