Zhaohui LI

W363 Westgate Building, University Park, PA 16801

Cell phone: (+1) 814-825-9537 Email: nkzhaohuilee@gmail.com Homepage: https://zhaohuilee.com

EDUCATION

School of Computer and Control Engineering, Nankai University, China Aug 2012-Jul 2016

Bachelor of Computer Science.

Department of Computer Science and Engineering, Pennsylvania State University Aug 2019-Present

The Natural Language Processing Lab

Ph.D. Student of *Prof. Rebecca J. Passonneau*

EMPLOYMENT

Institute of Computing Technology, Chinese Academy of Science, Beijing, China Jul 2016-Feb 2019

CAS Key Laboratory of Network Data Science and Technology

Research Assistant of *Prof. Jun Xu*

Big Data Lab, Baidu Research, Baidu Inc, Beijing, China Feb 2019-Aug 2019

Internship of advisor *Prof. Jun Huan*

RESEARCH INTEREST

My research interest resides at ***Nature Language Processing***. And I am now focusing on *Machine Reading Comprehension* and *Student Reports Auto-evaluation* task, which are based on *relation reasoning and Deep Reinforcement Learning* models. Moreover, I hope my research results can be applied to *Education Applications* and can benefit more and more people all over the world.

Machine Reading Comprehension; Nature Language Processing; Education;

PUBLICATIONS

- **Zhaohui Li**, Yue Feng, Jun Xu, YanYan Lan, Jiafeng Guo, Yue Feng, Xueqi Cheng. Teaching Machines to Extract Main Content for Machine Reading Comprehension. The 33th AAAI Conference on Artificial Intelligence (AAAI'19)
- **Zhaohui Li**, Jun Xu, YanYan Lan, Jiafeng Guo, Yue Feng, Xueqi Cheng. Hierarchical Answer Selection Framework for Multi-passage Machine Reading Comprehension. The 24nd China Conference on Information Retrieval. Springer, Cham, 2018: 93-104. (CCIR'18 best paper candidate).
- Tianyou Guo, Jun Xu, Xiaohui Yan, Jianpeng Hou, Ping Li, **Zhaohui Li**, Jiafeng Guo, and Xueqi Cheng. Ease the Process of Machine Learning with Dataflow. Proceedings of the 25th ACM International Conference on Information and Knowledge Management (CIKM'16). Indianapolis, USA, pp. 2437-2440, 2016.
- Yuan Wang, **Zhaohui Li**, Jie Liu, Zhicheng He, Yalou Huang, Dong Li. Word Vector Modeling for Sentiment Analysis of Product Reviews. The conference on Natural Language Processing and Chinese Computing (NLPCC'14). Springer, Berlin, Heidelberg, 2014: 168-180.

RESEARCH EXPERIENCE

QUESTION ANSWERING

Extract Key Information for Machine Reading Comprehension | ICT | Research Assistant Dec 2017-Oct 2018

Advisor: Prof. *Jun Xu*, Department of Computer Science and Technology, Renmin University of China.

- Surveyed Machine Reading Comprehension (MRC) methods and Deep Reinforcement Learning models in NLP.
- Developed special attention mechanism to get better questions and passages representations.
- Proposed a model that could identify the main content from passages through Markov Decision Process. Implemented experiment on MS-MARCO dataset and archived high performance.
- Paper: "Teach Machines to Learn Main Content for Machine Reading Comprehension" was accepted by AAAI'19. (1st author)

Multi-Passage Machine Reading Comprehension Pipeline| ICT | Research Assistant Oct 2017-May 2018

Advisor: Prof. *Jun Xu*, Department of Computer Science and Technology, Renmin University of China.

- Surveyed Question Answering (QA) and Machine Reading Comprehension (MRC) methods.
- Implemented a pipeline (Select passage->Select sentence->Predict answer) for a Chinese MRC task DuReader.
- Paper: "Hierarchical answer selection framework for Multi-passage Machine Reading Comprehension" was accepted by CCIR'18 as Best paper candidate. (1st author)

SENTIMENT ANALYSIS

Word Vector Modeling for Sentiment Analysis | Nankai University | Research Assistant Jan 2014-Dec 2014

Advisor: Prof. *Jie Liu*, School of Computer Science, Nankai University

- Surveyed Word Vector Modeling method and analysis statistical attributes over Amazon Product Review Dataset.
- Improved Word2Vec by adding positive/negative label which represents words' sentiment information on product reviews.
- Acquired the best F1 score in the Competition of NLPCC2014.
- Paper: Word Vector Modeling for Sentiment Analysis of Product Reviews was accepted by NLPCC'2014. (2nd author)

MACHINE LEARNING SYSTEM

BDA: big data analysis as a service (***1900 Stars on GitHub***) Aug 2015-Dec 2017

Advisor: Prof. *Jun Xu*, Department of Computer Science and Technology, Renmin University of China

- Responsible for the survey of Big Data Analysis System.
- Implemented an efficient, high scalability data analysis system and used Oozie to create a pipeline between HDFS and Spark.
- Core member of designing BDA Studio, which is a HCI to enable user to modify model on website visual interface.
- Developed BDA to an open source version system EasyML and created corresponding Docker images.
- Paper: Ease the Process of Machine Learning with Dataflow was accepted by CIKM'2016 as best Demo Paper Candidate. (1st undergraduate author)

Global Information Hunter: A news data capture, analysis and visualization system, which consists of Natural Language Understanding Engine, Knowledge Graph Engine, API Layer and Web App Layer Jan 2018-Jan 2019

Advisor: Prof. *Jiafeng Guo*, professor at CAS Key Laboratory of Network Data Science and Technology, ICT

- Team leader of Knowledge Graph Engine and API Layer (5 student members), one of the mentors of NLP Engine (2 mentors and 7 student members). And responsible for code review works.
- Designed the structures and methods of 3 parts in NLP Engine, which are Sequence Tagging, Document Classification and Sentiment Analysis.
- Established an Entity Knowledge Graph Schema and an Event Graph Schema, based on which Graph computing algorithms and API functions are implemented.

AWARDS AND HONORS

AAAI 2019 Travel Award and volunteer Jan 2019

"GongNeng" Scholarship of Nankai University Dec 2013

Interdisciplinary Contest in Modeling, "H" Award| Team member Feb 2015

Captain of College Soccer Team (Silver Medal) | Nankai University | Team Leader May 2013-Jul 2015

SKILLS

Expert in ***Reinforcement Learning***

Familiar with ***Pytorch*** and *TensorFlow*

Programming Languages: ***Python***