

Nguyen Le

+18626223765 | nguyenld.work@gmail.com | <https://linkedin.com/in/bomnguyen00/> | <https://github.com/nguyenld00>

Summary

Aspiring Machine Learning Engineer and Data AI Engineer Intern with substantial experience in enhancing AI solutions and data pipeline efficiencies. Demonstrated success in developing a real estate chatbot using reinforcement learning and implementing optimized real-time data ingestion for increased operational efficiency. Eager to leverage a strong foundation in AI and data engineering to drive innovation in machine learning applications.

Education

De Anza College

Computer Science

Cupertino, California

Jan 2025 - May 2027

Work Experience

Rockship

Data Engineer Intern

Ho Chi Minh City, Vietnam

Aug 2024 - Present

- Developed and implemented a data pipeline using **Singer SDK** to extract and transform over **50,000** ticket data entries from Zendesk and **300,000 form submissions** from Typeform, resulting in a **30% improvement in data processing efficiency**
- Created and deployed API endpoints that automatically received real-time data from Zendesk and Typeform, handling **up to 1,000 requests per minute**, and efficiently uploading the data to **AWS SQS** and **S3**
- Implemented serverless Python scripts using **AWS Lambda** to process data extraction, achieving a **25% reduction in operational costs** and ensuring **80% uptime** for critical data processing tasks

Rockship

Machine Learning Engineer Intern

Ho Chi Minh City, Vietnam

May 2024 - Aug 2024

- Developed a real estate chatbot consultant using **CrewAI** to construct a **single-agent** and **multi-agent reinforcement learning system**, enabling users to interact with the chatbot in a natural and informative way
- Implemented Langchain for document indexing in a real estate chatbot, improving query response relevance by **25%**.
- Designed and implemented a **Qdrant vector database** to optimize **RAG** techniques, reducing AI hallucination and improving reliability of information retrieval by **15%**.

Dizim.ai

Machine Learning Engineer Intern

Ho Chi Minh City, Vietnam

Feb 2024 - Apr 2024

- Automated data preprocessing and feature engineering processes, reducing model training time by **25%** and enhancing the efficiency and scalability of our data pipelines, which are crucial for processing large volumes of data
- Developed a crawling automation script that gathered **1000 videos** from Youtube and TikTok **30%** faster than manual methods, significantly reducing data collection time for AI model training
- Improved the preprocessing pipeline system, leading to a significant enhancement in speaker detection accuracy within videos, which subsequently boosted the performance of **lip-sync training models**

LightVision, Inc

Machine Learning Engineer Intern

Ho Chi Minh City, Vietnam

Dec 2023 - Feb 2024

- Facilitated seamless integration of computer vision solutions into multiple cross-functional projects, leading to a **10%** increase in project productivity and enhancing the team's expertise in computer vision technologies
- Proposed and implemented improvements to the fire detection function, resulting in a **15%** increase in detection sensitivity and minimizing the risk of fire-related incidents for the business

Projects

LLMs Benchmark for Medicine | HuggingFace

Mar 2024 - May 2024

- Contributed to a database of **30000+ medicine and pharmacy Q&A samples** in Vietnamese, enhancing its utility for evaluating LLM in answering and decision-making medicine and pharmacy questions
- Applied advanced quantization techniques, including the groundbreaking **GPQT method**, to optimize storage capacity and inference performance of large models on limited GPU resources

"Fintech" Challenge by Timo at HackYouth Hackathon | Github

Feb 2023 - Mar 2023

- Developed an Ethereum-based philanthropy platform that facilitates donations through **Fractional NFTs**, exchangeable for real-world assets, instead of **ERC-20 tokens**
- Developed, tested, and deployed a smart contract based on **ERC-1155 Multi-Token Standard** to the Goerli Testnet using Hardhat, with a user interface built in **ReactJS**, integrated via **Ethers.js** and **MetaMask SDK**

Skills

Languages: Python, C++, HTML/CSS, JavaScript, SQL

Developer Tools: VS Code, Google Colab, Google Cloud Platform, Jupyter Notebook

Technologies/Frameworks: PyTorch, TensorFlow, HuggingFace, Pandas