



KỸ SƯ MACHINE LEARNING

PHẠM DŨNG SƠN

MỤC TIÊU NGHỀ NGHIỆP

Tôi mong muốn tham gia vào việc triển khai mô hình Machine Learning lên môi trường production một cách an toàn, có thể theo dõi, giám sát hiệu suất mô hình và tự động cập nhật khi mô hình xuống cấp (model drift).

THÔNG TIN CÁ NHÂN

05/06/1986

Hà Nội

thaolinhh252512@gmail.com

0899882215

www.website.com

KINH NGHIỆM LÀM VIỆC

- **MACHINE LEARNING ENGINEER** TẠI CÔNG TY AI SOLUTIONS (2021-2023)

+ PHÂN TÍCH VÀ XỬ LÝ TẬP DỮ LIỆU LỚN CHO MÔ HÌNH DỰ ĐOÁN HÀNH VI NGƯỜI DÙNG

+ HUẤN LUYỆN MÔ HÌNH PHÂN LOẠI SỬ DỤNG SCIKIT-LEARN VÀ XGBOOST

+ TRIỂN KHAI MÔ HÌNH VÀO HỆ THỐNG BACKEND BẰNG FLASK VÀ DOCKER

- **AI ENGINEER** TẠI VISIONAI (2021-2022)

HỌC VẤN

- Khoa học máy tính tại Đại học FPT -
Học máy và khai phá dữ liệu tại Đại học Khoa học Tự nhiên - ĐHQG Hà Nội

KỸ NĂNG

- MLflow

- XGBoost

SỞ THÍCH

- Tham gia hackathon
- Sưu tầm sách lập trình
- Đọc sách
- Học ngoại ngữ

+ TRIỂN KHAI MÔ HÌNH NHẬN DIỆN KHUÔN MẶT BẰNG CNN VÀ PYTORCH

+ LÀM VIỆC VỚI DỮ LIỆU HÌNH ẢNH VÀ VIDEO LỚN, DÙNG AUGMENTATION ĐỂ CẢI THIẾN ĐỘ CHÍNH XÁC

+ TÍCH HỢP HỆ THỐNG VÀO ỨNG DỤNG GIÁM SÁT NỘI BỘ

NGƯỜI GIỚI THIỆU

- Ông Hồ Văn Tùng (AI Researcher Lead - AILab Việt Nam) - tung.ho@ailab.vn - 0977999000

- Bà Vũ Ngọc Mai (Quản lý sản phẩm AI - TechX) - mai.vu@techx.vn - 0967888999

- Bà Nguyễn Thị Lan (Data Science Manager - SmartRetail) - lan.nguyen@smartretail.vn - 0988111222

- Ông Phạm Minh Đức (Lead AI Engineer - FinTech Lab) - duc.pham@fintechlab.vn - 0912555666

DANH HIỆU VÀ GIẢI THƯỞNG

- **2023** - Vinh danh 'Kỹ sư AI sáng tạo nhất năm' tại SmartData Lab

CHỨNG CHỈ

- **2020** - Google Cloud Professional Machine Learning Engineer

HOẠT ĐỘNG

- **Người viết blog kỹ thuật tại Blog cá nhân - ml4vn.dev (2021 - nay)**

+ Viết bài chia sẻ về cách tối ưu mô hình Random Forest và XGBoost.

+ Tổng hợp tài liệu học sâu với PyTorch.

+ Phân tích các thuật toán học không giám sát như KMeans, DBSCAN.

- **Mentor đồ án tốt nghiệp về học máy tại CLB Sinh viên AI**

(2021)

- + Góp ý cải tiến mô hình dự đoán nhu cầu sản phẩm theo mùa vụ.
- + Hướng dẫn áp dụng pipeline sklearn để tổ chức mã rõ ràng.
- + Giúp nhóm sinh viên đạt điểm cao nhất lớp với sản phẩm AI ứng dụng.

- Tình nguyện viên huấn luyện mô hình AI mã nguồn mở tại Vietnam AI Community (2022)

- + Làm sạch và chuẩn bị dữ liệu văn bản tiếng Việt từ các nguồn báo chí.
- + Đào tạo mô hình phân loại văn bản bằng scikit-learn và TensorFlow.
- + Đóng góp mã nguồn và tài liệu trên GitHub cộng đồng.

- Thực tập sinh Machine Learning tại Công ty AI Insight (2020)

- + Hỗ trợ xử lý dữ liệu hình ảnh cho hệ thống nhận diện đối tượng giao thông.
- + Huấn luyện mô hình CNN sử dụng TensorFlow và Keras.
- + Viết script đánh giá mô hình và tối ưu hyperparameters.

- Thành viên nhóm tổ chức AI Hackathon tại CLB Công nghệ Thông tin UIT (2023)

- + Xây dựng đề bài và dữ liệu mẫu cho cuộc thi.
- + Chấm điểm các mô hình dự đoán, đánh giá accuracy và overfitting.
- + Tổ chức trao giải và hướng dẫn cải tiến mô hình cho các đội thi.

DỰ ÁN

- Tự động phân loại email hỗ trợ khách hàng (AI Engineer, HelpDeskAI) 2021

Ứng dụng NLP để phân loại email hỗ trợ khách hàng thành các nhóm: kỹ thuật, thanh toán, đăng ký, v.v.

- + Huấn luyện mô hình phân loại văn bản bằng Naive Bayes và BERT
- + Tích hợp hệ thống vào ứng dụng email nội bộ
- + Đánh giá độ chính xác và phản hồi lại các sai lệch phân loại

- Dự báo nhu cầu kho hàng thời gian thực (ML Engineer, LogiData) 2023

Dự báo số lượng hàng tồn kho theo ngày dựa trên dữ liệu lịch sử bán hàng.

- + Xây dựng mô hình ARIMA và Prophet để dự báo thời gian
- + Tự động hóa pipeline thu thập dữ liệu từ hệ thống ERP
- + Gửi cảnh báo đến quản lý khi dự đoán thấp hơn ngưỡng an toàn

- Phân tích rủi ro tín dụng bằng học máy (ML Specialist, FinPredict) 2021

Phân loại rủi ro vay vốn dựa trên dữ liệu lịch sử tài chính khách hàng.

- + Tiền xử lý dữ liệu thiếu và xử lý dữ liệu mất cân bằng bằng SMOTE
- + Huấn luyện mô hình với XGBoost và Logistic Regression
- + Xây dựng báo cáo trực quan bằng Streamlit phục vụ phòng tín dụng

