

CHAO ZHAO

📍 803 Zonghe Building, 92 West Dazhi St. , Harbin City, China, 150001

✉️ zhaochaocs@gmail.com

☎️ +86-187-0462-4509

🌐 <https://zhaochaocs.github.io/>

🎓 EDUCATION

Harbin Institute of Technology

📅 Sep 2016 - (Jun 2018)

M.Sc in Computer Technology, School of Computer Science and Technology

◦ 2 years, Harbin, China

- GPA: 83.0/100 Rank: 2/186
- Research fields: natural language processing, knowledge graph, health informatics, machine learning
- Core courses: Algorithms, Machine learning, Data mining, Natural language processing, Knowledge engineering, Artificial intelligence, Deep learning for NLP

Harbin Institute of Technology

📅 Sep 2012 - Jun 2016

B.E. in Flight Vehicle Design and Engineering, School of Astronautics

◦ 4 years, Harbin, China

- GPA: 87.8/100 3.7/4.0 Major GPA: 93.1/100 4.0/4.0 Rank: 5/85

📖 PUBLICATIONS

- **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)*. IEEE, 2017. (Accepted, Oral)
- 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, 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.
- Jingchi Jiang, Yi Guan, Jia Su, **Chao Zhao**, and Jinfeng Yang. “HIT-WI at TREC 2015 Clinical Decision Support Track.” In *TREC*. 2015.
- Jingchi Jiang, Yi Guan, and **Chao Zhao**. “WI-ENRE in CLEF eHealth Evaluation Lab 2015: Clinical Named Entity Recognition Based on CRF.” In *CLEF (Working Notes)*. 2015.

In Progress

- **Chao Zhao**, Min Zhao, and Yi Guan. “Classification of entities via their descriptive sentences.” arXiv:1711.10317, 2017.
- **Chao Zhao**, Jingchi Jiang, and Yi Guan. “EMR-based medical knowledge representation and inference via Markov random fields and distributed representation learning.” (*Artificial Intelligence in Medicine*, under review)
- 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*, under review)

🔧 PROFESSIONAL EXPERIENCE

Construction of Chinese medical knowledge base

📅 Sep 2017 - Now

Final year graduate project, Part II

◦ 3 months

- Acquiring knowledge automatically from medical texts, and integrating them as a medical knowledge base.
- Designing the descriptive schema of the knowledge base, as well as its corresponding storage, search and validation tools.

Entity classification based on its descriptive sentences

Internship at Knowledge Graph Group, Baidu Inc.

📅 May 2017 - Aug 2017

◦ 3 months

- Determined the category of arbitrary entities according to their descriptive sentences, with a text classification model.
- Designed a clustering module to filter the noise instances and alleviate the class imbalance problem.
- Applied this system to 2.1 million entities, and 1.1 million are successfully classified with a precision of 99.4%.

Concept association from user interest perspective

Internship at Knowledge Graph Group, Baidu Inc.

📅 Feb 2017 - Apr 2017

◦ 3 months

- Integrated the interests in user profiles as a network and then explored its structure by community detection.
- Labeled each interest community with relevant concepts, to depict the topics of interest at the concept level.

Knowledge representation and reasoning from Chinese electronic medical records

Final year graduate project, Part I

📅 Oct 2015 - Feb 2017

◦ 1.3 years

- Constructed a medical knowledge network containing the medical entities and entity relationships
- Regarded entities as random variables and modeled their relationships via Markov networks and Markov logic networks.
- Represented the medical entities as embeddings to depict their similarities, and then designed new energy functions for inference.

Error detection and correction of Chinese texts

With Zhongke Huilian Inc.

📅 Nov 2016 - Dec 2016

◦ 1 months

- Adopted n-gram language model with Kneser-Ney smoothing to detect and correct typos in Chinese 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

i2b2 2016 CEGS N-GRID De-identification Task, SUNY at Albany

📅 Jun 2016 - Oct 2016

◦ 4 months

- Implemented a de-identification system using the character-level bi-LSTM networks with enhanced word embeddings.
- Attained the F_1 measure of 0.899. The best score among the 15 participating teams was 0.914.

Construction of the knowledge graph on consumer-oriented consuming interest

With Ricoh Software Research Center Beijing Co.,Ltd.

📅 Mar 2016 - May 2016

◦ 3 months

- Analyzed the activities, similarities, importance, clustering, and purchase possibilities of the customers, according to the consumption behavior of each customer in e-commerce sites.

Biomedical literature retrieving for clinical question answering

2015 TREC Clinical Decision Support Track

📅 Jun 2015 - Aug 2015

◦ 2 months

- Retrieved the relevant medical literature of the given clinical records from 700, 000 articles via Lucene and Indri.
- Constructed a medical literature network and then identified the potentially relevant ones from the literature pool.

⚙️ SKILLS

Language: Chinese, English

Programming: Python (Tensorflow, theano), Java, C, C++, Matlab, Shell

Others: Git, Hadoop, \LaTeX , HTML, Photoshop

🏆 SELECTED AWARDS

- Innovation Scholarship (Top 1% in China) Dec 2017
- National Scholarship for Graduate Students (Top 1% in China) Nov 2017
- Outstanding Graduate Award (Top 10% of all graduates in HIT) Jun 2016
- Outstanding Final Year Project Thesis (Top 10% of all graduates in HIT) Jun 2016
- Top-grade Scholarship (Top 1% of all students in HIT) Sep 2015
- First National Prize for China Undergraduate Mathematical Modeling (Top 1.5% in China) Sep 2015
- National Scholarship for Undergraduate Students (Top 1% in China) Sep 2014