



KỸ SƯ MACHINE LEARNING

ĐỖ KHÁNH LAN

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

21/11/2000

Hà Nội

thaolinh252512@gmail.com

0723214598

www.website.com

HỌC VẤN

- Phân tích dữ liệu lớn tại Học viện Công nghệ Bưu chính Viễn thông - Khoa học dữ liệu và Trí tuệ nhân tạo tại Đại học Phenikaa

KỸ NĂNG

- MLflow

SỞ THÍCH

KINH NGHIỆM LÀM VIỆC

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

+ 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 THIỆN ĐỘ CHÍNH XÁC

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

- **JUNIOR MACHINE LEARNING ENGINEER** TẠI DATAWAVE (2022-2023)

- Nghe nhạc

- Tham gia hackathon

- Tập gym

+ XÂY DỰNG PIPELINE XỬ LÝ DỮ LIỆU VÀ HUẤN LUYỆN MÔ HÌNH VỚI MLFLOW

+ TỐI ƯU HIỆU SUẤT INFERENCE VÀ TRIỂN KHAI LÊN MÔI TRƯỜNG STAGING BẰNG FASTAPI

NGƯỜI GIỚI THIỆU

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

- Bà Nguyễn Minh Châu (Technical Project Manager - MLTech) - chau.nguyen@mltech.vn - 0933777888

+ VIẾT SCRIPT KIỂM THỬ MÔ HÌNH VÀ ĐÁNH GIÁ ĐỘ ỔN ĐỊNH THEO THỜI GIAN

- **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

- **MACHINE LEARNING SPECIALIST** TẠI FINTECH AI LAB (2020-2022)

+ PHÁT TRIỂN HỆ THỐNG PHÁT HIỆN GIAN LẬN GIAO DỊCH TÀI CHÍNH THEO THỜI GIAN THỰC

+ XÂY DỰNG MÔ HÌNH PHÂN CỤM KHÁCH HÀNG SỬ

DỤNG KMEANS VÀ DBSCAN

+ THEO DÕI HIỆU SUẤT MÔ HÌNH BẰNG DASHBOARD
CUSTOM SỬ DỤNG STREAMLIT

DANH HIỆU VÀ GIẢI THƯỞNG

- **2020** - Giải thưởng Đóng góp nổi bật cho hệ thống gợi ý sản phẩm nội bộ
- **2021** - Bằng khen vì hoàn thành dự án AI thương mại trước thời hạn 1 tháng

CHỨNG CHỈ

- **2022** - IBM Professional Certificate in Machine Learning – edX
- **2020** - Machine Learning Specialization – Coursera (Stanford University)
- **2023** - Introduction to MLOps – Coursera (DeepLearning.AI)
- **2021** - AWS Certified Machine Learning – Specialty
- **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.

- Thành viên nhóm nghiên cứu Machine Learning tại Phòng Lab AI, Đại học Bách Khoa TP.HCM (2021 - 2023)

- + Phát triển các mô hình dự đoán giá cổ phiếu dựa trên dữ liệu chuỗi thời gian.

- + Huấn luyện mô hình học sâu sử dụng LSTM với dữ liệu tài chính thực tế.

- + Viết báo cáo khoa học và trình bày tại hội thảo sinh viên nghiên cứu khoa học.

- 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.

- Tình nguyện viên trong dự án AI cộng đồng tại Dự án AI for Good – UNDP Vietnam (2020)

- + Tham gia xây dựng mô hình nhận diện nụ cười cho trẻ em khiếm thính.

- + Hỗ trợ annotation và xử lý ảnh bằng OpenCV.

- + Viết báo cáo trình bày tại buổi tổng kết dự án.

DỰ ÁN

- Hệ thống nhận diện cảm xúc trong video họp (ML Engineer, EmotionAI) 2023

Phân tích hình ảnh khuôn mặt trong video để xác định mức độ hài lòng của người tham dự họp.

- + Xử lý video, tách frame và detect khuôn mặt bằng OpenCV
- + Dùng CNN để phân loại biểu cảm khuôn mặt
- + Kết hợp thời gian thực bằng WebSocket để phản hồi live

- 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

- 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

