Gia Huy Phung

Dave Brubeck Way • Stockton, CA 95204 • jerryphung282@gmail.com • +1 (350) 217 9492 •
linkedin.com/in/huy-phung-gia • huyhoclaptrinh.github.io

Education

University of the Pacific — Stockton, CA
Master, Data Science. 4.0 — May 2026
Relevant Coursework: Machine Learning, NLP, Math for Data Science, Customer Analytics

University of Science and Technology of Hanoi — Hanoi, Vietnam
Bachelor, Information Communication and Technology. 3.0 — Dec 2023

Experience

TreeScape Power — Remote
Intern ML Engineer — June – Sept 2025

- Developed an end-to-end Computer Vision pipeline using **TensorFlow** and **OpenCV** to detect tree health defects, achieving **75% accuracy** on unstructured image data.
- Designed **offline experiments** to compare feature engineering techniques (HSV/LAB), utilizing statistical significance checks to select the optimal model architecture.
- Optimized model for edge deployment using **TensorFlow Lite**, reducing inference latency for mobile usage.

AHT Tech JSC — Cau Giay, Hanoi
Technical Consultant — Dec 2023 - Aug 2024

- Optimized complex **SQL** queries to improve database performance, significantly reducing response latency for critical ERP systems.
- Engineered automated data pipelines using **Python** and **PostgreSQL**, streamlining business workflows and minimizing manual intervention.
- Built automated **SQL dashboards** and data models to track key business metrics, serving as the source-of-truth for cross-functional stakeholders.

Projects

GoPark: Smart Parking Detection & Recommendation System — Stockton, CA
Computer Vision Developer — Dec 2025

- **Architected an AI-powered parking system** using **FastAPI** and **PostgreSQL/PostGIS** to orchestrate real-time spatial data and concurrent user requests.
- **Developed a dual-stage CV pipeline** (YOLOv8 & ResNet-18), achieving **97.2% mAP** with **5.4ms latency** for real-time vehicle detection and size classification.
- **Built a custom Regex-based NLP parser** that converts natural language to structured JSON, outperforming LLM latency by **500x** with a **96% extraction success rate**.
- **Optimized 24/7 reliability** via a Night-Time strategy with HSV augmentation, resulting in a **5.3-fold increase** in vehicle detections under low-light conditions.
- **Engineered a multi-factor scoring engine** that filtered **98% of invalid spots** based on vehicle size, distance, and ADA/EV constraints.

GOPHORA Stockton, CA
Full Stack AI Developer Oct 2025

- **Architected** a scalable microservices AI platform using **FastAPI**, **React**, and **Docker**.
- Engineered an **LLM-powered ranking system** using **Gemini Embeddings** and Vector Search, optimizing **match quality** between user profiles and opportunities.
- Designed **evaluation rubrics** to assess RAG retrieval accuracy, iterating on prompt engineering to improve relevance scores and reduce hallucination.

Predictive Modeling for English Premier League (EPL) 2022-2023 Standings Cau Giay, Hanoi
Data Scientist April 2022

- **Developed a predictive model** for the 2022-2023 EPL season by aggregating historical match, club, and player data from the 2020-2022 seasons.
- **Performed comparative analysis** between **Linear Regression** and **Random Forest Regressor**, identifying the latter as the superior model with an **R^2 score of 0.999.**
- **Optimized feature selection** by analyzing **Pearson correlation heatmaps**, identifying key predictors such as goals for, goal difference, and home possession to improve accuracy.
- **Visualized multi-dimensional team performance** using **Radar (Spider) charts** and interactive dashboards to identify historical trends and tactical strengths/weaknesses.
- **Conducted complexity experiments** to evaluate the trade-offs between model accuracy and training/evaluation time, ensuring an efficient predictive pipeline.

Skills & Interests

Languages & Frameworks: Python, SQL (PostgreSQL/PostGIS), PyTorch, TensorFlow, Scikit-learn, FastAPI, React.
AI & Statistics: Generative AI (RAG, Gemini), Computer Vision (YOLOv8, ResNet-18), NLP (Regex-based Parsing), Statistical Modeling (Linear Regression, Random Forest), Pearson Correlation.
Tools & Platforms: Docker, Git, CI/CD, PostGIS, Data Visualization (Plotly, Matplotlib, Seaborn), BeautifulSoup (Web Scraping).