

ZEXUE HE

Beijing Normal University (BNU) ◇ <https://zexuehe.github.io/> ◇ zexueh@mail.bnu.edu.cn

RESEARCH INTERESTS

Machine Learning, Information Retrieval, Natural Language Processing, Data Mining, and Bioinformatics

EDUCATION

College of Information Science and Technology, BNU, Beijing, China Sept. 2016 - Jun. 2019 Expt.
B.S. in Computer Science and Technology, last two years GPA: 3.92/4.00, major GPA: 3.96/4.00

- First-class Scholarship for Academic Excellence and Competition Excellence of BNU
- Straight As: Deep Learning, Pattern Recognition, Computer Graphics, Visualization and 4 programming courses
- Silver medal in 2016 International Collegiate Programming Contest at Beijing regional site (ACM/ICPC, Beijing)
- Best Female Team in 2016 China Collegiate Programming Final Contest (CCPC Final)

School of Life Sciences, BNU, Beijing, China Sept. 2015 - Jun. 2016
Major in Biology Science, overall GPA: 3.61/4.00, major GPA: 3.80/4.00

- Second-class Scholarship for Academic Excellence and Competition Excellence of BNU
- Gold medal in 2016 International Genetically Engineered Machine Competition (iGEM) at U.S., demo [here](#)

PUBLICATIONS

Learning Robust Representations by Projecting Superficial Statistics Out

Haohan Wang, **Zexue He**, Zachary C. Lipton, Eric P. Xing

Accepted to International Conference on Learning Representations (ICLR) 2019, oral presentation

[\[PDF\]](#)

Rapid and High-quality 3D Fusion for Human Brain's CT-MRI Heterogeneous Data

Zexue He, Minjie Li, Jinyao Li, Yiran Chen, Yanlin Luo

Accepted to International Conference on Virtual Reality and Visualization (ICVRV) 2018

[\[PDF\]](#)

Recommended to SCI Journal: SCIENCE CHINA Information Sciences

Presentation at ICVRV 2018, Qingdao, Shandong, China

Leveraging Gloss Knowledge in Neural Word Sense Disambiguation by Hierarchical Co-Attention

Fuli Luo, Tianyu Liu, **Zexue He**, Qiaolin Xia, Zhifang Sui, Baobao Chang

Accepted to Conference on Empirical Methods in Natural Language Processing (EMNLP) 2018

[\[PDF\]](#)

A Large-Scale Study of Mobile Search Examination Behavior

Xiaochuan Wang, Ning Su, **Zexue He**, Yiqun Liu, Shaoping Ma

Accepted to ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR) 2018 [\[PDF\]](#)

Presentation at SIGIR 2018, Ann Arbor, Michigan, U.S.

A Two-Stage Model for User's Examination Behavior in Mobile Search

Jiaxin Mao, Yiqun Liu, Noriko Kando, **Zexue He**, Min Zhang, Shaoping Ma

Accepted to ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR) 2018

[\[PDF\]](#)

Understanding Reading Attention Distribution during Relevance Judgment

Xiangsheng Li, Yiqun Liu, Jiaxin Mao, **Zexue He**, Min Zhang, Shaoping Ma

Accepted to ACM International Conference on Information and Knowledge Management (CIKM) 2018

[\[PDF\]](#)

OVERSEA RESEARCH COOPERATION

Carnegie Mellon University | Research Intern

Apr. 2018 - Oct. 2018

Robust Learning for Domain Generalization (DG) without Domain Information

Advisors: Prof. Zachary C. Lipton & Prof. Eric P. Xing, Machine Learning Department

Mentor: Haohan Wang, Ph.D. student, Language Technologies Institute

- Designed an algorithm to learn robust representations of cross-domain data under distribution shift without domain identifiers by excluding distribution-specific (superficial) statistics (e.g., texture) with orthogonal projection.

- Designed a new differentiable neural building block of traditional [GLCM](#) to extract superficial representations.
- Got comparable results with other DG methods that need domain knowledge on synthetic and standard datasets.

WORK EXPERIENCE

Google Inc. | Engineering Practicum Intern

Jul. 2017 - Sept. 2017

Wikipedia-like Sites Discovery and Analysis

Advisor: Jiang Bian, team manager of Dataz Group

- Built a regression model to identify high-quality websites which provide entities and relations for Google [Knowledge Graph](#) after analyzing and extracting features of Wikipedia. Finally predicted 5,793 qualified websites.
- Designed a parallel computing algorithm with [MapReduce](#) to count one-topic-page ratio of each website, inspired by [tf-idf](#), with NLP knowledge and techniques such as word segmentation and named entity recognition.
- Built a [MapReduce](#) pipeline to count internal link page-density of websites with HTML [DOM Tree](#) and [BigTable](#).
- Selected to attend [Grace Hopper Celebration Conference](#) 2017 in Orlando, FL, U.S. with Intern Scholarship.

RESEARCH EXPERIENCE

Tsinghua University | Research Assistant

Jun. 2017 - May 2018

Investigating Human Examination Behavior in Mobile Search

Advisor: Prof. Yiqun Liu, Computer Science and Technology Department

- Investigated users' interactions with search engine results page of 4 commercial mobile search engines - [Baidu](#), [Sogou](#), [Shenma](#), and [Haosou](#), after extracting features from interaction data collected in a lab-based user study.
- Built a two-stage model for examination behavior via [EM algorithm](#), [logistic regression](#), and published [paper](#).
- [Published](#) the analysis of user's examination behavior on mobile with large-scale click and viewport log of [Sogou](#).

Peking University | Research Assistant

Dec. 2017 - May 2018

Chinese Word Segmentation (CWS) with Character Glyph Embedding

Advisor: Prof. Baobao Chang, Key Laboratory of Computational Linguistics, Ministry of Education (MOE)

- Improved CWS by introducing the rich semantic information hidden in character glyph into character embedding.
- Designed an end-to-end neural network with an [autoencoder](#) to generate glyph embedding, BiLSTM, and [CRF](#).
- Got great CWS results especially the out-of-vocabulary rate on both simplified and traditional Chinese datasets.

Beijing Normal University | Research Group Leader

Jul. 2017 - Feb. 2018

Human Brain's CT-MRI Heterogeneous Data Fusion and Visualization

Advisor: Prof. Yanlin Luo, Engineering Research Center of Virtual Reality and Applications, MOE

- Proposed a heterogeneous data fusion algorithm for [CT](#) and [MRI](#) diagnoses by adding [MRI](#) to the center of [CT](#).
- Whole and layered visualized via designed trapezoid transfer function and CUDA ray-casting volume rendering.
- Implemented a new version of our visualization platform to conduct experiments and visualization. Results [here](#).

Beijing Normal University | Research Assistant

Nov. 2015 - Sept. 2016

Genetic Parallel Computing Model

Advisor: Prof. Xudong Zhu, School of Life Sciences

- Applied a biology parallel computing system to [NP-hard](#) problems using the information storage ability of DNA.
- Simulated [Hamiltonian Cycle Path](#) and [Vertex Coloring](#) problems with several [fluorescent proteins](#) and [PCR](#).

SELECTED PROJECTS

VR: 3D Virtual World System | Team Leader

Sept. 2016 - Mar. 2017

- Developed a 3D Virtual World System with [OpenGL](#), [OpenSim](#), [3D Studio Max](#), and [Linden Scripting Language](#).
- Made a movie with it about [Dream of the Red Chamber](#), one of [China's Four Great Classical Novels](#), demo [here](#).

Android Application: PDFCanvas | Team Member

Mar. 2016 - Mar. 2017

- Developed an [education App](#) which can share notes in class in real time, arrange slides, broadcast, and display.
- Designed an algorithm with stacks to store real-time data by different layers when taking notes and broadcasting.

OTHERS

Programming Languages: C++, C, Python, Java, Perl, SQL, R, HTML, Pascal, \LaTeX , MATLAB

Activities: a member of International Student Affairs at BNU Student Union, a volunteer in Google Family Day