Linqing Liu

http://likicode.com

Room 3301, Davis Center
University of Waterloo, ON, Canada
+1 (519) 580-6078 | likicode@gmail.com

EDUCATION

University of Waterloo **Waterloo, Canada**
M.Sc. in *Computer Science (thesis)* *Sept.2018 - May.2020 (Expected)*

- Supervisor: Prof. Jimmy Lin
- Related Courses: *Theory of Deep Learning, Data-Intensive Distributed Computing, Optimization for Data Science*

Tongji University **Shanghai, China**
B.Eng. in *Software Engineering* *Sept.2013 - June.2017*

- Related Courses: *Data Mining & Analysis, Data Warehouse Technology, Object-Oriented Programming, Data Structures, Operating System, Compiler Principle, Computer Architecture, Software Engineering*

PUBLICATIONS

1. **Linqing Liu**, Wei Yang and Jimmy Lin, *"Incremental Improvements to the Pairwise Word Interaction Model for Capturing Sentence Similarity"* (**NAACL**, under review), 2019
2. **Linqing Liu**, Yao Lu, Min Yang, Qiang Qu, and Jia Zhu, *"Generative Adversarial Network for Abstractive Text Summarization"*, The 30th AAAI Conference on Artificial Intelligence (**AAAI**, abstract), 2018
3. Yao Lu, **Linqing Liu**, Zhile Jiang, Min Yang and Randy Goebel, *"A Multitask Learning Framework for Abstractive Text Summarization"*, The 31th AAAI Conference on Artificial Intelligence (**AAAI**, abstract), 2019
4. **Linqing Liu**, Yao Lu, Ye Luo, Renxian Zhang, Jianwei Lu and Laurent Itti, *"Detecting Smart Spammers On Social Network: A Topic Model Approach."* (**NAACL-HLT**, Student Session), 2016

HONOURS & AWARDS

18' **David R. Cheriton Graduate Scholarship**, University of Waterloo, based on academic excellence
18' **AAAI Student Scholarship**, AAAI
16' **Google Anita Borg Memorial Scholarship: Asia-Pacific** , Google Inc (only 9 recipients in China)
16' **Top 3 of Beauty of Programming Competition 2016**, Microsoft Inc (Top 0.3%)
15' **First Prize in National Mathematics Competition**, Ministry of Education (Top 5%)

RESEARCH EXPERIENCES

Improvements to the Pairwise Word Interaction Model for Sentence Similarity **Waterloo**
Supervised by Prof. Jimmy Lin *Aug.2018 - Dec.2018*

- The highly-influential paper of Lan and Xu (2018) examined many neural network models for modeling sentence similarity and concluded that the ESIM and PWIM models are the best. We take the next logical step in trying to combine aspects of these two models to achieve the best of both worlds.
- We show that TreeLSTM consistently improves the performance of PWIM on all datasets. Visualization of a sample sentence pair also suggests that the TreeLSTM helps to leverage the syntactical matching signal of pairwise word interactions. We conduct thorough evaluations across eight datasets and establish new state-of-the-art results on the SICK dataset.

Generative Adversarial Network for Abstractive Text Summarization **Chinese Academy of Sciences**
Supervised by Prof. Min Yang, CAS *July.2017 - Dec.2017*

- We propose an adversarial process for abstractive text summarization, in which we simultaneously train a generative model G (as an agent of reinforcement learning, which takes the raw text as input and predicts the abstractive summarization) and a discriminative model D which attempts to distinguish the generated summary from the ground truth summary.

- Extensive experiments on the CNN/DailyMail Dataset shows that our model is able to generate more abstractive, readable and diverse summaries.

A Multitask Learning Framework for Abstractive Text Summarization **Chinese Academy of Sciences**
Supervised by Prof. Min Yang, CAS and Prof. Randy Goebel, University of Alberta *Oct.2017 - Feb.2018*

- We propose a Multi-task learning approach for Abstractive Text Summarization (MATS). Specifically, MATS consists of three key components: (i) a text categorization model that learns rich categoryspecific text representations using a LSTM encoder; (ii) a syntax labeling model that learns to improve the syntax-aware LSTM decoder; and (iii) an abstractive text summarization model that shares its LSTM encoder and decoder with the text categorization task and the syntax labeling task, respectively.
- The abstractive text summarization model enjoys significant benefit from the additional text categorization and syntax knowledge.

Topic Model Based Microblog Spammer Detection **iLab@Tongji**
Collaborate with Prof. Ye Luo, Tongji University *Oct.2015 - April.2016*

- The project aimed at detecting smart spammers, whose posting behaviors resemble that of legitimate users. The detection method is proposed based on their different topic distribution patterns.
- We extracted topic-based features (GOSS and LOSS) for spammer detection, which outperform state-of-the-art methods. We also built the first public available dataset of human-like fake accounts on Chinese Weibo (microblog) platform.

PROJECTS

Text-to-Image Information Retrieval from Large-Scale Web Archives **JCDL demo**
Nov.2018 - Jan.2019

- Analyzing large-scale web archives is a challenging task. We take advantage of the open-source "big data" infrastructure to support scholarly exploration of image information retrieval over the web archives.
- We utilize pre-trained object detection models (based on tensorflow/keras) and then parallelize it to different machines with Spark. The speed analysis shows that we can get the nearly linear speed up in the distributed environment. Each image is stored with its containing objects. We then build the retrieval model on top of the indexed data to get results with high accuracy.

Open Source Contribution

Differential network local consistency (DNLC) **R Package**
Authors: Yao Lu, Yusheng Ding, Linqing Liu and Tianwei Yu *Dec.2016*

- Using Local Moran's I to detect differential network local consistency
- CRAN: https://cran.r-project.org/web/packages/DNLC/index.html

TECHNICAL STRENGTH

Programming Languages	Python, R, C/C++, Matlab, LaTeX, HTML/CSS, Javascript
Platforms	Linux, Web, OS X(Unix)
Version Control	Git, SVN
Packages&Frameworks	Torch, Tensorflow, Scala, MapReduce, Spark