ZEXUE HE

Beijing Normal University (BNU) ◇ +86 18568235369 ◇ zexueh@mail.bnu.edu.cn

RESEARCH INTERESTS

Information Retrieval, Deep Learning, and Natural Language Processing

EDUCATION

College of Information Science and Technology, BNU, Beijing, China Mar. 2016 - Jun. 2019 Expt.
Bachelor of Science in Computer Science and Technology, overall GPA: 3.93/4.00, major GPA: 3.96/4.00
- The First-class Scholarship for Academic Excellence and Competition Excellence of BNU
- Straight A in 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 - Sept. 2016
Major in Biology Science, overall GPA: 3.61/4.00, major GPA: 3.80/4.00
- The Second-class Scholarship for Academic Excellence and Competition Excellence of BNU
- Gold medal in 2016 International Genetically Engineered Machine Competition (iGEM) at Boston, demo here

PUBLICATIONS

Learning Robust Representations by Projecting Superficial Statistics Out
Haohan Wang, **Zexue He**, Eric P. Xing
Submitted to International Conference on Learning Representations (ICLR) 2019, reviews 9, 7, 7 of 10 [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 of Virtual Reality and Visualization (ICVRV) 2018 [PDF]
Recommended to SCI Journal: CHINA SCIENCE Information Science
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 Examination Behavior in Mobile Web Search Based on Real Search Log
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 Abor, Michigan, U.S.

A Two-Stage Model for Users 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 Assistant 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 in Language Technologies Institute
- Designed a novel learning algorithm to learn robust representations of inputs by projecting superficial statistics (e.g., texture) out under distribution shift, which always cheat or even break down the well-known classifiers.

- Designed an easy-plugged neural block of traditional GLCM that can be optimized via back gradient propagation.
- Got comparable performance with other DG methods with 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 distinguish high-quality websites to provide entities 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 techniques like 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 on Mobile Search
Advisor: Prof. Yiqun Liu, Computer Science and Technology Department
- Investigated users' interactions with Search Engine Results Page (SERP) of 4 commercial mobile search engines (SE) - *Baidu, Sogou, Shenma and Haosou*, after extracting features of collected interaction data in lab experiments.
- 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 — Jan. 2018 - 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 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.

SELECTED PROJECTS

Simplified Search Engine (SE) for BNU School of Government Website — Oct. 2017 - Dec. 2017
- Implemented a simplified keyword search engine after crawler and text operations like word segmentation and stop word removal. Indexed documents with inverted index and ranked results based on tf-idf.

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.

SKILLS

Programming Languages: C++, C, Python, Java, Perl, SQL, HTML, Pascal, LaTeX, MATLAB
Toolkits/Software: TensorFlow, Keras, PyTorch, Tableau, Gephi