



KỸ SƯ MACHINE LEARNING

NGUYỄN THẢO NGÀ

MỤC TIÊU NGHỀ NGHIỆP

Tôi mong muốn góp phần cải tiến trải nghiệm người dùng bằng cách áp dụng các mô hình recommendation system sử dụng Collaborative Filtering, Matrix Factorization hoặc Deep Learning.

THÔNG TIN CÁ NHÂN

28/03/1981

Hà Nội

thaolinh252512@gmail.com

0756045595

www.website.com

HỌC VẤN

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

KỸ NĂNG

- Python

- SQL

- Airflow

KINH NGHIỆM LÀM VIỆC

- **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)

SỞ THÍCH

- Chơi cờ vua
- Chơi đàn guitar
- Trồng cây
- Sưu tầm sách lập trình
- Chụp ảnh

NGƯỜI GIỚI THIỆU

- Ông Nguyễn Văn Thịnh (Head of Machine Learning – AI Solutions) - thinh.nguyen@aisolutions.vn - 0908777888
- Bà Nguyễn Minh Châu (Technical Project Manager – MLTech) - chau.nguyen@mltech.vn - 0933777888
- Bà Lê Thùy Trang (Chuyên gia AI – InsightPro) - trang.le@insightpro.vn - 0955333444
- 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

+ 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

- **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

- **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

- **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

- **2022** - Top 3 kỹ sư xây dựng hệ thống phát hiện gian lận tốt nhất năm

- **2020** - Giải thưởng Đóng góp nổi bật cho hệ thống gợi ý sản phẩm nội bộ

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

CHỨNG CHỈ

- **2021** - Advanced Machine Learning on Kaggle – Coursera

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

- **2021** - Deep Learning Specialization – Coursera (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.

- **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 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.

- **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

- 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

- 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.

- + Tiễn 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

