Hoang Le Duc • Machine Learning Engineer • Data Engineer

Gainesville, FL 32612 • (786) 674-9470 • Leduchoang011@gmail.com • linkedin.com/in/hoangleduc • github.com/LeDHoang

Personal Summary

Ambitious Computer Science student with industry experience in Machine Learning, Data Science. Passionate about leveraging AI to develop impactful solutions, such as healthcare chatbots and fraud detection systems, and seeking opportunities to advance innovation in AI-driven technologies.

Education

B.Sc. Computer Science - University of Florida, FL   GPA 3.6 | Expected Aug 2025
Herbert Wertheim College of Engineering Dean's List | Sigma Alpha Mu Fraternity Gamma Tau Chapter
Certifications: *IBM AI Engineering Professional Certificate (Ongoing), Ironhack Data Analytics Bootcamp*

Professional Experience

Machine Learning and AI Internship   **FL, United States**
Verifone   *Aug 2024 - Now*
- Designed and implemented a centralized, multi-site fraud detection model using **Python, TensorFlow, and TCNN**, boosting detection accuracy by **20%**.
- Streamlined **ETL** processes and improved model efficiency with **hyperparameter tuning** via **Optuna**, cutting training time by **50%** and leveraging **ClearML** for performance evaluation.
- Worked with diverse teams to integrate a user-friendly **UI** into the **C-Suite** system, enhancing oversight and expediting anomaly detection across **16** sites.

Machine Learning and AI Internship   **Ha Noi, Vietnam**
Institute of Technology and Science   *May 2022- July 2022*
- Engineered a **Python**-based facial identification system for touch-free COVID-safe applications using **PyTorch** and **CNN**, achieved **87%** accuracy under diverse conditions, outperforming existing systems by **10%**.
- Curated a dataset of over **5000 facial keypoint** photos from public sources and onsite volunteers in order to improve performance at different conditions and subjects.

Personal Projects

[RAG-based] PDF Query System for Healthcare Insurance   *Dec 2024*
- Developed a **NLP-powered RAG** application enabling users to query healthcare insurance PDFs with precise and contextually relevant responses through an interactive web interface, powered by **Streamlit** and **Docker** for containerized deployment.
- Implemented Admin application for PDF ingestion, text chunking, GenAI embedding generation with **FAISS**, **Amazon Titan**, and **AWS S3**-based index storage.
- Built a User application integrating **LangChain** and **Claude** to deliver contextually accurate responses to user queries.

[DeepSeekResearch] Multi-agent Deep Research Assistant   *Jan 2025*
- Developed a multi-agent AI-powered research assistant using **CrewAI**, integrating advanced search tools and **LLM** models to gather, analyze, and synthesize comprehensive research findings.
- Built an interactive **Streamlit** interface that enables users to refine their research prompts and generate detailed, structured reports for data-driven insights.

Skills

Programming Languages: Python, SQL, R, HTML, CSS, Java, C++, JavaScript.
Frameworks & Libraries: TensorFlow, PyTorch, Scikit-learn, LangChain, React.js, Next.js, Pandas, NumPy.
Tools: Docker, Git, AWS S3, FAISS, Tableau, Power BI, ClearML.
Domains: AI, Machine Learning, Deep Learning, Computer Vision, Anomaly Detection, NLP.