

Ho Dang Cao

✉ dangcaoho151202@gmail.com | 📞 097 2367 154 | 🌐 [HoDangCao](https://HoDangCao.com) | 🌐 [leetcode/HoDangCao](https://leetcode.com/HoDangCao)

EDUCATION

University of Science, Student in Information Technology.

Oct 2020 – Oct 2024

- Major: Data Science
- GPA: 8.18/10

Achievement: Received an academic **excellence scholarship** in the 3rd semester of 2020-2021.

English Certificate: VSTEP - B2 level.

TECHNICAL SKILLS

- **Programming languages:** Python, SQL, C/C++
- **Frameworks/ Platforms:**
 - Familiar with Machine/Deep Learning frameworks such as Pytorch, Tensorflow, Scikit-Learn, LangChain, HuggingFace.
 - Data manipulation: Crawling (Selenium, BeautifulSoup), Processing (Numpy, Pandas, Excel), Statistics (Statsmodel), Big data (Hadoop, Pyspark), Visualization (Matplotlib, Seaborn, Tableau, Power BI).
- **Technologies:** Git, Docker, MSSQL, MongoDB

PROGRAMMING PROJECTS

LARGE LANGUAGE MODELS SYSTEM

- **Team size:** 1 member
 - [Details](#)
- **Description:** This project involves the development and exploration of **Large Language Models (LLMs)**, focusing on key components such as building and fine-tuning models, data preprocessing, and evaluation. It includes hands-on implementation of **Retrieval-Augmented Generation (RAG)**, **Chain of Thought (CoT) prompting**, and the creation of a **Llama 3 model** from scratch. The project also delves into advanced **fine-tuning strategies** using **PEFT** techniques like **LoRA** and **Adapters**, showcasing practical expertise in LLM architecture, optimization, and reasoning.
- **Technologies:** Pytorch, Numpy, Pandas, Spacy, Sklearn, Transformers, Matplotlib.
- **Responsibilities:** Build components of a LLMs system step by step from scratch: Llama 3 model, Retrieval-Augmented Generation (RAG), Chain of Thought (CoT).
- **Products:** 2 websites that allows users to
 - Upload pdf files and return information corresponding to questions from users - [Details](#).
 - Upload to DB, return SQL code and user query results in seconds without exposing data to LLM - [Details](#).

OPENCV PROJECTS

- **Team size:** 1 member
 - [Details](#)
- **Description:** Explore fundamental computer vision tasks using OpenCV in this project. From edge detection to object tracking, this repository showcases a range of applications such as Hand Gesture Recognition, Lane Detection for Self-Driving Cars, Augmented Reality (AR), and Real-Time Facial EXpression Recognition (FER) using CNN and Haar Cascade classifiers. Each demo focuses on a unique aspect of image processing, providing both theoretical insights and practical coding implementations.
- **Technologies:** OpenCV, Numpy, Pandas, Keras, Sklearn, Matplotlib.
- **Responsibilities:** Implemented various projects, including object tracking, lane detection, gesture recognition and AR; Built and trained a CNN model for FER, integrating Haar Cascade classifiers for face detection.

TEXT TO VIDEO

- **Team size:** 1 member
 - [Details](#)
- **Description:** Developed a Text-to-Video AI Model from scratch using Generative Adversarial Networks (GANs) to convert textual prompts into dynamic video frames. The project involved creating a custom dataset and training the model to generate realistic videos with moving shapes based on text inputs.
- **Technologies:** NumPy, PyTorch, Pillow, OpenCV.
- **Responsibilities:**
 - Designed and implemented a GAN architecture with Generator and Discriminator for video generation.
 - Developed a custom dataset and pre-processed data for training.
 - Built and trained the model using adversarial loss.
 - Evaluated model performance by generating video outputs from unseen textual prompts.