Chao Zhao

Room 386, Engineering 2, 1156 High St, Santa Cruz, CA, 95064

zhaochaocs@gmail.com https://zhaochaocs.github.io/

Education

University of North Carolina at Chapel Hill — Jan 2020 -
Ph.D. student in Computer Science — Chapel Hill, NC, USA

- Natural language understanding and generation; knowledge-grounded models.

University of California, Santa Cruz — Sep 2018 - Dec 2019
Ph.D. student in Computer Science (withdrawal with a M.Sc degree) — Santa Cruz, CA, USA

- Natural language understanding and generation.

Harbin Institute of Technology — Sep 2016 - Jun 2018
M.Sc in Computer Technology, School of Computer Science and Technology — 2 years, Harbin, China

- Research fields: natural language processing, knowledge graph, health informatics

Harbin Institute of Technology — Sep 2012 - Jun 2016
B.E. in Flight Vehicle Design and Engineering, School of Astronautics — 4 years, Harbin, China

Selected Publications

- **Chao Zhao**, Snigdha Chaturvedi. "Weakly-Supervised Opinion Summarization By Leveraging External Information" *Thirty-Fourth AAAI Conference on Artificial Intelligence*. Feb 2020.
- **Chao Zhao**, Jingchi Jiang, Yi Guan, Xitong Guo, and Bin He. "EMR-based medical knowledge representation and inference via Markov random fields and distributed representation learning." *Artificial Intelligence in Medicine*, 87 (2018): 49-59.
- **Chao Zhao**, Jingchi Jiang, Zhiming Xu, and Yi Guan. "A study of EMR-based medical knowledge network and its applications." *Computer Methods and Programs in Biomedicine* 143 (2017): 13-23.
- **Chao Zhao**, Min Zhao, and Yi Guan. "Constructing a Hierarchical User Interest Structure based on User Profiles." *2017 IEEE 17th International Conference on Data Mining Workshops (ICDMW)*. pages 156−162, Nov 2017.
- Zhipeng Jiang*, **Chao Zhao***, Bin He, Yi Guan, and Jingchi Jiang. "De-identification of medical records using conditional random fields and long short-term memory networks." *Journal of Biomedical Informatics*, S75 (2017): S43-S53, **co-first author**
- Jingchi Jiang, Jing Xie, **Chao Zhao**, Jia su, Yi Guan, and Qiubin Yu. "Max-Margin Weight Learning for Medical Knowledge Network." *Computer Methods and Programs in Biomedicine* 156 (2018): 179-190.
- Jingchi Jiang, Jichuan Zheng, **Chao Zhao**, Jia Su, Yi Guan, and Qiubin Yu. "Clinical-decision support based on medical literature: A complex network approach." *Physica A: Statistical Mechanics and its Applications* 459 (2016): 42-54.
- Jingchi Jiang, Xueli Li, **Chao Zhao**, Yi Guan, and Qiubin Yu. "Learning and inference in knowledge-based probabilistic model for medical diagnosis." *Knowledge-Based Systems* 138 (2017): 58-68.

Professional Experience

Knowledge-enhanced BERT encoding — June 2019 - Sep 2019
Internship at Knowledge Graph Group, Baidu Inc. — 3 months

- Designed a whole-word tagging schema for natural language, where the tagging labels can provide corresponding linguistical knowledge and commonsense;
- Applied the tagging results as complementary features of BERT to boost its performance on reading comprehension (+0.2%), question answering (+0.4%), natural language inference (+1.4%), and similarity assessment (+0.6%).

Unsupervised opinion summarization for online product reviews Nov 2018 - May 2019
Accepted by AAAI-20 ◦ 6 months

- Developed a summarization method for online product reviews by leveraging external knowledge, which is easily obtained from the Internet without any human efforts;
- Achieved higher ROUGE score compared with the strong baselines with human supervision.

Medical knowledge representation and reasoning Mar 2016 - Jun 2018
Final year graduate thesis, published at AIM & CMPB ◦ 2 years

- Constructed a medical knowledge network from the real EMR data, which is then converted as a Markov network to support clinical decision making;
- Extracted triple knowledge automatically from medical texts and then integrated it as a large-scale medical knowledge graph. Designed strategies for knowledge description, storage, and validation.

Entity Linking based on corresponding descriptive sentences May 2017 - Aug 2017
Internship at Knowledge Graph Group, Baidu Inc. ◦ 3 months

- Determined the category of arbitrary entities according to their descriptive sentences with a CNN-based model, with specially designed modules to make the model robust to noisy and imbalanced data;
- Applied this system to 2.1 million entities which were hard to be linked. The method linked 1.1 million out of them with a precision of 99.4%.

Concept association from the perspective of user's interest Feb 2017 - Apr 2017
Internship at Knowledge Graph Group, Baidu Inc., published in DaMNet workshop@ICDM-17 ◦ 3 months

- Integrated the interests of two million users as a network and then explored its structure by community detection.
- Labeled each interest community with relevant concepts automatically, to depict the interest relevancy at a concept level.

Error detection and correction for short texts Nov 2016 - Dec 2016
With Zhongke Huilian Inc. ◦ 1 months

- Adopted n-gram language model with Kneser-Ney smoothing to detect and correct possible typos in texts.
- Corrected about 80% of the errors of test data, with only a small corpus (about 4M) to train the language model.

Removing of protected health information from psychiatric evaluation records Jun 2016 - Oct 2016
i2b2 2016 CEGS N-GRID De-identification Task, published at JBI ◦ 4 months

- Implemented a de-identification system using the character-level bi-LSTM with enhanced word embeddings.
- Attained an F_1 of 0.899, which is competitive with the best score among the 15 participating teams (0.914).

Teaching

- TA for CMPS12B: Introduction to Data Structures, UCSC Fall 2018

Skills

Programming: Python (pytorch, tensorflow, keras, theano) > Java > Matlab > C > C++
Others: Git, Linux, Hadoop, LaTeX, HTML, Photoshop

Selected Awards

- The Regents' Fellowship (University of California, Santa Cruz) Dec 2018
- Innovation Scholarship (Ministry of Industry and Information Technology, China) Dec 2017
- National Scholarship for Graduate Students (Ministry of Education, China) Nov 2017
- Outstanding Graduate Award (Harbin Institute of Technology) Jun 2016
- Outstanding Final Year Project Thesis (Harbin Institute of Technology) Jun 2016
- Top-grade Scholarship (Harbin Institute of Technology) Sep 2015
- First National Prize for China Undergraduate Mathematical Modeling Sep 2015
- National Scholarship for Undergraduate Students (Ministry of Education, China) Sep 2014