



## KỸ SƯ MACHINE LEARNING

VŨ NGÂN THU

### MỤC TIÊU NGHỀ NGHIỆP

Tôi muốn xây dựng các mô hình dự báo chuỗi thời gian (time-series forecasting) để phục vụ các bài toán như dự báo doanh thu, tồn kho, nhu cầu khách hàng,... ứng dụng thực tế trong doanh nghiệp.

### THÔNG TIN CÁ NHÂN

26/11/1986

Hà Nội

thaolinht252512@gmail.com

0337581159

www.website.com

### HỌC VẤN

- Trí tuệ nhân tạo tại Đại học Bách khoa Hà Nội - Công nghệ phần mềm tại Đại học Sư phạm Kỹ thuật TP.HCM

### KỸ NĂNG

- Google Colab

- XGBoost

### SỞ THÍCH

### KINH NGHIỆM LÀM VIỆC

- **ML ENGINEER INTERN** TẠI SMARTRETAIL (2020-2021)

+ HỖ TRỢ XÂY DỰNG HỆ THỐNG GỢI Ý SẢN PHẨM DỰA TRÊN HÀNH VI KHÁCH HÀNG

+ THỰC HIỆN EXPLORATORY DATA ANALYSIS VÀ TIỀN XỬ LÝ DỮ LIỆU

+ SỬ DỤNG LIGHTGBM VÀ GRID SEARCH ĐỂ CẢI THIẾN ĐỘ CHÍNH XÁC MÔ HÌNH

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

- Trồng cây

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

## NGƯỜI GIỚI THIỆU

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

+ XÂY DỰNG MÔ HÌNH PHÂN CỤM KHÁCH HÀNG SỬ DỤNG KMEANS VÀ DBSCAN

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

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

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

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

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

+ 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

- **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)

+ 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

## DANH HIỆU VÀ GIẢI THƯỞNG

- **2023** - Danh hiệu 'Mô hình AI tối ưu hiệu suất nhất năm' – tại FinTech AI

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

## CHỨNG CHỈ

- **2023** - Introduction to MLOps – Coursera (DeepLearning.AI)

- **2022** - Natural Language Processing with Classification and Vector Spaces – DeepLearning.AI

- **2020** - Machine Learning Specialization – Coursera (Stanford University)

## 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.

- **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.

- **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 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**

#### **- 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

#### **- 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

#### **- 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.

- + Tiễn 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

#### **- 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

#### **- 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

