Trịnh Ngọc Các

(+84) 941-849-078 | cactrinh5@gmail.com | github.com/Ngoc-Cac
Ho Chi Minh City, Vietnam

PERSONAL INTRODUCTION

I am a fresh Data Science graduate at the University of Economics Ho Chi Minh City, with an interest in researching the application and integration of AI for practical purposes, mainly in the fields of Computer Vision and Signal Processing. Due to limited practical experience in the development and deployment of AI systems, I am seeking opportunities to gain hands-on experience in these areas while further honing my existing knowledge.

I mainly use: Git, Python, Lua, Bash

I have had experience with: Docker, Docker Clusters, Hadoop HDFS, Spark, LaTeX

EDUCATION

University of Economics Ho Chi Minh City
College of Technology and Design   Aug, 2022 – Mar, 2026
Bachelor of Science in Data Science

- Graduation Thesis: "A Framework for Radiology Image De-identification using Generative AI"
- Overall GPA: 3.86/4.0

WORK EXPERIENCE

<u>AAI Labs</u>, Engineering Department   Vilnius, Lithuania
Junior Researcher, Remote   Nov, 2025 – Present
Intern Researcher, Remote   Aug, 2025 – Nov, 2025

- **Scope**: Responsible for the prototyping and development of AI-integrated tech solutions.
- **Details**:
  - At AAI Labs, we use **Scrum**, an Agile framework, to manage and plan our projects. During development, we follow the principles of **TDD**, focusing on writing tests before any development takes place.
  - I mainly take part in the investigation, analysis and development of a first version for said solutions. This is where I leverage most of my DA/ML/DL knowledge to solve problems both me and my team encounter throughout the project.
  - I participated in the deployment of an traffic sign detection model for a dash-cam system, using YOLO-based models.
  - I also investigated and developed algorithms to generate discount trends based on historical sales data.
- **Technologies Used**: Git, Google Cloud Storage, Flask, Flask-Restful, Pytest, Numpy, Pandas, OpenPyXL, Scikit-learn, PyTorch

PROJECTS

Radiology Image De-identification using Generative AI   Aug, 2025 – Oct, 2025

- **Overview**: This project investigates the topic of removing sensitive information present within radiology images, using image segmentation and image inpainting. The project provides a proof of concept for a convenient and scalable de-identification pipeline, which can serve medical experts with no prior experience to generative AI.
- **Methodology**: The project combines Segment Anything Model with inpainting models such as LaMa and Stable Diffusion to perform the de-identification process. To investigate Stable Diffusion's performance on capturing anatomical information, I also fine-tuned the model on X-ray images provided by the MURA dataset.
- **Libraries Used:** PyTorch, Diffusers, Accelerate, bitsandbytes, Gradio
- **Link to repository**: https://github.com/Ngoc-Cac/gen_ai_deidentification

Big Data Application for Sentiment Analysis on KFC Reviews in HCMC   Apr, 2025 – May, 2025

- **Overview**: This group project investigates the development and deployment of a Big Data system to support the storage and processing of KFC reviews in Ho Chi Minh City.
- **Methodology**: In this project, we used Big Data technologies such as **Apache Hadoop** and **Apache Spark** to store and process large amount of unstructured data. We followed a basic **ETL process** where reviews are extracted from Google Maps by means of web crawling into our locally-hosted HDFS, deployed using pre-built Docker images. Afterwards, the data is preprocessed and saved for future analysis and model training with **PySpark** and **PyTorch**.
- **Responsibilities**: I am responsible for the implementation of the application infrastructure, including: the building environment and service containers for analysis, reviews crawling functionality and model training functionality.
- **Technologies Used:** Docker, Apache Hadoop, Apache Spark, Selenium, PySpark, PyTorch, Transformers
- **Link to repository**: https://github.com/Ngoc-Cac/kfc-reviews-with-bigdata

High-Performance Computing Weather Prediction System Apr, 2025 − May, 2025

- **Overview**: This group project experimented with **Docker Clusters** to deploy a high-performance computing system running a weather prediction service.
- **Methodology**: We experimented with a cluster-based computing system using Docker Clusters to deploy two main services: a backend API running our trained **LSTM model** and a frontend web UI for interactive demonstration. While the objective is to deploy a multi-node cluster, we decided to simulate this environment using VMs built with VirtualBox, due to limited time and budget.
- **Responsibilities**: My role in this project revolves around the backend functionalities, which includes the building and maintenance of necessary containers and clusters in order to deploy and run the application, the implementation and deployment of API endpoints needed in the frontend UI. I am also responsible for the training of the models used.
- **Technologies Used:** Docker, Docker Clusters, FastAPI, Chakra UI, PyTorch
- **Link to repository**: https://github.com/Ngoc-Cac/weather-prediction

RESEARCH ACTIVITIES

Proceedings of STAIS national conference 2024 May, 2024 − Aug, 2024

Citation: Cac, T. N., Khanh D. T. M., Nguyen, T. N. T., Yen, N. V., Thanh D. N. H. (2024). *Rerouting Airline Passengers of Canceled Flights with Genetic Algorithms*. STAIS 2024

- **Role**: Lead author and presenter of paper.
- **Description**: The paper investigates the use of Genetic Algorithms to reroute passengers so that most passengers can still finish their journeys. We achieved this by modelling the problem as a Maximum Flow Problem, and applied a self-implemented genetic algorithm to compute the edge weights. We found the implemented algorithm capable of solving networks with roughly 20-50 edges in reasonable time.
- **Code for paper**: https://github.com/Ngoc-Cac/MaxFlowGA

AWARDS

UEH Young Researcher Awards 2025 Dec, 2024 − Apr, 2025

- **Dec, 2024 − Jan, 2025**: Preparation and submission of work titled "*Rerouting Airline Passengers of Canceled Flights with Genetic Algorithms*".
- **Jan, 2025**: Results announcement of Rank A.

LANGUAGE SKILLS

Vietnamese: Native
English:

- TOEIC 2-Skill: 965 overall (taken in 2024)
- IELTS: 8.0 overall (taken in 2021)
- Fluent in English, can converse on a daily basis.
- Intermediate-level academic writing.