NIE LUN YIU

Email: flitternie@gmail.com Website: flitternie.github.io

EDUCATION

The Chinese University of Hong Kong (CUHK) 08/2015-07/2020

B.B.A. in Professional Accountancy, Second Major in Computer Science

- Admitted at the age of 15 with grades skipped
- CGPA: 3.58 / 4.00, Computer Science Major GPA: 3.67 / 4.00
- Core Courses (Average GPA: 3.85/4.00): Computer Organization, Data Structures, Fundamentals of Artificial Intelligence, Fundamentals of Machine Learning, Text Mining Models & Application

RESEARCH EXPERIENCE

Pretrained Transformer for Source Code to Text Generation 11/2019-Present

Supervised by Prof. Wai Lam, in collaboration with researchers from CMU and NTU Singapore

- Incorporating abstract syntax tree (AST) and language model for contextualized pre-train of source code embedding, in order to improve the performance of Transformer on downstream source-code-to-text generation tasks (e.g. code summarization, commit message generation) under low-resource setting.
- Achieved the state-of-the-art performance with a preliminary result of 18.42% increase in BLEU-4 score and 20.56% increase in ROUGE-L score on commit message generation benchmark.

Multi-grained Attention for Domain Adaptive Sentiment Analysis 05/2019-Present

Supervised by Prof. Wai Lam

- Developing a multi-grained transfer network that integrates the sentence, aspect, and word representation with a hierarchical attention mechanism for cross-domain sentiment analysis.
- Designed research proposal independently and investigated the role of post-training and fine-tuning in the knowledge transferability of pre-trained BERT under low-resource setting.

Commit Message Generation with AST and Hybrid Ranking 09/2019-11/2019

Supervised by Prof. Wai Lam, in collaboration with researchers from NTU Singapore

- Proposed ATOM, which is the first hybrid retrieval-generation model that generates commit message from source code differences and achieves the new state-of-the-art results (BLEU-4 score increased by 30.72%).
- Incorporated AST into the attentional Seq2Seq model as the encoder to represent code differences by the path and leaf node changes in AST.
- Integrated the results from both retrieval-based method and NMT-based method via a hybrid re-ranking module that computes the similarity score between code difference and commit message with a CNN.

Stock Index Return Forecasting with BERT 08/2019-05/2020

CUHK Undergraduate Final Year Project, supervised by Prof. Leizhen Cai

- Researching on a domain-specific pre-train method that masks specific features instead of a whole token, in order to improve the performance of BERT in stock forecasting and improve the model's domain adaptability among different stock markets.
- Lead a team of 6 students in developing an open-source framework that transfers state-of-the-art model architectures from natural language processing domain to quantitative finance domain.

E-commerce Product Description Generation System 07/2019-08/2019

Research Project for Sinovation Ventures DeeCamp organized by Dr. Kai-Fu Lee

- Developed an extensible and customizable hybrid retrieval-generation system that generates product description from keywords.
- Implemented a bi-attentional Transformer that incorporated external knowledge base and allowed customized generation style based on user preference (e.g. function-oriented or appearance-oriented).

- Outperformed the existing systems developed by Alibaba and JD.com with an average increase of 82.19% in terms of overall quality score in a double-blinded evaluation by 181 disinterested human annotators.

WORK EXPERIENCE

Text Mining Group, Information Systems Laboratory, CUHK 05/2019-05/2020

Research Assistant, supervised by Prof. Wai Lam

- Replicated the previous state-of-the-art models as baselines, proposed and validated new research ideas for improving the current works by designing and doing experiments independently.
- Summarized latest papers and presented research findings to other researchers in lab's weekly meeting.

Institute of Computing Technology, Chinese Academy of Sciences 07/2018-08/2018

Research Intern, supervised by Prof. Yungang Bao

- Researched on existing blockchain technologies and developed smart contracts on Ethereum to boost the trading efficiency of a blockchain transaction system with an optimized user interface.

Shanghai Genus Information Technology 04/2017-07/2017

Part-time Research Intern

- Participated in the development of a quantitative trading execution platform for equity investors
- Implemented Long Short-Term Memory network for the simulation of equity market fluctuation

PUBLICATION

- Shangqing Liu, Cuiyun Gao, Sen Chen, **Lun Yiu Nie**, and Yang Liu. 2019. ATOM: Commit Message Generation Based on Abstract Syntax Tree and Hybrid Reranking. *arXiv preprint arXiv: 1912.02972.* (Under Review of IEEE Transactions on Software Engineering)
- **Lun Yiu Nie**, Zhicong Zhang, Cuiyun Gao, Pengcheng Yin, Yang Liu, and Wai Lam. 2019. TARGET: Pre-trained Transformer for Contextualized Code Commit Message Generation. (Submitted to ACL 2020)

AWARDS & ACHIEVEMENTS

- Dean's List (Top 10%) 2018 & 2019
- Wu Yee Sun College Master's List (1 awardee/major/class) 2019
- Department Academic Excellence Scholarship 2017 & 2018
- Champion - Hong Kong Big Data for Business Challenge 2019 (Student Stream) 2018
- EY Scholarship (2 awardees/year) 2016
- Wu Yee Sun College Global Learning Award 2016

LEADERSHIP EXPERIENCE

Founder & Leader, CUHK Machine Learning Study Group 02/2019-Present

Planned and taught workshops for undergraduate students on machine learning algorithms and deep learning models; Organized weekly meetings to study natural language processing research papers.

Debater, CUHK Mandarin Debate Team 10/2015-12/2018

Represented the university in international debate competitions and awarded remarkable prizes, including the champion of Chinese Debate World Cup 2019, Hong Kong Division, the champion and outstanding debater of 2018 Hong Kong Inter-collegiate Mandarin Debate Competition.

Co-founder, College Society of Business Administration 11/2015-02/2017

Organized large-scale activities with around 600 students engaged to enrich their extra-curricular lives.

INTERESTS AND SKILLS

Volunteer: Organized voluntary teachings to the students living in Chinese remote area in Yan Chak Service Learning Programme and CUHK Business School Service Trip.

Programming: Proficient in C, C++, Python; Familiar with Scikit-Learn, TensorFlow, PyTorch, etc.

Language: English (Fluent, TOEFL: 110/120), Mandarin (Native), Cantonese (Fluent)