HA MINH QUAN

Ho Chi Minh City, 12 District
0376-316-144 | haminhquan12c7@gmail.com | www.linkedin.com/in/ha-minh-quan-b10717294 | https://github.com/CatKingio

SUMMARY

I graduated with a degree in artificial intelligence engineering, specializing in deploying and developing AI models, with a particular emphasis on natural language processing. I have solid knowledge in machine learning, deep learning, along with proficiency in frameworks such as TensorFlow, Keras, Scikit-Learn, especially PyTorch, extensive experience working with API. With experience evenly spread across projects, including computer vision and natural language processing, and specializing in LLM chatbots. I am eager to learn and contribute value to your company in this field.

SKILLS & PROFICIENCIES

Machine Learning / DeepLearning Skill:

- Knowledgeable in: Prediction, Clustering, Classification model. Text / Image Processing, Computer Vision, Expertise in Natural Language Processing, Supervised / Unsupervised Learning Training, Fine-tuning, Evaluation
- Model Architectures: CNN, RNN, LSTM, BERT, ResNet, Yolo, Transformers.
- Library: TensorFlow, Keras, Scikit-Learn, PyTorch, PaddleOCR, OpenCV.

Data Skill:

- Knowledgeable in: Data collecting, Data cleaning, Data Preprocessing, Transformation, Analytics, Visualization.
- Library: Scapy, BeautifulSoup, SQL, Numpy, Pandas, Matplotlib, Seaborn.

Other:

- Python, Github, Linux, Docker, FastAPIs, Cloud Platform (Basic), Improve Chat GPT Prompts.
- Html, Css, Sass, Js, Flask, ASP.NET, Django.
- Teamwork, Problem Solving, Good Communication Skills, Handle high work pressure.

EXPERIENCE

TMA Intern AI Engineering **11/2023 - 1/2024**

Company: *TMA Solutions District 12, Ho Chi Minh City*

- Application for extracting data from prescriptions through captured images.

Project description: Extracting prescription information from captured images using computer vision (Library PaddleOCR V4) and arranging the information into the correct format using LLM (API GPT-3.5-turbo-1106).
Tasks performed:

- Implemented and applied fuzzy logic algorithms to refine and standardize the results synthesized by LLM.
- Optimized prompts to achieve higher standards, reduced token requests, increased coverage of use cases, and implemented parallel multi-threaded API requests. Optimized system design, testing and statistical results.

Achievements: Accuracy increased from 58% to 76.3%, The speed increases by 9% depending on traffic. The processing speed for handling single or multiple types of medications is nearly equivalent.

- Extracting data from unstructured text documents, invoices.

Project description: Extracting data from text documents and invoices via images using dynamic templates based on keyvalue relationships and origin coordinates..
Tasks performed:

- Developed the application interface using Streamlit, featuring drag-and-drop functionality to extract keyvalue on each document and invoice, along with dynamic template creation, saving of created dynamic templates, and application of templates to new documents.
- Extracted information about the content coordinates within the identified keyvalue.
- Conducted research, proposed ideas, and contributed to the development of a vector matching algorithm for keyvalue in the product.

- Other participated projects: Predicting fruit ripeness based on sound, internal organizational chatbot.

PERSONAL PROJECTS

Virtual Assistant Chatbot (LLM Mistral 7B, AutoGen, RAG, Transformers Classification, LM Studio, Docker)

Member Team: 1 *8/2023 -11/2023*

- Project description: A chatbot utilizing LLM model capable of classifying various types of requests, integrated with multi-agent architecture and retrieval-augmented generation to support automated code debugging optimization, user assistance, etc.
- Working methods and procedures:
  - The Chatbot is based on the LLM Mistral 7B model through LM Studio API.
  - Agents are created using the AutoGen library to assign roles to functions, enabling the LLM model to interact and collaborate until completion of assigned tasks.
  - RAG enables the chatbot to access external data beyond the trained model's data, including PDFs, DOCX files, and web-scraped data.
  - The transformers model classifies task command sentences, helping the chatbot categorize tasks that need to be processed before being assigned to group chat agents or executing other tasks.
  - If the assigned task involves coding, the code will be executed and debugged within a Docker environment.

Time Series - Predicting Specific Stock Prices (ARIMA, LSTM, FLASK)

Member Team: 1 *10/2023 -11/2023*

- Project description: Using VNQuant to fetch Vietnamese stock market data for training, visualizing the data, and conducting analysis, building a predictive model using ARIMA and LSTM for stock price forecasting, and creating a prediction interface.
- Working methods and procedures:
  - Utilized VNQuant to retrieve stock prices, performed data cleaning, filled missing values, removed outliers, visualized and gained insights from the data.
  - Employed LSTM and ARIMA models for stock price prediction, compared their performance and results.
  - Building the system and user interface application using Flask.

Predicting diabetic retinopathy from images (EfficientNet, Cv2, Django)

Member Team: 5 – Role: Team Manager, Model development *6/2022 -8/2022*

- Project description: Building a healthcare website with the functionality of predicting diabetic retinopathy in patients through retinal fundus images and classifying the disease level using the EfficientNet model.
- Working methods and procedures:
  - Utilizing OpenCV to read images, enhance brightness, resize, crop circular images according to retinal fundus, convert to grayscale, and highlight blood vessel features in the retinal fundus.
  - Developing the EfficientNet model, designing a bucket training loop to reduce memory load, evaluating and enhancing model performance.
  - Designing the web interface and backend to embed the trained model for image prediction.

EDUCATION

Nguyen Tat Thanh University **2020 - 2024**

Artificial Intelligence Engineer

ACHIEVEMENTS

Encouragement award at NTTU AI Hackathon Using model to predict COVID-19. **8/2022**

Role: Team leader

Fourth prize at NTTU Software Hackathon Using model to Predicting diabetic retinopathy from images. **4/2023**

Role: Team Manager, Model development.