

# Hai Nam Cao Viet

✉ [hainamkaist@gmail.com](mailto:hainamkaist@gmail.com)  [hai nam cao viet](#)  [MrClumsy](#)

## EDUCATION

### Korea Advanced Institute of Science and Technology (KAIST)

*Bachelor of Science, Advanced major in Computer Science*

– GPA: 3.62

Daejeon, South Korea

*Aug. 2020 – May 2024*

### Phan Boi Chau High School for the gifted

*Specialize Math class*

Vinh, Viet Nam

*Aug. 2016 – May 2019*

## EXPERIENCE

### Individual study

*Data Intelligence Lab at KAIST*

[Web page](#)

*Dec 2021 – January 2022*

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

### Web development Intern

*Onsquare in Seoul, Korea*

[Project: Looina](#)

*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.

## PROJECTS

### Emoji Recommendation

*Final project for CS492 - Introduction to deep learning*

[Github repo](#)

*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

*Final project for CS470 - Introduction to AI*

[Poster code](#)

*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 project for CS475 - Machine learning for NLP*

[Final report](#)

*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

*Final project for CS376 - Machine learning course*

[Github repo](#)

*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

### Vietnam Mathematical Olympiad 2018 and 2019

*Ministry of Education and Training*

*Jan. 2018 and Jan. 2019*

- 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