



## KỸ SƯ MACHINE LEARNING

VŨ PHƯƠNG THU

### MỤC TIÊU NGHỀ NGHIỆP

Tôi hướng đến vai trò thiết kế pipeline xử lý dữ liệu và huấn luyện mô hình tự động hóa bằng cách áp dụng các kỹ thuật MLOps hiện đại như MLflow, TFX, Airflow,... giúp giảm thời gian triển khai và tăng độ tin cậy.

### THÔNG TIN CÁ NHÂN

11/11/1990

Hà Nội

thaolinh252512@gmail.com

0741562757

www.website.com

### HỌC VẤN

- Phân tích dữ liệu lớn tại Học viện Công nghệ Bưu chính Viễn thông

### KỸ NĂNG

- Model Deployment (Flask, FastAPI)

- Python

### 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

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

- Xem phim khoa học viễn tưởng
- Học ngoại ngữ
- Chơi đàn guitar

+ 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

## NGƯỜI GIỚI THIỆU

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

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

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

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

## DANH HIỆU VÀ GIẢI THƯỞNG

- **2021** - Được Giám đốc Kỹ thuật ghi nhận vì cải tiến hệ thống tự động hóa mô hình

## CHỨNG CHỈ

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

## HOẠT ĐỘNG

- **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.

- **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.

## **- Cộng tác viên nghiên cứu với doanh nghiệp tại Dự án hợp tác cùng Công ty RetailAI (2021 - 2022)**

- + Xây dựng hệ thống dự đoán doanh thu theo khu vực và mùa vụ.
- + Kết hợp dữ liệu từ thời tiết và lịch sử bán hàng để tạo đặc trưng mới.
- + Triển khai mô hình Gradient Boosting và trình bày kết quả cho khách hàng.

## **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

#### **- 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 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

