



## KỸ SƯ MACHINE LEARNING

HOÀNG DŨNG SƠN

### MỤC TIÊU NGHỀ NGHIỆP

Tôi đặt mục tiêu phát triển hệ thống pipeline training + serving hoàn chỉnh trên cloud, kết hợp với Kubernetes và Docker để đảm bảo khả năng mở rộng (scalability) và tính ổn định cao.

### THÔNG TIN CÁ NHÂN

16/03/1982

Hà Nội

thaolinhh252512@gmail.com

0917539762

www.website.com

### HỌC VẤN

- Khoa học máy tính tại Đại học FPT

### KỸ NĂNG

- SQL

### SỞ THÍCH

- Nấu ăn

- Tham gia hackathon

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

## NGƯỜI GIỚI THIỆU

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

+ 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

- **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

- **2021** - Được Giám đốc Kỹ thuật ghi nhận vì cải tiến hệ thống tự động hóa mô hình
- **2020** - Nhân viên tiên phong triển khai MLOps nội bộ thành công
- **2022** - Top 5 nhân sự có mô hình dự đoán chính xác nhất trong đội ML
- **2022** - Top 3 kỹ sư xây dựng hệ thống phát hiện gian lận tốt nhất năm

## CHỨNG CHỈ

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

## HOẠT ĐỘNG

- **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.

**- 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.

## DỰ ÁN

**- 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

