

Đinh Nhật Thành

0363956098 | dingnhathanh248@gmail.com | [linkedin.com/in/d-nhatthanh248](https://www.linkedin.com/in/d-nhatthanh248) | github.com/MathematicGuy

HỌC VẤN

CMC University

Bằng Cử Nhân Công Nghệ Thông Tin / GPA: 3.09

Đạt giải khuyến khích cuộc thi nghiên cứu khoa học cấp trường

Hà Đông, Hà Nội

09/2022 – 10/2025

Động lực: Sẽ ra sao nếu mình có thể giảm thiểu được tình trạng mất định hướng trong giới trẻ. Đó chính là động lực khi còn nhỏ, có rất nhiều học sinh tiềm năng nhưng vì không xác định rõ được động lực hay mục tiêu rõ ràng mà lãng phí quá nhiều thời gian vào giải trí mà không thực sự sống. Đối với 1 đứa từng là học sinh cá biệt, em thấy được vấn đề này nằm ở 2 nguyên nhân nằm ở việc thiếu định hướng rõ ràng và nhận thức được những môn mình học sẽ áp dụng được gì trong tương lai, các giải pháp hiện tại đều yêu cầu tài chính cao. Khi em thấy được tiềm năng của AI trong giáo dục, em biết đây là con đường mà mình sẽ chọn. Việc sử dụng AI sẽ giảm thiểu được phần lớn chi phí tư vấn, tạo môi trường học từ việc trích xuất content học tập qua internet, gợi ý phương pháp học cho mỗi lứa tuổi, thời gian ngủ tối ưu, tạo roadmap sử dụng kinh nghiệm của những người đi trước và vân vân. Đây là mục tiêu lớn nhất của em, điều này sẽ không chỉ giúp phát triển đất nước mà cả các bậc cha mẹ cảm thấy bất lực với những đứa trẻ có tiềm năng chưa được khai phá.

KINH NGHIỆM

Thực tập sinh

CMC ATI

09/2024 – 12/2024

Hà Đông, Hà Nội

- Phát triển hệ thống trích xuất thông tin thẻ CCCD từ A-Z, sử dụng YOLOv11 để định vị và VietOCR để nhận dạng ký tự.
- Tối ưu hóa quy trình xử lý ảnh (hiệu chỉnh phối cảnh, tăng cường chất lượng), giúp tăng độ chính xác nhận diện góc lên **83%** và trích xuất vùng thông tin (ROI) đạt **93%**.
- Xây dựng API backend bằng FastAPI và quản lý cơ sở dữ liệu với PostgreSQL để hoàn thiện luồng xử lý và lưu trữ dữ liệu.

Nghiên Cứu Sinh

CMC University

03/2025 – 7/2025

Hà Đông, Hà Nội

- Học cách đọc báo nghiên cứu và viết báo cáo latex, docx, slides chuẩn format nghiên cứu khoa học cấp trường
- Xây dựng bộ dữ liệu tùy chỉnh sử dụng Matplotlib để hình dung kết quả và xác định vấn đề thử nghiệm trích xuất keypoints của Yolopose11m cho 74 videos giúp mô hình nắm bắt được đặc trưng của mỗi hành động tốt hơn.
- Nghiên cứu và phát triển module nhận diện hành vi bạo lực học đường, kết hợp YoloPose11m để trích xuất khung xương và các mô hình chuỗi thời gian (LSTM/GRU) để phân loại hành vi.
- Thực hiện Feature Engineering chuyên sâu từ dữ liệu keypoints, giúp cải thiện Mean Accuracy của mô hình từ **43%** lên **78%**.
- Tối ưu các mô hình học sâu bằng TensorFlow với các kỹ thuật (Dropout, L2 Regularization) và đánh giá hiệu suất tin cậy qua K-Fold Cross-Validation, tận dụng CUDA để tăng tốc xử lý.

DỰ ÁN

Naive RAG for Question Answering & Multiple Choice Question Generation

06/2024 – 09/2024

- Mục đích: Tóm tắt tạo sinh câu hỏi trắc nghiệm dựa trên nội dung PDF.
- Xử lý và phân tích file PDF bằng 'PDFReader' và 'Pandas' để xác định tần suất token, đảm bảo phù hợp với context window của mô hình LLM.
- So sánh hiệu quả giữa hai phương pháp 'Recursive Chunking' và 'Dynamic Chunking' để tối ưu hóa quá trình trích xuất ngữ cảnh.
- Áp dụng 'Flash Attention', 'CUDA' và lượng tử hóa 'nf4' để tăng tốc mô hình LLM và so sánh các mô hình embedding trên các tập dữ liệu STS-B, QQP và MRPC.
- Tinh chỉnh đầu ra của LLM bằng 'Prompt Engineering' để đảm bảo câu hỏi trắc nghiệm được tạo ra có cấu trúc chuẩn và dễ dàng trích xuất bằng 'Regex'.

CHỨNG CHỈ

Samsung Innovation Campus AI Program <i>Samsung Innovation Campus</i>	06/2023 – 09/2024 <i>Samsung</i>
Linear Algebra for Machine Learning and Data Science <i>DeepLearning.ai</i>	07/2024 – 09/2024 <i>Coursera</i>
Calculus for Machine Learning and Data Science <i>DeepLearning.ai</i>	09/2024 – 10/2024 <i>Coursera</i>

KỸ NĂNG CHUYÊN MÔN

Languages: Java, C#, Python, SQL (Postgres), MongoDB (NoSQL), JavaScript, HTML/CSS
AI/ML Frameworks & Libraries: PyTorch, TensorFlow, Scikit-learn, LangChain, OpenCV, MediaPipe, Streamlit, Spacy
Developer Tools: Git, Docker, Flask, FastAPI, PostmanAPI
Big Data & Cloud (Optional): Hadoop, Google Cloud Platform