Hai-Nam V. Cao

291 Daehak-ro, Daejeon, South Korea

Email LinkedIn Github Website

Education

Korea Advanced Institute of Science and Technology (KAIST) **Aug. 2020 – Aug. 2024**

B.S. in Computer Science *Daejeon, South Korea*

- **GPA**: 3.64/4.3
- Certificate in Artificial Intelligence

Experience

Undergraduate Researcher **June 2023 – Dec 2023**

KAIST Interaction Lab, Advisor by Prof. Junho Kim *Daejeon, Korea*

- Conducted research on on two specific topics around natural language interactions with charts: (1) chart question-answering (CQA) and (2) chart captioning.
- Design a version of code language specifically for chart analysis covering some of the following aspects of human-computer interaction: multimodal interaction (e.g., chart, text, mouse gestures), intent disambiguation, explanation generation, and human-readable presentation of results & explanations.

Web Development Intern **June 2022 – Sep 2022**

Onsquare Co. *Seoul, Korea*

- Developed a feature for editing images and Scalable Vector Graphic (SVG) objects with a range of functionalities similar to those found in the Photoshop app using fabricJS, ImageMagick framework.
- Employing modern type theory for all the objects created in Looina project including image, sound, spreadsheet, .etc.
- Writing documentation for most of the implemented functionalities.

Individual Study **Dec 2021 – Jan 2022**

KAIST Data Intelligence Lab *Daejeon, Korea*

- Conducted research on fairness and robustness in unstructured data. Preprocessing and cleaning data to minimize discrimination and bias in machine learning models.

Projects

EvolveUniTest **Oct 2023 – Dec 2023**

Final project for CS453 - AI Based Software Engineering, KAIST *Github repo*

- Developed an innovative unit test generation system utilizing LLMs and incorporating genetic algorithm.

Personality recognition **Feb 2023 – June 2023**

Qualcomm-KAIST Kaggle Hackathon 2023 *Kaggle competition*

- Developed the model for personality recognition from personaltiy(mbti)-labeled text data of 15,000 Korean question-answer pairs, outperformed the baseline accuracy by 23%

Emoji Recommendation **Aug 2021 – Dec 2021**

Final project for CS492 - Introduction to deep learning, KAIST *Github repo*

- Developed a sentence-based emoji suggestion feature using a fine-tuned BERT model and deployed it as a REST API using FastAPI.
- Created a web-based chatting-box application to showcase the model's capabilities.

Semi-Supervised Semantic Segmentation **Aug 2022 – Dec 2022**

Final project for CS470 - Introduction to AI, KAIST *Poster Code*

- Replicated results from the original paper: Semi-Supervised Semantic Segmentation with Cross-Consistency Training
- Improved the result of the paper with GAN network and proposed other method of Temporal Ensembling.

Robustness of Reading Comprehension Models to Entity Renaming **Aug 2022 – Dec 2022**

Final project for CS475 - Machine learning for NLP, KAIST *Final report*

- Conducted experiments to evaluate and compare the robustness of BERT, RoBERTa, and SpanBERT against entity renaming across five distinct datasets.

- Proposed a novel improvement method for model's robustness, namely Anonymized Training with optional Anonymized Inference.
- Achieved a substantial mitigation in performance drop from over 30% to less than 5% for SpanBERT against entity renaming, with BERT and RoBERTa's performance drop being further reduced.

Shopee Price Match Guarantee **Mar 2022 – June 2022**

Final project for CS376 - Machine learning, KAIST *Github repo*

- Developed a product matching system that extracts image and text embeddings to determine if two products are the same based on their images and descriptions.
- Implemented and integrated the triplet loss function into the training of ResNet18, resulting in a significant improvement of the model's validation F1 score from 60% to 71%.
- Achieved the highest validation F1 score of 72.46% for images matching and 68.27% for titles matching.

Awards

Qualcomm-KAIST Innovation Award (2023): Winner.
KAIST International Undergraduate Scholarship (2019): Full-ride Scholar.
Vietnam National Mathematical Olympiad (2019): Bronze medal.
Vietnam National Mathematical Olympiad (2018): Bronze medal.

Technical Skills

Fields of Interest: Software Engineering, Machine learning/Deep learning
Programming languages: Python, C/C++, TypeScript, JavaScript, SQL, PHP, HTML, CSS
Technologies/Frameworks: PyTorch, Linux, Git, Vim, Pandas, NumPy, Scikit-learn, Matplotlib, Tensorflow, FabricJs, TensorflowJs, RE:DOM
Certificates: 7.5 in IELTS, TOPIK level 2, 800 in SAT subject Math and Chemistry.