Lim Zhe Xun

Email: zlim116@e.ntu.edu.sg
Portfolio: https://hersheyzinc.github.io/
Mobile: +65 96997574
Github: github.com/HersheyZinc

Education

- **Nanyang Technological University**
  Bachelor of Computer Science - Data Science and Artificial Intelligence; GPA: 4.67 *Aug 2021 - May 2025*
  Courses: *Artificial Intelligence, Machine Learning, Neural Networks and Deep Learning, Computer Vision, Data Analytics*
- **Hwa Chong Junior College**
  GCE 'A' Level - 88.75/90 *Jan 2017 - Dec 2018*

Skills Summary

- **Languages**: Python, Java, C++, R Studio, SQL
- **Frameworks**: Pytorch, Tensorflow, Scikit, NLTK, Langchain, HuggingFace, OpenCV, Pandas, Flask, Streamlit

Work Experience

- **Digital Trust Centre (NTU)**
  Undergraduate Researcher *May 2024 - Dec 2024*
  - **LLM Machine Unlearning**: Conducted comprehensive survey on machine unlearning methods for LLMs.
  - **LLM Finetuning**: Reproduced machine unlearning method on Llama 3 via relabelled gradient descent.
- **Robert Bosch ASEAN**
  Machine Learning Intern *May 2023 - Dec 2023*
  - **Synthetic Data Generation**: Developed synthetic 3D object data generation via Blender.
  - **Object Detection**: Finetuned YOLOv4 models to compare performance between real and synthetic datasets.
  - **LLM**: Integrated LLaMA 2 and LangChain to develop a Retrieval-Augmented Generation (RAG) system for efficient querying of in-house documents.
- **DSO National Laboratories**
  Machine Learning Intern *May 2022 - Aug 2022*
  - **Video Segmentation**: Integrated FCAnet and STCN models into advanced video segmentation pipeline.
  - **System Architecture**: Designed application architecture using FastAPI and React.
- **NAS Daily**
  Localization Intern *Mar 2021 - Jul 2021*
  - **Automation**: Developed Python applications that streamlined workflows, reducing localization workload by 80%.
  - **Computer Vision**: Developed computer vision techniques to enhance the quality of translated video subtitles.

Projects

- **Autonomous PowerPoint Agent (LLM)**: Designed and implemented a Large Language Model (LLM)-powered agent framework to autonomously create, manage, and operate PowerPoint presentations. Enabled seamless interaction using natural language queries, allowing users to generate slides, edit content, rearrange layouts, and navigate presentations via chat interface. Tech: Python, OpenAI, Streamlit
- **Sales Lead Generation (LLM, Prompt Engineering, Web Scraping)**: Developed a web scraping pipeline using BeautifulSoup to collect potential prospect business websites and leveraged LLMs to automatically identify their Applicant Tracking System (ATS), optimizing lead generation and recruitment insights. Tech: Python, OpenAI, BeautifulSoup
- **Image Segmentation for Precision Farming (Image Processing, Image Segmentation)**: Collected and labeled an image segmentation dataset of hydroponic crop growth and benchmarked multiple image segmentation models for performance evaluation. Tech: Python, OpenCV, PyTorch, OpenMMLabs
- **DEEP Project (Image Processing, Image Classification)**: Collected and labeled an image dataset of various screw types, trained a custom CNN model from scratch with 95% accuracy, and deployed it as a Telegram bot for user-friendly access. Tech: Python, Tensorflow, OpenCV, Telegram API.
- **ECG Heartbeat Classification (Time Series)**: Developed a pipeline to classify ECG signals from the MIT-BIH Arrhythmia Database for early arrhythmia detection. Conducted an ablation study on RNN, CNN, and hybrid CNN-LSTM architectures to optimize model performance. Tech: Python, Pytorch
- **Urban Mobility Analytics System (Data Mining, Data Analytics)**: Developed a data-driven urban planning system that features real-time visualization of urban mobility patterns, interactive analysis of POI distributions and relationships, and predictions based on historical movement data. Tech: Python, Trackintel, Streamlit
- **Multilingual Analysis of Low-Resource Language (NLP, Data Analytics)**: Developed and evaluated automatic part-of-speech taggers for Pnar, a low-resource Indian language, by processing datasets, performing exploratory data analysis, and incorporating rule-based improvements.. Tech: Python, NLTK, Streamlit
- **Telegram Chess Bot (Automation)**: Telegram bot to automatically generate and share daily Chess and Othello puzzles in group chats, featuring adjustable difficulty levels and .gif-based puzzle solutions. Tech: Python, Telegram API