Mengmeng Kuang

Email: kuangmeng@msn.com | Phone: (852) 9574 1537

GitHub: https://github.com/kuangmeng | Home Page: http://about.meng.uno

Education

Sept. 2020 **M.Phil** Degree in Computer Science, **The University of Hong Kong**
Supervisor: Prof. Hing-fung Ting
Thesis: Data-centric Approaches for better Multiple Sequence Alignment
Research Interests: Machine Learning, Deep Learning, Bioinformatics

Jun. 2018 **B.Eng** Degree in Computer Science, **Harbin Institute of Technology**
Thesis Advisor: Prof. Tiejun Zhao
Thesis: Cross-domain High Precision Chinese Word Segmentation
GPA: 89.6/100

Experience

Research

[J1] **Mengmeng Kuang**, Yong Zhang, Tak-wah Lam and Hing-fung Ting. "MLProbs: A Data-centric Approach for better Multiple Sequence Alignment". *IEEE/ACM Transactions on Computational Biology and Bioinformatics*. **The reviewers rated the paper as "Good", and we have submitted the revision and are waiting for the final decision.**

[C1] **Mengmeng Kuang** and Hing-fung Ting. "A data-centric pipeline using convolutional neural network to select better multiple sequence alignment method". *In Proceedings of the 11th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics (BCB '20)*.

[C2] **Mengmeng Kuang**, Yong Liu and Lufei Gao. "DLPAlign: A Deep Learning based Progressive Alignment Method for Multiple Protein Sequences". *In Proceedings of the Eleventh International Conference on Computational Systems-Biology and Bioinformatics (CSBio 2020)*.

Internship

Oct. 2020 - Cur | *Research Intern*, SmartMore Tech., Hong Kong
Engaged in generating handwriting text/image samples by conditional GANs and Auto-encoder models for an OCR task.

Jul. 2018 - Aug. 2018 | *Research Assistant*, Machine Intelligence and Translation Lab
Responsible for integrating a medical information processing platform, including medical relationship extraction, disease knowledge map, medical recommendation, medical document translation, and other functions.

Jan. 2017 - Apr. 2017 | *Algorithm Engineer Intern*, Richer Tech., Harbin
Participated in designing and implementing the target tracking system based on the CAMShift algorithm and the OpenCV framework for a climbing machine.

Project

Deep Learning based Progressive Alignment for Multiple Sequences | Oct. 2019 - Jun. 2020

To improve the quality of multiple sequence alignment on the "low similarity" protein family, the most influential part was selected from more than a dozen previous progressive methods of multiple sequence alignment. The CNN and bidirectional LSTM were used to establish a classification decision model to select the best calculation method for this part. Finally, a new progressive multiple sequence alignment tool was published and the average accuracy can be improved by 2.8%.

A Data-centric Pipeline for better Multiple Sequence Alignment | Sept. 2018 - Oct. 2019

To improve the quality of multiple protein alignment, random forest classifiers were used to select the most appropriate strategy for different families to obtain the temporary results. In order to further improve the quality of the temporary multiple sequence alignment, different similarity regions were extracted from the results, and random forest classifiers were used for selecting better strategies from "Realign Reliable Regions" and "Realign Unreliable Regions". Finally, a new pipeline, which could improve the accuracy of 1.65% compared with others in the test of 1356 protein families, was released.

Cross-domain High Precision Chinese Word Segmentation | Nov. 2017 - Jun. 2018

To improve the accuracy of Chinese Word Segmentation, a system based on conditional random field and Viterbi algorithm was developed with Java training from the artificial word segmentation results of The People's Daily. In order to further improve the adaptability in specific fields (medicine, law and finance), heuristic rules and specific guidelines were added. Finally, this word segmentation system could get an average accuracy of 97% in these specific fields.

Teaching

Jan. 2020 - Jun. 2020	COMP1117 Computer programming
Feb. 2019 - Jun. 2019	COMP7606 Deep learning

Scholarships and Certificates

Postgraduate

Nov. 2018	Certificate of Teaching and Learning in Higher Education
Sept. 2018	Postgraduate Scholarship

Undergraduate

Jun. 2018	Enterprise Scholarships
Nov. 2016	National Encouragement scholarship
Dec. 2015	Merit Student
Nov. 2015	National Encouragement scholarship

Languages

Chinese	Mother tongue
English	Fluent

Skills

Basic	TensorFlow, Programming (C/C++, Java, Python etc.)
	Office Softwares
Intermediate	RESTful Web-service, Keras, Scikit-Learn
Advanced	Modeling of Machine Learning and Deep Learning
	Multiple Sequence Alignment