

BUI CHI MINH

Ha Noi, Viet Nam

(+84) 886870997 ♦ buichiminh.cntt@gmail.com

EDUCATION

Posts and Telecommunications Institute of Technology

2015 - 2020

Major

Bsc, Computer Science

Programming languages and courses :

Python , JavaScript/Java , SQL , Data Structure , Algorithms

VNU University of Engineering and Technology

2023 - now

Major

Msc, Computer Science

Programming languages and courses :

Natural Language Processing using Deep Learning , Algorithms

EXPERIENCE

AI Engineer, Viettel Group

Jun 2022 -now

1. News categories classification

- Developed and optimized the model for CPU servers with improved to 85% accuracy overall and latency less than 30% compared with old proposed models.

- Automated the labelling phase, the application has effectively diminished the financial and labor investments required from individuals and already integrated into the website system.

- Designed model for recommendation news focus on private customers.

Product: The application has successfully integrated into the Reputa system to serve the other customers for social listening purposes.

2. End2End QA system

- Designed deep learning models for a end to end question answering system (top3 in a AI challenge).

- Leveraged and utilized the GPUs and optimized code for cross- training and parallel training.

- Implemented bi-encoder and tested cross-encoder for the retrieval phase.

- Built docker for this application push to company hub.

3. Court Judgement for Predictions and Explanation

- Used some big sparse transformer models combined with lexical model to enrich the information and improved the accuracy for the task.

- Reached top 1 in the competitions with a scientific paper accepted in SemEval (ACL)

4. Research about LLM

- Experimented the RAG pipeline included : finetuning LLM models and augmented the data for specific usecases (legal qa) incorporated with langchain. Eg : BloomZ, Llama2, MPT, ...

- Build a wrapper for evaluation tool using HELM with custom datasets.

- Deployed and optimized the speed inference of Llama2 using vLLM, exllama.

Data Scientist, Vingroup

Jun 2021 - Jun 2022

1. Tracking human workforce using vband

- Worked with Human Activity Recognition (HAR) problems to helped the manager tracking employees workforce. (<https://news.microsoft.com/vi-vn/2021/09/06/vantix-ung-dung-tri-tue-nhan-tao-nang-cao-nang-suat-lao-dong/>)

- Used various timeseries model and approached to improved the accurate of the predictions. (Worked with RNN, CNN, ensemble models and XGBoost)

- Worked with GPS correction problems for tracking worker in the construction (applied in VinHomes

construction).

- Used Kalman-Filtering algorithms to improved the roadmap for GPS and also reduced the GPS flashing, flicking problems.
- Built a FastAPI demo for map tracking and dockerized the application into production.

Product: The application has served all the VinPearls and some VinHomes projects.

AI Engineer, Viettel Group

Jan 2020 - Jun 2021

1. Duplicate news and video recommendation

- Designed an algorithm for video recommendation within the stipulated time. Used and applied tensor-flow as main application and flask for api. Worked with a 5-members research team for accomplishing this.

- Increased the average time per devices by 5-10% and the total amounts of time by 13% by A/B testing.

- Detected duplicate news cover 95% accuracy over a thousand news per day. This helped the moderators to reduced 80% working hours to check the duplicate ones.

Product: The features has applied in Mocha (app for Android and IOS) and Netnews.vn

2. Auto tuning firewall project

- Developed a system that detect anomaly user behaviors for auto tuning WAF system.

- Built a system that collected the data and detect bad requests to reduce false positive from traditional WAF (Web ApplicationFirewall).

- Built a profile prototype and a model to tracked IP to rank users rating.

- Reduced the human workforce in monitoring with an acceptable accuracy.

- Applied quantization to the model to reduce the latency (support around 100 mils user requests per day).

- Developed a DNS monitor module (tracking DNS to alert redirect site) using cron job.

Products: The application has integrated into the WAF and served more than 100 millions requests per day and denied around 100 to 1000 IP attackers per day.

SKILLS

Machine Learning

- Proficiency: Natural Language Processing
- Intermediate: Data analytics, visualization

Programming Languages

- Proficiency: Python
- Intermediate: Java, bash
- Basic: JS, SQL

Tools and Frameworks

- Proficiency: Pytorch, Git, Docker
- Intermediate: Tensorflow, Keras, Flask, FastAPI

Database & Cloud

- Proficiency: MySQL
- Intermediate: MongoDB
- Basic: AWS, Azure

Others

- Crontab, Ubuntu, Excel, Web dev, airflow

LANGUAGES

- **Vietnamese : Mothertongue**
- **English : Toeic/CEFR B2**

AWARDS AND SCHOLARSHIPS

SemEval-2023 Task 6: LegalEval , ACL	<i>2023</i>
Awards: Top-1 ranking in the subtask C1 (Legal Judgment Prediction) and top-2 ranking in the subtask C2 (Court Judgment Prediction & Explanation).	
Zalo AI Challenge 2022. , Zalo	<i>2022</i>
Awards: Impressive Solution, Top 3/103.	
Task: End-to-end Question Answering, finding an answer for each question in Vietnamese Wikipedia Corpus.	
BKAV Online CTF competitions , BKAV	<i>2020</i>
Awards: Top 5 leaderboard for CTF.	
Scholarship for Global Problems Based Learning (gPBL 2019) , Shibaura University	<i>2019</i>
PTIT CTF , PTIT	<i>2018,2019</i>
Awards: Third prize and second prize	
PTIT ACM/ICPC , PTIT	<i>2018,2019</i>
Awards: 2 Consolation prizes	

PUBLICATION

1. Thanh Dat Hoang, Chi Minh Bui, Khac Hoai Nam Bui. "Viettel-AI at SemEval-2023 Task 6: Legal Document Understanding with Longformer for Court Judgment Prediction with Explanation". Proceedings of the 17th International Workshop on Semantic Evaluation (SemEval-2023).
2. Thanh-Do Nguyen, Chi Minh Bui, Thi-Hai-Yen Vuong and Xuan-Hieu Phan. "Passage-based BM25 Hard Negatives: A Simple and Effective Negative Sampling Strategy For Dense Retrieval". Proceedings of the 37th Pacific Asia Conference on Language, Information and Computation. (PACLIC 37)