YU-HUNG WU

Graduate Institute of Communication Engineering, National Taiwan University

yhwuntu@gmail.com ◇ https://yhwu.cc

EDUCATION

National Taiwan University (NTU) — Sep. 2024 - Present
M.S. in Graduate Institute of Communication Engineering — Advisor: Shao-hua Sun

National Taiwan University of Science and Technology (NTUST) — Sep. 2017 - Jun. 2021
B.S. in Computer Science and also Double Major in Finance

- **Computer Science major GPA: 4.05**, last 2-year GPA: 4.05, cumulative GPA: 3.88
- I participated in the following undergraduate research:
    1. BERT-Based Visual Question Answering (VQA) (Advised by Dr. Kuan-Yu Chen)
    2. Video Activity Recognition in Sports Videos (Advised by Dr. Bi-Ru Dai)
    3. High-Frequency Trading (An interdisciplinary project in CS and finance. Advised by Dr. Wei-Zhong Miao)
- Got **4th place** among 60 teams in the undergraduate research competition of the computer science department.

ACADEMIC RESEARCH EXPERIENCE

Full-Time/Part-Time Research Assistant — Aug. 2021 - Feb. 2024/Mar. 2024 - Aug. 2024
CKIP Lab, Institute of Information Science, Academia Sinica — *Taipei, Taiwan*

- I am involved in the following projects:
    1. *Dynamic-Window Transformer*: The model adopts a differentiable method to select different window sizes for each input token, and reaches memory efficiency by adopting the sliding-window algorithm.
    2. *Data Augmentation on Visual Question Answering (VQA)*: I devised a method to generate VQA samples from existing datasets like VQA 2.0 and trained a discriminator to evaluate the quality of each sample.
    3. *Bible NER*: To assist Christians in gaining a deeper understanding of the Bible, we developed a system that annotates various labels within its text. We collaborated with the Sagos organization in the United States.
    4. *Chinese Llama-2 Pretraining*: I became part of the team focused on pretraining LLMs. Given the significance of data selection during the pretraining phase, we are exploring strategies to select data samples that align with specific patterns.

WORK EXPERIENCE

Machine Learning Engineer Internship — Jul. 2020 - Jun. 2021
Industrial Technology Research Institute (ITRI) — *Taipei, Taiwan*

- Responsible for developing a Deep Reinforcement Learning (DRL)-based dialogue system. Our goal is to develop a new agent to solve dialogue policy learning, which is based on Q-Learning.

Software Engineer Internship — Sep. 2018 - Mar. 2020
Asia Pacific Machine Intelligence Company (AP-MIC) — *Taipei, Taiwan*

- AP-MIC is a business partner with Google, which focuses on developing chit-chat dialogue engines according to user requirements. For example, users can choose different "chatbot skills" to design their chatbot.
- I developed the concatenated system between Actions on Google (AoG) and our platform, letting the users create an AoG chatbot in an end-to-end fashion on our platform.
- At the early stage of the coronavirus pandemic, I developed a COVID-19 chatbot, which had reached 82,000 conversation sessions and 71% user return rate.

TEACHING EXPERIENCE

Teaching Assistant   Sep. 2020 - Jan. 2021
Data Structures Course in NTUST (120 Students)   *Instructor: Prof. Kuan-Yu Chen*

- Invited to join the teaching team in Data Structures course because I got "A+" when I took this course.

- I designed innovative homework assignments and helped students to solve the problems.

HONOR & AWARDS

- **Top 1%** (5th place among 409 participants), E-sun Tbrain NLP Competition   Aug. 2020

- **Top 2%** (Top 12 teams among 501 participants), Post Office Big Data Competition   Jun. 2019

- **Top 6%** (4th place among 60 teams), Undergraduate Research Competition   Jan. 2021

INVITED TALKS

- **"Can Emotional Labor be Outsourced to AI?"** at Eslite Xinyi Store, Taipei, Taiwan   Nov. 2023

LEADERSHIP EXPERIENCE

- **Class representative** of computer science department class 2020, which is my junior year.

- **Representative** of the computer science department in freshman orientation.

- **Team leader** of several class projects, e.g. full-stack development course, data science course, operating system course, and software engineering course.

SELECTED RESEARCH PROJECTS

Dynamic Window Transformer on Long Sequence Modeling   Supervised by Dr. Wei-Yun Ma

- While Transformers excel across various benchmark datasets, they frequently grapple with quadratic complexity in terms of time and memory usage. To address this, we employ the sliding window method, enabling Transformers to handle long sequences.

- However, the window size is currently determined heuristically. To improve upon this, we've devised a differentiable approach that enables the learning of varied window sizes for different tokens.

- We adopted the Query Chunking and Input Checkpointing strategies to reduce peak memory usage in the Transformer architecture. With these adaptations, we can seamlessly integrate our method into the Transformer.

- The results from both Question Answering and Translation outperform the baseline performance.

Data Augmentation on Visual Question Answering   Supervised by Dr. Wei-Yun Ma

- Visual Question Answering (VQA) is a task in computer vision and NLP that involves answering questions about an image. Our objective is to implement data augmentation techniques for this task.

- We begin by identifying potential "question types" for a given image through "question type embedding". Subsequently, we employ a generation model to produce QA pairs based on the image.

- Additionally, we trained a discriminator to selectively retain high-quality QA pairs.

Video Activity Recognition in Sports Videos   Supervised by Prof. Bi-Ru Dai

- We developed a CNN-based model specifically designed to detect distinct motions within sports videos.

- In our project, we have the capability to annotate multiple motions along with their durations within a single video. We've crafted an algorithm that not only detects transition animations but also identifies the specific sport featured in each segment.

- We employed an ensemble of several models and crafted a pipeline approach specifically tailored for processing long sports videos.