



KỸ SƯ MACHINE LEARNING

BÙI LAN NGÀ

MỤC TIÊU NGHỀ NGHIỆP

Tôi đặt mục tiêu làm chủ các kỹ thuật tối ưu hóa như Bayesian Optimization, Grid Search, Random Search để tìm cấu hình mô hình tối ưu nhất trong thời gian huấn luyện hạn chế.

THÔNG TIN CÁ NHÂN

12/09/1992

Hà Nội

thaolinh252512@gmail.com

0713241597

www.website.com

HỌC VẤN

- Khoa học dữ liệu và Trí tuệ nhân tạo
tại Đại học Phenikaa

KỸ NĂNG

- MLOps

- SQL

SỞ THÍCH

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

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

- Tham gia hội thảo công nghệ
- Tham gia cộng đồng lập trình
- Sưu tầm sách lập trình
- Chơi đàn guitar

NGƯỜI GIỚI THIỆU

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

+ 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

- **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

- **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

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

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

CHỨNG CHỈ

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

- **2023** - Applied Data Science with Python – University of Michigan (Coursera)
- **2020** - Machine Learning Specialization – Coursera (Stanford University)
- **2021** - AWS Certified Machine Learning – Specialty
- **2020** - Google Cloud Professional Machine Learning Engineer

HOẠT ĐỘNG

- 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

- 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

- 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

- 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