**PHAM MINH HOANG**

+84 394205905 • hoang1.phamminh@toshiba.co.jp

Summary

I am currently working at Toshiba Software Development Vietnam (TSDV) as an AI Engineer with 3 years of experience.

Experience

**Toshiba Software Development Vietnam (TSDV)** Hanoi/Vietnam

AI Engineer October 2022 – Present

Responsible for researching and developing application (Proof-of-concept level) for AI-related solutions.

* ***Advanced Chatbot Development and Natural Language Processing***: Developed LLM-based conversational systems leveraging RAG (Retrieval Augmented System), including:
  + RAG system with database using agent-based architecture that can function as both normal conversational chatbot and database analytic specialist with on-premise open-source LLMs, achieving over 85% improvement in response accuracy compared to starting baseline.
  + Graph-RAG system for technical Japanese documents utilizing graph technique, with special developed chunking strategy for maximizing structural information preserving.
  + Graph VQA (Visual Question Answering) with multi-modal LLMs, capable of producing non-trivial analyzing for multi-variate time-series data that has been converted to graph form.
* ***MLOps Infrastructure and Platform Development***:
  + Deployed in-house MLOps platform on-premise and provided alternative solutions for outdated services.
  + Migrated existed end-to-end workflows for model training process onto Toshiba Kubeflow-based self-developed MLOps platform.
  + Customized automated large-scale data versioning control (DVC) workflow with LakeFS.
  + Streamlined the training outputs and artifacts logging by utilizing the open-source platform MLFlow.
* ***Research and Proof of Concept (PoC) Development***:
  + Investigated and built Android mobile PoC application for real-time HAR (Human Activity Recognition) through available sensors on smartphone such as triaxial accelerometer and triaxial gyrocope using Deep Neural Network (DNN).
  + Investigated and built Windows PoC application using Windows .NET framework for real-time SED (Sound Event Detection) that can take audio input from both microphone and playback stream. To maximizing the inference speed, developed DNN models was deployed in C through Tensorflow Lite library.
  + Built Linux-Windows cross-platform compatible PoC application that integrates both SED and Driver Drowsiness Detection (DDD) using self-developed models on open-source dataset. The application was developed with PyQt and FFMPEG.

Education

**Hanoi University of science and technology**

* Major: Electrical and Telecommunication
* GPA: 3.3/4

Additional

* Academic incentive scholarship of Hanoi University of Science and Technology (2019-2022)
* TSDV Best Engineer of 23B term