Xiusi Chen

3551 Boelter Hall, UCLA, Los Angeles, California 90066
(+1) (310)-913-7838 • ✉ xchen@cs.ucla.edu

RESEARCH INTERESTS

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

EDUCATION

Peking University 2018 - 2023 (expected)
Ph.D. in Computer Science

- **Fields:** Data Mining (especially Graph Mining and Text Mining), Database, Machine Learning
- **Advisor:** Prof. Wei Wang

Peking University 2015 - 2018
M.S. in Computer Science **Rank: Top 2/54**

- **Fields:** Data Mining(especially Graph Mining), Database, Machine Learning, Recommender Systems
- **Advisor:** Prof. Jun Gao

Peking University 2011 - 2015
B.S. in Computer Science

- **Thesis:** The Design and Implementation of Indexes on Distributed Big Graph Processing Framework
- **Advisor:** Prof. Jun Gao

PUBLICATIONS

Fangbo Tao*, Chao Zhang*, **Xiusi Chen**, Meng Jiang, Tim Hanratty, Lance Kaplan and Jiawei Han, *Doc2Cube: Automated Document Allocation to Text Cube via Dimension-Aware Joint Embedding*. In Proceedings of the 17th IEEE International Conference on Data Mining (**ICDM**), 2018. Oral.

Chao Zhang, Fangbo Tao, **Xiusi Chen**, Jiaming Shen, Meng Jiang, Brian Sadler, Michelle Vanni and Jiawei Han, *TaxonGen: Constructing Topical Concept Taxonomy by Adaptive Term Embedding and Clustering*. in Proceedings of 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD**), 2018. Oral.

Chang Zhou, Jinze Bai, Junshuai Song, Xiaofei Liu, Zhengchao Zhao, **Xiusi Chen**, and Jun Gao, *ATRank: An Attention-Based User Behavior Modeling Framework for Recommendation*. To appear on the 32nd AAAI Conference on Artificial Intelligence(**AAAI**), 2018. Oral.

Xiusi Chen, Qimu Zheng, Xinxi Jiang and Jun Gao, *The Design and Implementation of Index on Distributed Big Graph Systems* In Proceedings of the 34th National Database Conference(**NDBC**), 2017. Oral.

Xiusi Chen, Xiaoyu Li, and Jun Gao, *Demo of a Semi-Supervised Method Based System to Detect Advertisements from QA pairs* In Proceedings of the 33rd National Database Conference(**NDBC**), 2016. Oral.

PREPRINTS

Xiusi Chen, Xiaoyu Li, Chang Zhou, Xiaofei Liu and Jun Gao, *Learning Item Embedding with Heterogeneous Information for Recommendation*. under review

Xiusi Chen*, Xiaoyu Li*, Chang Zhou and Jun Gao, *Scalable Graph Embedding Enhanced by Content-Preserving Locality Sensitive Hashing*. under review

RESEARCH EXPERIENCE

Topic taxonomy construction of massive scientific literature Jun. 2017 - Present
Advisor: Prof. ChengXiang Zhai *Department of Computer Science,* ***University of Illinois at Urbana-Champaign***

- Proposed to use multi-task learning to make the produced word embedding better capture the hypernym-hyponym relation between word pairs.
- Implemented the multi-task model and evaluate the quality of the embeddings both qualitatively and quantitatively.

- Exploring pattern-based and embedding-based stacking method to guarantee a high F-score on the classification of hypernymy relation.
- A 10% boost of relation accuracy and term coherency is perceived.
- **Paper submitted to VLDB'18.**

Merchandise Recommendation and Serving on Taobao.com Sept. 2016 - Jun. 2017
Mentor: Dr. Chang Zhou *Recommendation and Serving Group, Cloud Retail Services,* ***Alibaba Group***

- Proposed a Locality-Sensitive Hashing based approach to enhance graph embedding with text to improve similarity measuring of long-tail items under a collaborative filtering setting.
- Built a deep neural network based recommendation system.
- The system has been on production since 2016.12 and serves more than 200million users on over 300 million products per day.
- A 10% boost of CTR(click through rate) is perceived after the deployment.
- **Paper submitted to IJCAI'18 and ICDE'18.**

Semi-Supervised Advertisement Discrimination System Mar. 2016 - Jun. 2016
Advisor: Prof. Jun Gao *Institute of Network Computing and Information Systems,* ***Peking University***

- Developed semi-supervised(only positive examples are given) iterative algorithms to discriminate advertisements from normal answers in QA communities.
- Conventional machine learning methods like Logistic Regression, Random Forest, Support Vector Machine, and ensemble methods are involved.
- **Invited demo paper for NDBC 2016.**

The Design and Implementation of Indexes on Distributed Big Graph Processing Framework Mar. 2015 - May 2015
Advisor: Prof. Jun Gao *Institute of Network Computing and Information Systems,* ***Peking University***

- Investigated different indexes suiting the distributed graph processing systems.
- Implemented different types of indexes on Giraph.
- The modified Giraph support both synchronous and asynchronous execution modes.
- The proposed index mechanism brought an 18 to 70 times query execution boost.
- **Paper accepted by NDBC 2017.**

HONORS AND AWARDS

Leo KoGuan Fellowship Sept. 2016
top 2%, Peking University

Merit Student Sept. 2016
top 5%, Peking University

Graduate Student Academic Fellowship Sept. 2015, 2016, 2017
Peking University

2nd Place of National Database Conference (NDBC) Cup 2016 Jun. 2016
Out of 566 teams, China Computer Federation

Kaggle State Farm Distracted Driver Detection Jul. 2016
Ranked top 13%(191/1440)

3rd Prize of 15th Peking University ACM Programming Contest May. 2016
Peking University

3rd Place of 17th Chinese University Basketball Association (CUBA) Championships, Beijing Division Oct. 2014
On behalf of Peking University, Federation of University Sports of China

6th Place of the Middle School Students Swimming Contest of Hunan Province, 100M Breaststroke Sept. 2010
The Education Department of Hunan Province

TEACHING EXPERIENCE

Introduction to Database Systems (Honor Track) Spring 2016
Teaching Assistant, Peking University

Database Systems Principles and Technologies Fall 2016
Teaching Assistant, Peking University

SKILLS

Programming Languages	C/C++, Java, Python
Web	Django, Flask, Bootstrap, PHP, JavaScript, HTML
Platform/Tools	GNU/Linux (Ubuntu, Debian, CentOS), Mac OS, Git, GDB, Vim, SQL, LaTeX
Distributed Systems	Hadoop, Giraph
Libraries	Scikit-Learn, Keras, TensorFlow
Languages	Chinese (native), English (professional)