



KỸ SƯ MACHINE LEARNING

VŨ MINH THỊ

MỤC TIÊU NGHỀ NGHIỆP

Tôi muốn thành thạo việc xử lý và phân tích dữ liệu lớn (big data) trên nền tảng cloud như AWS Sagemaker, Google Cloud AI Platform để huấn luyện mô hình hiệu quả với chi phí tối ưu.

THÔNG TIN CÁ NHÂN

15/01/1991

Hà Nội

thaolinh252512@gmail.com

0380437034

www.website.com

HỌC VẤN

- Học máy và khai phá dữ liệu tại Đại học Khoa học Tự nhiên – ĐHQG Hà Nội
- Công nghệ phần mềm tại Đại học Sư phạm Kỹ thuật TP.HCM

KỸ NĂNG

- Feature Engineering
- Scikit-learn
- Model Deployment (Flask, FastAPI)
- PyTorch
- TensorFlow

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

- **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

SỞ THÍCH

- Chụp ảnh
- Chơi đàn guitar

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

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

NGƯỜI GIỚI THIỆU

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

- Ông Đỗ Trung Kiên (Trưởng nhóm Kỹ thuật AI - DeepVision) - kien.do@deepvision.vn - 0944666777

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

- Bà Trần Thu Hằng (Trưởng phòng Phân tích Dữ liệu - DataWave) - hang.tran@datawave.vn - 0933222444

- Bà Lê Thùy Trang (Chuyên gia AI - InsightPro) - trang.le@insightpro.vn - 0955333444

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

+ 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

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

DANH HIỆU VÀ GIẢI THƯỞNG

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

- **2020** - Nhân viên tiên phong triển khai MLOps nội bộ thành công

- **2021** - Kỹ sư Machine Learning xuất sắc quý II tại Công ty AI Solutions

- **2020** - Giải thưởng Đóng góp nổi bật cho hệ thống gợi ý sản phẩm nội bộ
- **2021** - Được Giám đốc Kỹ thuật ghi nhận vì cải tiến hệ thống tự động hóa mô hình

CHỨNG CHỈ

- **2022** - TensorFlow Developer Certificate – DeepLearning.AI
- **2020** - Google Cloud Professional Machine Learning Engineer
- **2021** - Advanced Machine Learning on Kaggle – Coursera

HOẠT ĐỘNG

- 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.

- Diễn giả hội thảo sinh viên nghiên cứu ML tại Hội thảo AI4Students (2022)

- + Trình bày đề tài phân tích hành vi khách hàng bằng clustering.
- + Hướng dẫn sinh viên tiếp cận pipeline xử lý dữ liệu lớn với Spark MLlib.
- + Giao lưu và giải đáp các thắc mắc về sự khác biệt giữa ML và DL.

- 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.

DỰ ÁN

- Hệ thống gợi ý sản phẩm cá nhân hóa (Machine Learning Engineer, ShopAI) 2022

Xây dựng hệ thống recommendation engine dựa trên hành vi duyệt web và lịch sử mua hàng của người dùng.

- + Tiền xử lý dữ liệu hành vi và xây dựng profile người dùng
- + Huấn luyện mô hình Collaborative Filtering và Matrix Factorization
- + Triển khai mô hình trên server bằng Flask và Docker

- 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

- 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

