



KỸ SƯ TRÍ TUỆ NHÂN TẠO

ĐỖ SƠN THỊ

MỤC TIÊU NGHỀ NGHIỆP

Tôi kỳ vọng xây dựng mô hình NLP tiếng Việt có độ chính xác cao, áp dụng trong chatbot, phân loại nội dung, tổng hợp văn bản và các hệ thống hỗ trợ khách hàng.

THÔNG TIN CÁ NHÂN

26/02/1989

Hà Nội

thaolinh252512@gmail.com

0885097479

www.website.com

KINH NGHIỆM LÀM VIỆC

- **MACHINE LEARNING ENGINEER** TẠI CÔNG TY FINAL (2020-2022)

+ XÂY DỰNG HỆ THỐNG DỰ ĐOÁN RỦI RO TÍN DỤNG BẰNG RANDOM FOREST VÀ XGBOOST

+ TIỀN XỬ LÝ DỮ LIỆU TÀI CHÍNH LỚN TỪ NHIỀU NGUỒN KHÁC NHAU

+ TRIỂN KHAI MÔ HÌNH TRÊN CLOUD SỬ DỤNG FASTAPI VÀ DOCKER

HỌC VẤN

- Toán - Tin ứng dụng tại Đại học Khoa học Tự nhiên - ĐHQG TP.HCM

KỸ NĂNG

- Docker cho AI

- Scikit-learn

SỞ THÍCH

DANH HIỆU VÀ GIẢI THƯỞNG

- **2022** - Giải nhì cuộc thi mô hình AI dự đoán rủi ro tín dụng tổ

- Xem phim khoa học viễn tưởng

- Nấu ăn

chức bởi ngân hàng lớn

CHỨNG CHỈ

NGƯỜI GIỚI THIỆU

- Ông Đỗ Minh Dũng (Giám sát kỹ thuật AI - Công ty AI4Life) - dung.do@ai4life.vn
- 0977999000

- Bà Lê Thị Kim Yến (Data Science Manager - Công ty BigData Lab) - yen.le@bigdatalab.vn - 0933666999

- Bà Vũ Mai Lan (Trưởng nhóm NLP - Công ty VietAI) - lan.vu@vietai.vn - 0966888999

- **2023** - AWS Certified Machine Learning - Specialty

- **2021** - TensorFlow Developer Certificate - Google

- **2021** - Applied Data Science with Python - University of Michigan (Coursera)

HOẠT ĐỘNG

- **Thành viên dự án AI phát hiện gian lận tài chính tại Công ty FinAI (2021 - 2022)**

+ Xử lý dữ liệu giao dịch tài chính để phát hiện hành vi bất thường.

+ Huấn luyện mô hình Random Forest và Gradient Boosting với dữ liệu nhãn lịch sử.

+ Triển khai hệ thống real-time prediction phục vụ kiểm duyệt giao dịch.

- **Cộng tác viên dự án AI cộng đồng tại Vietnam OpenAI Community (2022)**

+ Tham gia chú thích dữ liệu âm thanh tiếng Việt cho mô hình chuyển giọng nói thành văn bản.

+ Viết tài liệu mô tả pipeline xử lý dữ liệu đầu vào trước khi huấn luyện.

+ Thử nghiệm các mô hình pre-trained của HuggingFace và tổng hợp kết quả benchmark.

- **Người chia sẻ nội dung trên blog AI Việt tại aivietnam.tech (2021 - nay)**

+ Viết bài phân tích kiến trúc Transformer và các ứng dụng

NLP.

- + Tổng hợp tài liệu học thuật và hướng dẫn đọc paper AI.
- + Giới thiệu cách sử dụng PyTorch và TensorFlow trong dự án thực tế.

- 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.

DỰ ÁN

- Hệ thống dự đoán giá bất động sản (Machine Learning Engineer, PropAI) 2021

Dự án ứng dụng machine learning vào việc định giá bất động sản theo vị trí và đặc điểm kỹ thuật.

- + Thu thập và xử lý dữ liệu nhà đất từ nhiều nguồn
- + Huấn luyện mô hình hồi quy (Linear, Random Forest)
- + Đánh giá độ chính xác và triển khai giao diện thử nghiệm

- 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

