Hai-Nam V. Cao

291 Daehak-ro, Daejeon, South Korea

Email LinkedIn Github Website

Education

Korea Advanced Institute of Science and Technology (KAIST) **Daejeon, South Korea**
Bachelor of Science, Advanced major in Computer Science *Aug. 2020 – May 2024*

- **GPA**: 3.64

Phan Boi Chau High School for the gifted **Vinh, Viet Nam**
Specialize Math class *Aug. 2016 – May 2019*

Experience

Undergraduate Researcher **Daejeon, Korea**
KAIST Interaction Lab *June 2023 – Dec 2023*

- 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 **Seoul, Korea**
Onsquare Co. *June 2021 – September 2022*

- Employing modern type theory for all the objects created in Looina project including image, sound, spreadsheet, .etc.
- 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.
- Writing documentation for most of the implemented functionalities.

Individual study **Daejeon, Korea**
KAIST Data Intelligence Lab *Dec 2021 – January 2022*

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

Projects

EvolveUniTest **Github repo**
Final project for CS453 - AI Based Software Engineering, KAIST *Oct 2023 – Dec 2023*

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

Personality recognition **Kaggle competition**
Qualcomm-KAIST Kaggle Hackathon 2023 *Feb 2023 – Present*

- 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 **Github repo**
Final project for CS492 - Introduction to deep learning, KAIST *Aug 2021 – Dec 2021*

- 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 **Poster code**
Final project for CS470 - Introduction to AI, KAIST *August 2022 – December 2022*

- 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 **Final report**

Final project for CS475 - Machine learning for NLP, KAIST *August 2022 – December 2022*

- 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 **Github repo**

Final project for CS376 - Machine learning, KAIST *March 2022 – June 2022*

- 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.

Skills

Fields of Interest: Software Engineering, Machine learning/Deep learning
Languages: Vietnamese (native), English (IELTS: 7.5).
Programming languages: Python, C/C++, TypeScript, JavaScript, PHP, HTML, CSS, MySQL
Ability to utilize libraries: PyTorch, Pandas, NumPy, Scikit-learn, Matplotlib, Tensorflow, FabricJs, TensorflowJs, RE:DOM