Weizhi PENG

Vancouver, BC, Canada
Phone: +1 (778)-885-8117; Email: jimpeng98@gmail.com Website: pengwz.info/about/

Education

New York Institute of Technoloy (Vancouver) **Vancouver, BC, Canada**
MSc in Cybersecurity *2024.01 - now*

King's College London **London, UK**
MSc in Artificial intelligence *09.2021 – 01.2023*

University of Birmingham **Birmingham, UK**
Major in BSc Computer Science *09.2019 – 07.2021*
B.Sc. Computer Science with Honors, Class I. GPA: 4.25/4.25

Harbin Institute of Technology at Weihai **Weihai, China**
Major in BEng Software Engineering *09.2017 – 07.2021*
GPA: 85/100
Coursework includes Neural Networking, Machine Learning and Intelligent Data Analysis, Advanced Networking, Computer Vision and Imaging, Mathematic Modelling & Decision Making, Software Engineering, System Programming in C/C++, Algebra and Geometry, Discrete Mathematics, Probability and Statistics, Java Programming Design, Python Programming, Principle of Computer Organization, Data Structures, etc.

Work Experience

Shengqu Game

Machine Learning Internship **Shanghai, China**
Unreal Department – Artificial Intelligence Group *08.2023 – 11.2023*

- Led the development of a 7-billion parameter language model for code completion, using PyTorch. Improved user interaction by fine-tuning the model for instructional chatting.
- Streamlined the model training process with Accelerate and DeepSpeed in a Linux environment, using four A100 GPUs, and formulated a unique evaluation method to ensure model accuracy.
- Deployed the model on a high-performance server using VLLM for fast responses.
- Engineered a Visual Studio Code extension for real-time coding assistance with company API documents, mirroring GitHub Copilot's functionality, and added a chat feature using the fine-tuning model for user API queries.
- Github: https://github.com/Miraclove/DevAssistant

BOC International

Quantitative Machine Learning Internship **Shanghai, China**
Fintech Department *04.2023 – 07.2023*

- Design, training, optimization, deployment and trading algorithm design, development and optimization of time series based Transformer, GPT, BERT architecture for high frequency trading models.
- Analyze large amounts of financial market data, discover and interpret patterns, and work with a team of software engineers to embed the models into trading systems.

Research Experience

King's College London

Few-shot Learning for Text Generation **London, UK**
Researcher *10.2020 – 06.2021*

- In academic research, pre-training models such as T5, GPT2, Bart, etc are used to generate Table to Text generation by Few-shot Learning and memory is utilized to store pre-training knowledge and training instance selection is used to improve training performance and achieve higher performance.
- Pre-trained models are trained and tested using books, songs and other public datasets, evaluated and optimized using BLEU-4, ROUGE-4 and other methods.

University of Birmingham

Development of the COVID Corpus Website **Birmingham, UK**
Developer *10.2019 – 06.2020*

- Deployed NLP functions on a Django server, providing robust APIs for seamless analysis result retrieval. Spearheaded the NLP module development for the COVID Corpus website using the React framework, working closely with the development team to ensure accurate and effective data visualization.
- Cleaned and preprocessed COVID-19 research data with NLP pipelines for quality analysis. Applied LDA algorithms for effective topic modelling in large datasets. Implemented BERT-based summarization tools to distill detailed academic texts into brief overviews. Built a recommendation system using TF-IDF and document similarity to improve the ease of finding related research.
- Project Demonstration: https://www.miraclove.com/service/nlp-analysis/

Additional Skills

- **Computer Language:** Python, Java, C++, C, SQL, JavaScript, HTML, Haskell
- **Software:** Pytorch, Linux, Visual Studio, Eclipse, MATLAB, Wireshark, SQL Server, React, NumPy, Pandas, Django