Bhuribhat Ratanasanguanvongs

1024 Pattanakarn 24 Road, Suanluang, Suanluang, Bangkok 10250
+66 95-753-0883
bhuribhat@gmail.com
github.com/Bhuribhat
linkedin.com/in/bhuribhat-ratanasanguanvongs

Objective

I am a graduated computer engineering student at Chulalongkorn University. I am interested in Data Scientist, Computer Vision, and Machine Learning. I am actively engaged in expanding my expertise in Computer Vision and Natural Language Processing (NLP). Currently immersed in learning advanced techniques such as LLM fine-tuning and RAG (Retrieval-Augmented Generation). I am committed to continuous learning and growth, seeking opportunities to apply my engineering background to the exciting and evolving domains of Data Science and Machine Learning.

Education

Bachelor of Engineering in Computer Engineering 2020 – 2024

Chulalongkorn University – GPAX 3.31

Applied Science Program, Triam Udom Suksa School – GPAX 3.59

Related courses: Software Engineering, Software Define System, Database Systems, Statistics, Data Science and Data Engineering, Neural Network, Pattern Recognition, Computer Vision, Digital Image Processing

Certificates: google data analytics, neural networks and deep learning, convolutional neural network

Publication: "End-to-End Development of Chatbots Using Large Language Models: A Use Case in Thai Legal Documents", published in ACL 2024 Student Research Workshop

Work Experience

Machine Learning Engineer Internship, bit.studio 2023

- Developed a recognition model for American Sign Language fingerspelling utilizing OpenCV, MediaPipe, and Transformer technology, Kaggle competition hosted by Google I/O 2023
- Developed an advanced video-to-video with text prompt control animation pipeline. Utilized generative models like stable diffusion in combination with the YOLO multi-pose estimation model
- Finetuned Stable Diffusion model using Dreambooth, Textual Inversion, and LoRA
- Utilized Neural style transfer to compose one image in the style of another image

PET-CT brain scan, 2023

- PyTorch implementation of the Siamese U-Net for image semantic segmentation
- Utilized DeeplabV3_Resnet50 model to detect the source of epilepsy cases from a brain scan
- Identified specific regions of interest within the images on Chula Brain PET/CT scan dataset
- Trained the model using Dice loss as the loss function and evaluated its performance with IoU
- Worked collaboratively within a diverse team consisting of students, professors, and a doctor

Capstone Project

Reliable Domain-Specific Chatbot for Thai Language, 2023 – 2024

- Developed a Thai law chatbot to answer questions accurately and precisely within a specific and latest Thai legal documentation without having hallucination in its responses
- Read research papers related to mitigating hallucinations in the ACL Anthology
- Developed Large Language Models (LLMs) and Retrieval-Augmented Generation (RAG) using LangChain with open-sourced LLMs such as Llama2-7b, SeaLLM-7b, Typhoon-7b, OpenThaiGPT
- Developed a web scrapper for Thai legal documentation then generated a dataset using GPT-4 API
- Finetuned using QLoRA to increase LLM expertise in Thai legal domain to enhance its capabilities
- Developed a prompt optimization to increase robustness and ensure the LLM follows instructions
- Developed an automate evaluation metric to assess the result using GPT-4 as a judge
- Utilized HPC for finetuning and inferencing Chatbot such as LANTA, TARA, and NT-SERVER
- Optimized inference runtime with TensorRT-LLM via Docker in linux environment
- Deployed a chatbot using FastAPI and Docker on LINE official account application
- User preference evaluation using A/B testing in Chatbot Arena interface
- Cooperated with Datamind Lab in Chulalongkorn University and NECTEC Thailand as co-advisor

Projects

Traffy Fondue Clustering, 2023

- Utilized PyTorch to develop an EfficientNet model for image classification on Kaggle dataset
- Developed a data pipeline using numpy, scipy, pandas, Airflow, and containerized using Docker
- Developed a machine learning pipeline using MLflow and Scikit-Learn to perform K-means clustering
- Visualized through interactive dashboard using folium, seaborn, and streamlit running on localhost
- **github.com/Bhuribhat/Traffy-Clustering**

Text Classification, 2023

- Developed a language detection model using Scikit-Learn for language identification
- Utilized TensorFlow to develop a LSTM model for text classification
- Created REST API using FastAPI and dockerized web application for deploying using Docker
- **github.com/Bhuribhat/Text-Classification**

Real-Time Car Detection, 2022

- Gathered Kaggle dataset to create cascade classifier model and background subtraction algorithm
- Utilized OpenCV to perform image processing for real-time detecting and counting on video footage
- **github.com/Bhuribhat/Vehicle-Detection-Project**

Discord Bot, 2020

- My personal python-based discord assistant bot running on my own server
- Scrape websites for filtered job listings using BeautifulSoup
- Automated the process of sending daily schedules and important assignment notifications
- **github.com/Bhuribhat/Discord-Bot**

Skills

- Software knowledge: MS Office, MS Excel, MS PowerPoint, Power BI, Tableau, Gephi
- Programming Languages: Python, Java, JavaScript, C, C++, Scala, R, SQL
- Language: Thai (Native), English (CU-TEP 77 / 120), basic Chinese and German

Extracurricular activities

- Teacher Assistant in NLP Individual Study, 2023
- Individual Study PET-CT brain scan, 2023
- Member of CU Basketball Club, 2020 - 2021
- Member of CU Game Developer Club, 2020
- Hobbies: engage in gaming, basketball, movie, and music