Yuxuan Lou

Tel: +65-82600153

Email: yuxuanlou98@gmail.com; yuxuanlou@u.nus.edu

EDUCATION

National University of Singapore, Singapore (M.Sc. in Statistics)

- School of Statistics and Probability 2020.08 – Present

Fudan University, Shanghai, China (B.S. in Applied Mathematics)

- School of Data Science 2018.09 – 2020.07

  Core courses: C Programming (A), Introduction to Statistical Learning and Machine Learning (A), Data Structure (A), Computational Statistics (A-), Statistics: Principles, Methods and R (A-), Foundations of Probability Theory (A-), Introduction to Artificial Intelligence (B+)

- School of Mathematical Science 2016.09 – 2018.07

PUBLICATION

- ***Yuxuan Lou***, *Fuzhao Xue*, *Zangwei Zheng, Yang You, 2021.* Sparse-MLP: A Fully-MLP Architecture with Conditional Computation *arXiv preprint* ***arXiv:2109.02008*** *(AAAI 2022 In Submission)*

- *Fuzhao Xue*, *Ziji Shi*, ***Yuxuan Lou***, *Yong Liu*, *Yang You*, *2021.* Go Wider Instead of Deeper *arXiv preprint* ***arXiv:2107.11817*** *(AAAI 2022 In Submission)*

RESEARCH EXPERIENCES

Neural Network Model Scaling with Mixture of Experts

HPC LAB, National University of Singapore, Advisor: Prof. Yang You 2021.03 – Present

- Reviewed and reproduced modern Vision Transformer models and MLP-like models.
- Designed large-scale models (Sparse-MLP, Widenet) based on Mixture of Experts.
- Proposed a fully-MLP architecture with conditional computation in two directions and extended MoE to spatial dimension of image representation.
- Introduced parameter sharing to ViT-MoE models and proposed an explanation of why specific LayerNorm parameters had better performance.
- Distributed model training on TPU clusters.
- Detailed ablation study to further investigate the contribution of different model components.
- 2 papers submitted to AAAI 2022

Neural Network based Image Compression and Image Query System

DAS LAB, Harvard University, Advisor: Prof. Stratos Idreos 2019.07-2020.01

- Constructed the neural-network based image compression models which include Auto-Encoder, adaptive arithmetic coding, and adaptive code length regularization.
- Built neural network model based on Pyramid Convolutional Network and Generative Adversarial Network for different query tasks according to compressed image representation.
- Introduced spp-net and inverse spp-net, which is designed to better understand and summarize the multiscale

knowledge of images.

- Tested different model training methods to adjust the parameters of the model and improve model capacity.
- The new compressed image representation of our model is 4 times smaller than that of the baseline model without loss of digit capacity.

Score System of Figure Skating Sports Base on LSTM

CV LAB, School of Data Science, Fudan University Advisor: Prof. Yanwei Fu 2018.05 – 2019.01

- Reviewed video analysis methods including SVR, CNN, 3D convolution, and LSTM.
- Constructed the dataset by searching and downloading figure skating videos, including NHK, TEB, COC, 4CC, etc., and filtered the dataset by removing the videos that are not fluent or coherent.
- Assisted to propose a deep architecture that includes two complementary components, Self-Attentive LSTM and Multi-scale Convolutional Skip LSTM.
- Compared different pooling and regression methods (Max vs. Avg pooling, RBF vs. Linear SVR, SENetvs. C3D and TES vs. PCS).
- Results validated the effectiveness of proposed architecture.

Design of Toolkit (fastNLP) for Natural Language Processing

School of Data Science, Fudan University Advisor: Prof. Xipeng Qiu 2018.09 – 2018.12

- Learned to establish dataset SQuAD, a dataset of questions for machine comprehension of text, and analyze dataset based on sliding window baseline and logic regression.
- Reviewed pre-training language models and methods including ELMO, OpenAIGPT, etc
- Implemented a language representation model Bidirectional Encoder Representations from Transformers (BERT).
- Analyzed the model with the tasks of masked LM, next sentence prediction on SQuAD, GLUE, etc.
- Participated in designing FastNLP, a modularized and extensible toolkit for Natural Language Processing, to achieve a better performance.

INTERNSHIP

Interactive Entertainment Group, Tencent

Machine Learning Engineer Intern 2020.3 - 2020.6

- Data mining and data cleaning of users' comments on specific games.
- Built machine learning models to classify emotional levels of comments.
- Built deep learning abstractive text summarization models to extract the summary of comment contexts

SKILLS

Programming language & Database: C++, Python, Pascal, R, SQL, Spark

Deep learning: Tensorflow, Pytorch, Keras