



## KỸ SƯ TRÍ TUỆ NHÂN TẠO

BÙI SƠN SƠN

### MỤC TIÊU NGHỀ NGHIỆP

Tôi mong muốn học sâu về Explainable AI (XAI), giúp mô hình trở nên minh bạch và có thể giải thích được trong môi trường tài chính, y tế và pháp lý.

### THÔNG TIN CÁ NHÂN

23/01/1988

Hà Nội

thaolinh252512@gmail.com

0335780101

www.website.com

### KINH NGHIỆM LÀM VIỆC

- **AI RESEARCH ASSISTANT** TẠI VIỆN TRÍ TUỆ NHÂN TẠO VIỆT NAM (2021-2022)

+ THAM GIA NGHIÊN CỨU ỨNG DỤNG TRANSFORMER TRONG XỬ LÝ NGÔN NGỮ TỰ NHIÊN TIẾNG VIỆT

+ HUẤN LUYỆN MÔ HÌNH BERT TÍNH CHỈNH CHO PHÂN LOẠI VĂN BẢN VÀ CHATBOT

### HỌC VẤN

- Khoa học và Kỹ thuật dữ liệu tại Đại học VinUni - Khoa học máy tính tại Đại học Tôn Đức Thắng

### KỸ NĂNG

- Xử lý ngôn ngữ tiếng Việt

- OpenCV

- Machine Learning

- TensorFlow

- PyTorch

+ ĐỒNG TÁC GIẢ 2 BÀI BÁO TẠI HỘI NGHỊ AI QUỐC GIA

### DANH HIỆU VÀ GIẢI THƯỞNG

- **2022** - Bằng khen đóng góp cho dự án nhận diện khuôn mặt thời gian thực

- **2023** - Nhận giải 'Mô hình AI hiệu quả nhất' tại Hackathon ngành tài chính

## SỞ THÍCH

- Sưu tầm sách lập trình
- Trồng cây
- Tham gia cộng đồng lập trình

## NGƯỜI GIỚI THIỆU

- Bà Phan Thanh Thảo (AI Technical Lead - Công ty OpenAI Việt Nam) - thao.phan@openai.vn - 0988111223
- Ông Trịnh Quốc Hưng (Senior AI Engineer - Công ty EcomAI) - hung.trinh@ecomai.vn - 0909777666
- Bà Nguyễn Hồng Ánh (AI Project Manager - Công ty DataInsight) - anh.nguyen@datainsight.vn - 0909222333
- Ông Phạm Văn Tùng (Lead Data Scientist - Công ty FinAI) - tung.pham@finai.vn - 0933777888
- Ông Đỗ Minh Dũng (Giám sát kỹ thuật AI - Công ty AI4Life) - dung.do@ai4life.vn - 0977999000

## CHỨNG CHỈ

- **2022** - MLOps Fundamentals - DataCamp

## HOẠT ĐỘNG

### - Trợ giảng khóa học Machine Learning cơ bản tại Trung tâm Tin học UIT (2020)

- + Hướng dẫn học viên làm quen với scikit-learn, pandas và numpy.
- + Giải đáp các thắc mắc về quá trình huấn luyện mô hình ML cơ bản.
- + Chuẩn bị dữ liệu và hệ thống máy chủ cho lớp học thực hành.

### - Mentor nhóm sinh viên thi AI Hackathon tại Hội Sinh viên Kỹ thuật (2023)

- + Hướng dẫn chọn đề tài liên quan đến phân loại hình ảnh thực phẩm.
- + Góp ý cải thiện hiệu năng mô hình CNN trên tập dữ liệu tự thu thập.
- + Hỗ trợ đội thi giành giải nhì toàn quốc trong cuộc thi.

### - Thành viên nhóm tổ chức học thuật về AI tại Câu lạc bộ Sinh viên AI (2019 - 2021)

- + Tổ chức hội thảo về thị giác máy tính, xử lý ngôn ngữ tự nhiên.
- + Mời chuyên gia từ các công ty AI đến chia sẻ kinh nghiệm.

- + Phụ trách truyền thông và viết báo cáo sau hội thảo.

**- Thành viên nhóm Nghiên cứu AI ứng dụng tại Phòng thí nghiệm AI, ĐH Bách Khoa (2021 - 2023)**

- + Nghiên cứu mô hình học sâu (deep learning) áp dụng vào nhận dạng hình ảnh y tế.

- + Tham gia huấn luyện mô hình CNN trên dữ liệu ảnh chụp X-quang và MRI.

- + Công bố 1 bài báo khoa học tại hội thảo quốc gia về AI năm 2022.

## DỰ ÁN

**- Ứng dụng nhận diện khuôn mặt chấm công (Computer Vision Engineer, SmartHR) 2021**

Phát triển hệ thống nhận diện khuôn mặt thời gian thực cho doanh nghiệp vừa và nhỏ.

- + Xây dựng mô hình CNN với TensorFlow để nhận diện khuôn mặt

- + Triển khai nhận diện trên camera realtime bằng OpenCV

- + Kết nối mô hình với hệ thống chấm công và ghi log ra cơ sở dữ liệu

**- Hệ thống nhận diện biển số xe (AI Engineer, VisionAI) 2020**

Xây dựng mô hình OCR để nhận diện biển số xe từ ảnh giám sát giao thông.

- + Xử lý ảnh và trích xuất vùng chứa biển số bằng YOLO

- + Nhận diện ký tự bằng mô hình CNN kết hợp CRNN

- + Triển khai mô hình trên cloud để xử lý theo batch mỗi ngày

**- Dự án chatbot hỗ trợ khách hàng tiếng Việt (NLP)**

### **Engineer, CustomerAI) 2023**

Xây dựng chatbot thông minh xử lý ngôn ngữ tự nhiên để hỗ trợ khách hàng 24/7.

- + Huấn luyện mô hình seq2seq và fine-tune BERT tiếng Việt
- + Tích hợp RASA framework để điều hướng hội thoại
- + Triển khai API chatbot lên hệ thống website của khách hàng

### **- Hệ thống phân loại văn bản tự động cho doanh nghiệp (AI Engineer, BizAI Solutions) 2022**

Xây dựng hệ thống phân loại email và tài liệu nội bộ bằng mô hình học máy.

- + Tiền xử lý dữ liệu văn bản bằng Python và NLTK
- + Huấn luyện mô hình SVM và BERT để phân loại chủ đề tài liệu
- + Triển khai mô hình trên FastAPI và tích hợp vào dashboard nội bộ

