Truong Duc Tai

Ward 15, Go Vap district, Ho Chi Minh city
0327341107
taitruong256@gmail.com
github.com/taitruong256

Education

Industrial university of Ho Chi Minh city, Vietnam — September 2021 - Present
Bachelor of Engineering in **Data Science** (Year 4) — GPA: 3.48/4.0

Awards & Achievements

LOTTE Scholarship — 2024
Awarded by LOTTE Foundation for academic excellence. — Scholarship

Top 12 Vietnam – International Collegiate Programming Contest (ICPC) — 2024
Asia Hanoi Regional Contest — Certificate

Third Prize – International Collegiate Programming Contest (ICPC) — 2023
Southern Vietnam Regional Contest — Certificate

Second Prize – Vietnam National Collegiate Programming Contest — 2022
National Round – Specialized Track — Certificate

Consolation Prize – Vietnam National Collegiate Programming Contest — 2021
National Round – Non-specialized Track — Certificate

Top 5% - The scholarship for encouraging study. — 2 semesters
Awarded for outstanding academic performance in the program.

Project

Multi Large Language Model Code Solver — March 2025 – April 2025
Kaggle competition — *[Github]*

- Built an evaluation system to test AI models' programming skills on the CodeMMLU dataset (multiple-choice questions); input includes questions and choices, output provides answers (A, B, C, D).
- A four-step process: created standardized prompts; queried DeepSeek-V3, Gemini 2.0 Flash, GPT-4o, and Llama 3 70B Instruct via API with error handling; validated responses for single-character format (A, B, C, D); aggregated results using majority voting, applying confidence-weighted voting when needed.
- DeepSeek-V3 scored 75% accuracy, Gemini 2.0 Flash 72%, GPT-4o 70%, Llama 3 62%, voting ensemble 71%; answer distribution (A, B, C, D) was balanced, showing no bias.

Mammography Breast Cancer Detection — January 2025 – March 2025
Kaggle Competition — *[Github]*

- Developed a deep learning model to classify mammography images as cancerous or non-cancerous, supporting radiologists in improving diagnostic accuracy and efficiency.
- Implemented a three-step process: preprocessed DICOM images by converting to PNG and filtering for specific view and laterality; applied data augmentation (flipping, rotation, brightness adjustments, distortions); fine-tuned a pretrained EfficientNet-B0, Resnet50, HR-Net, Swin-transformer

Utilizing a probabilistic encoding model combined with a linear classifier and extreme value theory for the open set recognition problem. — March 2024 - September 2024
Young Scientists Conference 2024 — *[Certificate]* *[Kaggle]*

- Successfully achieved the goal of accurately classifying closed sets and correctly identifying open sets. Using beta-VAE for closed set training, encoding the input images into a latent space with normal distribution. For the encoding and decoding network, Wide ResNet (WRN) was used.
- For open set recognition, we applied the Extreme Value Theory (EVT). An image is identified as belonging to the open set if its probability threshold exceeds a predefined limit.
- The results showed that the model achieved a 99.7% classification accuracy for the closed set and a 94.0% accuracy for open set recognition.

Restoring Vietnamese Accents with Deep Learning April 2024 - June 2024
Final project of machine learning course *[Kaggle]*

- Built a deep learning system to recover Vietnamese accents from unaccented text by training on parallel data of unaccented and accented sentence pairs.
- Preprocessed the dataset by building vocabularies, tokenizing and mapping sentences to IDs, embedding sequences, then trained and evaluated both LSTM and Transformer models on the same data.
- Based on the learning curves and validation checkpoints, the Transformer consistently outperformed the LSTM: it achieved lower training and validation loss, and higher accuracy throughout training. Final validation accuracy was 0.9332 for Transformer versus 0.9011 for LSTM.

Leadership & Activities

Programming Lab, Industrial university of Ho Chi Minh city

- Co-leader of IUH Vung Tau 2024 codecamp of the school's Informatics Olympic team November 2024
- Seminar on graph knowledge in competitive programming for students in lab like DFS, BFS, shortest path, minimum spanning tree, joints and bridges May 2023
- Mentor of the club's first year c++ programming course October 2022

Competitive Programming

- Top 500 Meta Hacker Cup 2024 Qualification Round 2 Certificate
- Codeforces Expert (Rating 1613) Profile
- Honors PROCON Vietnam 2024 Certificate

Skills & Interests

Machine Learning/AI Frameworks: TensorFlow, Pytorch, Transformers, Keras, Scikit-learn, OpenCV,
Data Science: Pandas, NumPy, Matplotlib, Seaborn, Plotly
Version Control: Git, GitHub
Databases: MySQL, MongoDB, SparkSQL
APIs & Web Services: Django Rest Framework, FastAPI
Crawl data: Selenium, BeautifulSoup, Requests
Laboratory: Programing Lab, Data Innovation Lab
Interests: Reading, playing chess, swimming, walking