

Dinh Khai Nguyen

+84 37838454 nguyendinhkhaiqt2005@gmail.com <https://github.com/ngwkhai> Hanoi, Vietnam

OBJECTIVE

I aim to work in a position as an AI Engineer or Data Science Engineer where I can apply my technical skills and gain practical experience. Over the next 3–5 years, my objective is to become a key team member capable of taking on technical leadership or project management roles, and to make long-term contributions to the growth of the company.

EDUCATION

VNU University of Engineering and Technology	Aug 2023 - Present
The Degree of Bachelor in Artificial Intelligence	
GPA: 3.44/4	

HONORS & AWARDS

Academic Encouragement Scholarship, University of Engineering and Technology	2023-2024
Second Prize, Thai Nguyen Provincial Excellent Student Contest in Physics	2023
Third Prize, Thai Nguyen Provincial Excellent Student Contest in Informatics	2021

SKILLS

Programming	Python, C/C++, Java, SQL
Tools	Git, Docker, VSCode, Jupyter Notebook, LaTeX, Microsoft Office
Frameworks	PyTorch, Scikit-learn, TensorFlow, Keras, Hadoop, PySpark
Languages	Vietnamese: Professional, English: Professional
Soft Skills	Problem Solving, Teamwork, Public Speaking, Critical Thinking

PROJECTS

Machine Translation using Sequence to Sequence Models (https://github.com/ngwkhai/Neural-Machine-Translation)	Jun 2025 - Jul 2025
Position: Leader (Team of 2)	
Description: Built the first large-scale Vietnamese diacritic restoration dataset (~12.5M sentence pairs without → with diacritics). Implemented Seq2Seq models (RNN, GRU, LSTM, Transformer) from scratch using PyTorch. Designed and conducted comprehensive model evaluation with BLEU and ChrF++ metrics. Achieved BLEU 81.31 and ChrF++ 88.57 with Transformer, significantly outperforming RNN and GRU baselines. Contributed across data construction, model development, and performance evaluation , leading to a robust prototype system for Vietnamese text restoration. Technologies: Python, PyTorch, Jupyter Notebook, NumPy, Pandas, Matplotlib, BLEU Score, ChrF++.	
License Plate Recognition System (https://github.com/ngwkhai/License-Plate-Recognition)	Apr 2025 - May 2025
Position: Leader (Team of 5)	
Description: Developed a real-time license plate recognition system by integrating a fine-tuned YOLOv5 model with OCR for text extraction. Focused on system development : designed the pipeline, built REST APIs with FastAPI, managed data in PostgreSQL, and created a ReactJS-based frontend. Deployed the system with Docker for scalability and ease of integration. Achieved real-time processing at ~12 FPS with reliable recognition accuracy. Enhanced traffic management efficiency and enabled citizens to conveniently check their traffic violation records online. Technologies: Docker, FastAPI, OpenCV, ReactJS, PostgreSQL, YOLOv5.	
Credit Card Fraud Detection Using Machine Learning Algorithms (https://github.com/ngwkhai/Card-fraud-detection)	Oct 2024 - Dec 2024
Position: Leader (Team of 2)	
Description: Conducted fraud detection on the Kaggle dataset (284,807 transactions, only 492 fraud → 0.172%). Applied SMOTE and ADASYN to handle severe class imbalance. Implemented and compared Logistic Regression, Decision Tree, Random Forest, and XGBoost . Improved Recall to 0.816 with ADASYN while maintaining high ROC-AUC (0.9744). Contributed to data preprocessing, resampling, model development, and evaluation , delivering a robust prototype for financial fraud detection. Technologies: Python, NumPy, Pandas, Scikit-learn, imbalanced-learn, Matplotlib, Seaborn.	