YAN Yuhang (Henry)

https://yany-henry.github.io | yanyuhang2002@link.cuhk.edu.hk | +86 18945222225

Education

The Chinese University of Hong Kong (CUHK) Sep. 2021 - Jul. 2025 (expected)
Bachelor of Science in Computer Science (CSCI) Hong Kong SAR, China

- Cumulative GPA: **3.757**/4.000 (**Top 10%**); A member of **S.H. Ho College**
- **Dean's List** of Faculty of Engineering; **ELITE Stream** of Faculty of Engineering
- **Spring semester exchange** to École Polytechnique Fédérale de Lausanne (**EPFL**), 2023-24
- **Spring semester exchange** to Yuanpei College, Peking University (**PKU**), 2022-23

Research Experience

Evaluation on the Vulnerability of Current Generative Models Jan. 2024 - present
Research Assistant to Prof. Sabine Süsstrunk (EPFL) Lausanne, Switzerland

- **"Jailbreak" Analysis & Security Enhancement**: Experimented with more than ten large language model (LLM) "jailbreak" techniques to identify persistent vulnerabilities. This led to help technology companies develop enhanced defence mechanisms that improve the security of their LLMs and the ability to resist the generation of harmful content.
- **Advanced Alignment Circumventing Techniques**: Utilized reinforcement learning and greedy search to identify alternatives to sensitive words within the latent space to circumvent the alignment of the LLMs. These techniques circumvent traditional alignment mechanisms, exposing critical security vulnerabilities within LLMs.
- **Attention Diversion Techniques for LLMs**: Designed an array of intricate input prompts and conducted extensive experiments with potentially harmful statements. These efforts effectively diverted LLM attention, facilitating precise identification of weaknesses and enabling targeted attack strategies.

Efficient Video Analytics Jun. 2023 - Sep. 2023
Research Assistant to Prof. Eric Chi Lik LO (CUHK) Hong Kong SAR, China

- Won the **Best Project Award 2023** among 58 undergraduate projects.
- **Multi-modal AI System for Lost-and-Found at Hong Kong International Airport**: Utilized CLIP and OWL-ViT models to craft a multi-modal lost-and-found system, harnessing user texts and images to search through extensive video data, resulting in a significant advancement in applying AI to improve airport services.
- **Custom Dataset and AI Model Evaluation**: Created a tailored airport video dataset and established multi-modal algorithm evaluation standards. This effort led to the selection of high-performing multi-modal Zero-Shot Object Detection and NLP models, saving the airport lost-and-found system processing time by 5-8 hours.
- **Frame Detection Optimization with Increased Tolerance**: Developed an algorithm to identify potential frames for lost items, overcoming challenges in pinpoint accuracy by experimenting with tolerance levels. This improved fault tolerance and made searching through large video footage more efficient, while ensuring high accuracy and recall rate.

Traffic Characteristics Analysis of the Network Apr. 2023 - Jun. 2023
Research Assistant to Prof. Tong YANG (PKU) Beijing, China

- **Traffic Characteristics Analysis and Platform Development**: Leveraged a high-performance traffic analysis platform built from scratch. Conducted detailed statistics and analysis of network traffic characteristics at the link, traffic, and packet levels using core traffic data to recognize campus network user preference.
- **TCP/IP and DNS Packet Analysis for Data Mapping**: Utilized campus network traffic data to extract and map IP addresses to domain names from pcap files, focusing on DNS packet structures. This task involved detailed examination and application of TCP/IP protocols to analyze network traffic patterns.
- **Campus Network Traffic Analysis**: Identified Top-K domains with the highest traffic volumes to understand user preferences in the campus network. Utilized C++ to parse DNS-containing pcap files, mapping IP addresses and domain names, analyzing network usage patterns, including peak request times and high-traffic sites.

Work Experience

AIJobTech Oct. 2023 - present
CTO & Co-founder Hong Kong SAR, China

- **AI-Driven Job Matching Startup Launch and Funding**: Implemented AI recommendation systems algorithm to facilitate precise job matching for business positions.
- **Resume Optimization with Large Language Models**: Led the deployment of large language models to polish resumes and personal statements, crafting a full-stack platform that elevates job search success and user satisfaction.
- **Technical Leadership in Product Development**: Led the technical team in developing a business-focused and AI-powered job matching platform, integrating innovative tech to enhance product quality and market alignment.

Skills

Languages	English (Fluent), Mandarin (Native) and Cantonese (Intermediate).
Programming	C/C++, Java, JavaScript, MATLAB, Python{PyTorch, TensorFlow}, Scala and SQL.
Tools	LaTeX, Anaconda, Git, Linux and Spark.