Junhao RAN

Junhao.Ran@rice.edu | LinkedIn | GitHub | (+1) 346 479 2175 | Houston, TX

RESEARCH VISION

Understand Deep Learning and devise powerful & reliable deep learning solutions to real-world challenges

EDUCATION

Rice University, Houston, TX 08/2023 - 12/2024 (Expected)
Master of Electrical and Computer Engineering, Data Science Track

The Chinese University of Hong Kong, Shenzhen (CUHK-SZ), China 09/2019 - 06/2023
B.Eng. in Computer Science and Engineering
Honors: Bo Wen Scholarship, Dean's List, Undergraduate Research Awards

NLP INTERNSHIP

Ping An Technology, China | *NLP Algorithm Engineer Intern* 02/2023 - 07/2023
Topic: End-to-End Topic Segmentation with Transformer for Long Sequence

- Contributed to the construction of an in-house automatic video editing platform, spearheading investigations into effective methods for topic segmentation
- Applied **Transformer for long sequences – Big Bird** to topic segmentation, showcasing its superiority over traditional hierarchical architectures and achieving state-of-the-art (SOTA) performance – Pk:14.1% on the Wiki727k Dataset

Topic: Early Rumor Detection with Reinforcement Learning

- Headed the Early Rumor Detection project, a research project aimed at rapidly identifying and limiting the spread of rumors
- Constructed a comprehensive model that interprets social media posts (e.g., tweets) as data streams. This model encompasses two pivotal modules:
  - Rumor Detection Module: Utilized **Transformer for long sequences – Big Bird** to classify events, represented by a series of posts, determining their validity as rumors
  - Checkpoint Module: Employed **reinforcement learning** with **deep Q-learning** to dictate the optimal activation timing for the rumor detection module, ensuring rumors are identified at their inception
- Achieved a competitive accuracy of 83.30% in early rumor detection using, on average, just 4.7% of an event's posts. This performance nearly matches the 83.96% accuracy of the Rumor Detection Module when processing the full set of posts

COMPUTER VISION RESEARCH

Shenzhen Research Institute of Big Data, China | *Research Assistant* 08/2021 - 01/2023
Topic: Hierarchical Transformer with Knowledge Guidance for **Fine-Grained Image Recognition** (FGIR)

- Headed a research project focusing on enhancing FGIR by integrating **knowledge graphs** with **vision Transformer**, bridging knowledge-driven and data-driven methods
- Developed the model Hierarchical Transformer with Knowledge Guidance (HTKG) for FGIR. The model utilizes a **Swin Transformer** Encoder to process image features. It further combines these features with a bird category-attribute bipartite knowledge graph, aiming to capture detailed category and attribute distinctions within images
- Benchmarked HTKG on the Caltech-UCSD Birds-200-2011 (CUB-200-2011) dataset, achieving an accuracy of 92.47%
- **Award:** Undergraduate Research Awards, CUHK-SZ

HANDS-ON PROJECTS & COURSEWORK

AthenaLite: Investigating the Training of Reasoning Capabilities in Tiny Language Models | *Co-developer* 05/2023

- Developed AthenaLite, a tiny language model aimed to achieve reasoning capabilities rivaling large models
- Constructed a rich, multi-tiered reasoning dataset using an automatic annotation pipeline based on **GPT-4** and **GPT-3.5-turbo**
- Adopted a multi-phase training approach for AthenaLite, initially utilizing diverse dialogue data, then escalating to more complex reasoning tasks
- Evaluated AtheLite using human assessments and observed its use of reasoning strategies for basic math questions, surpassing the baseline model, **Alpaca-LoRA**, but still trailing significantly behind GPT-4

Operating Systems | *Individual Projects,* Operating System Design and Implementation with **C** 11/2021

- **Multithreading:** Implemented a Jumper Frog game, created a char map to store the location information of frogs and logs, and created 9 threads to move them
- **Virtual Memory Simulation:** Created virtual memory for **NVIDIA GPU** with read, write, and memory snapshot functions
- **File System:** Designed a file system for **CUDA** GPU with open, read, write, deleting, and file view functions

SKILLS

ML Algorithms & DL Architectures: Reinforcement Learning / Transformer / CNN / RNN
Programming Languages: Python / Java / JavaScript / C / C++ / CUDA / SQL (MySQL) / Verilog
Libraries & Frameworks: PyTorch / Java Spark / NumPy / Pandas / Keras / Matplotlib / WandB / TensorBoard