VU THANH TUYEN

AI RESEARCH ENGINEER

Hanoi, Vietnam • 0896023142 • tuyenvt455@gmail.com • https://github.com/SuZeAI

As a responsible and passionate individual, I am proactive in seeking knowledge and learning, always open to feedback from others. I continuously strive to accumulate knowledge and experience with the goal of becoming a TechLead AI Research Engineer. I hope to fully leverage my capabilities while working at the company.

WORK EXPERIENCE

AI Research Engineer, Viettel Cyber Security — Jan 2024 - Dec 2024

- Developed Self-Learning and Problem-Solving Skills: Cultivated the ability to autonomously learn, explore new concepts, and identify unresolved challenges, while enhancing personal soft skills.
- Pretrained Language Models (PLMs) Implementation: Researched and applied various Pretrained Language Models, including BERT, RoBERTa, GPT-2, and T5, to key Natural Language Processing (NLP) tasks such as Named Entity Recognition, Text Classification, and Text Summarization.
- Intelligent Vulnerability Detection in Code: Conducted surveys to identify issues, followed by comprehensive reviews of relevant research papers and methodologies. Collected and analyzed data to construct an appropriate model. Utilized graphs such as Abstract Syntax Trees (AST), Control Flow Graphs (CFG), and Data Flow Graphs (DFG), employing tools such as PyTorch, PyTorch Lightning, Scikit-learn, and Torch Geometric for model construction, along with configuration tools like Hydra and unit testing using Pytest.
- Chatbot Development: Engaged in research and development of Knowledge Distillation (KD) models from large models like GPT and Gemini, aiming to identify suitable methods for data generation, model construction, and effective benchmarking for chatbot applications.

AI Researcher, PTIT IEC & Naver Lab — Sep 2023 - Present

- Foundational Knowledge in Natural Language Processing (NLP): Acquired a robust understanding of NLP through advanced coursework, including Stanford's CS224N and CS224U, mastering essential topics such as data preprocessing, word segmentation, Word2Vec, and Transformer-based models.
- Participation in Knowledge Seminars: Attended weekly seminars to present findings and engage in discussions on emerging knowledge in the field.
- Support for GCN-Computer Vision Project: Contributed to data processing related to object relationships within images, utilizing tools such as NumPy, Pandas, and JSON for effective data manipulation and filtering.
- Malware Detection in Android Applications: Co-led a project employing Graph Neural Networks (GNNs) with fine-tuned Word2Vec to generate enriched vector representations of function names in call graphs, leveraging shared naming conventions in malware. Achieved an impressive 98.8% F1-score in performance. Collaborated closely with a co-researcher and a Master's level advisor.
- Utilization of Advanced Frameworks: Leveraged frameworks such as PyTorch, Torch Geometric, Deep Graph Library (DGL), Google Colab, and Kaggle for model training, evaluation, and analysis.

Personal Projects

- AI Tetris Project: Developed a model within the Tetris game environment using genetic algorithms and Q-learning, achieving recognition and winning an award at the AI Contest PTIT competition.
- LLM Pipeline Development for Slide Generation: Fine-tuned the LLaMA 3.1 8B and Mistral NeMo 12B models on a custom dataset to automate the generation of HTML slides. Planned enhancements include implementing image retrieval capabilities for presentations.
- ADM Permissions 2024 Project: Focused on malware detection in APK files by employing feature extraction methods through large language models such as BERT and RoBERTa, moving away from traditional graph-based approaches to create features suitable for recognition by conventional models.
- Comment Classification: Customized and applied the BERT architecture for the classification of comments, determining whether they are positive or negative in sentiment.

EDUCATION

Bachelor : Information Technology — 2021 - Present

Posts & Telecommunications Institute of Technology (PTIT) - Hanoi
GPA: 3.23

SKILLS

- **Research:** Proficient in reproducing, fine-tuning, and extending research on Large Language Models (LLMs) for customized applications.
- **Practical Applications:** Specializes in Text Generation, Question Answering with Retrieval-Augmented Generation, and Contextual Understanding utilizing LLMs.
- **Technical Proficiency:** Skilled in using tools such as NumPy, Pandas, Scikit-learn, and advanced NLP libraries for efficient model training and evaluation.
- **Collaboration Skills:** Strong collaboration and communication abilities, particularly within LLM-focused teams, with a demonstrated aptitude for identifying and researching problems.
- **Programming Languages:** Proficient in Python, C, C++, and Java.
- **Large Language Model (LLM) Expertise:** Extensive experience in developing and fine-tuning LLMs, including LLaMA, Mistral, RoBERTa, and Sentence-BERT, for a variety of NLP tasks.
- **Deep Learning Proficiency:** Well-versed in Transformer architectures, Recurrent Neural Networks (RNNs), and Long Short-Term Memory (LSTM) networks, with a strong emphasis on large-scale NLP applications.
- **Frameworks and Tools:** Skilled in utilizing frameworks such as PyTorch, TensorFlow, Hugging Face Transformers, Unsloth, and TensorRT for the development and deployment of LLMs.

ACHIEVEMENTS

- **Mathematics Awards:**
  - Encouragement Prize in Mathematics at the district level (7th grade).
  - 1st Prize in Mathematics at the district level (8th grade).
  - 2nd Prize in Mathematics at the provincial level (12th grade).
- **Academic Achievements:**
  - Received a PTIT scholarship for first-semester GPA of 3.6/4 (Excellent).
- **Competitions:**
  - Awarded a prize in the AI Contest held at PTIT in 2022.
- **Certifications**:
  - Obtained Samsung's Python for AI certification (2022).
  - Completed the "Crash Course on Python" certification from Google.
  - Certified as a member of the organizing committee for events at JAIST and PTIT.
  - Served as a mentor for the Samsung for AI course organized by PTIT in collaboration with Samsung.