TRAN NGUYEN DUY BAO

15/02/2003 • Nha Be District, Ho Chi Minh City • trannguyenduybao152@gmail.com • +84 829 782 878
Linkedin • Github • Portfolio

AI Engineer with 2 years of experience building and deploying ML solutions in Computer Vision, NLP, and Big Data. Proven ability to deliver real-world applications in healthcare, legal, and media domains. Strong in end-to-end development, from data engineering to model deployment.

TECHNICAL SKILLS

- **Languages & Frameworks:** Python, JavaScript (Node.js, React.js), Flask, FastAPI, SQL, NoSQL
- **AI/ML Libraries:** TensorFlow, PyTorch, OCR, HuggingFace, Scikit-learn, YOLO Ultralytics, DeepFace, XGBoost, LangChain, vLLM, Ollama, Transformers, SpaCy, NLTK
- **Tools & Platforms:** Git, Docker, GCP, Azure, OpenAI API, RabbitMQ, Chromium, Scrapy, Selenium,
- **Data Engineering:** PySpark, Pandas, Matplotlib, Seaborn
- **Soft Skills:** Critical thinking, communication, research-driven, self-learning, scientific paper reading

WORKING EXPERIENCE

AI Engineer – Junior Saigon, Vietnam
MCV Complex – Space1 Jul 2023 – Present

- **Built LLM healthcare chatbot** using Flask and vector DB (RAG pipeline) utilizing AI Graph Agent.
- **Developed FaceID event check-in** with YOLOv8 + DeepFace for spoofing detection.
- **Multimodal approach** for video + audio summarization.
- Boosted content violation detection **by 75%** with a sequential data pipeline.
- **Created sentiment and legal classifiers** finetuning PhoBERT.
- **Built full-stack AI tools** for data crawling and classification.

PROJECTS

Medical Abnormality Detection from Chest X-rays Ton Duc Thang University, 2023
Technologies: Python, YOLOv8

- Fine-tuned YOLOv8; applied augmentation for 6-class detection.
- Preprocessed data using image scaling and data augmentation techniques, achieving ~70% accuracy with ~60% recall.

Vietnam Stock Forecasting Sep 2023 – Jan 2024
Technologies: Scikit-learn, Tensorflow, OLS, Scrapy, Selenium

- Collected financial data via Scrapy and Selenium from multiple sources.
- Cleaned and processed tabular data with Pandas; performed regression analysis using OLS.
- Visualized trends and anomalies using Matplotlib and Seaborn.
- Predict next time period price using LSTM resulted in ~3 point loss, approximately 5% of real data

EDUCATION

Ton Duc Thang University – B.Sc. in Computer Science (2025)
GPA: 8.0 / 10 | English: C1 | Top 10 – Recursion 2022

REFERENCES

Dr. Nguyen Trang Thao - Former PhD at TDTU Lab - ORCid - thao.nguyentrang@vlu.edu.vn

Mr. Huynh Hai Huynh - CTO at MCV Group - haihuynh.huynh@mcvnetworks.com