Xi (Sheryl) Zhang

Contact Information

Cornell University
Address: 425 E 61st Street, Suite 301, New York, NY 10065
Homepage: xi-zhang.net
Email: sheryl.zhangxi@gmail.com

Research Interests

Machine Learning Theory: Geometric Deep Learning, Causal Inference, Generative Models;
Applications: Health Care, Natural Language Processing, Recommender and Search Systems.

Work Experience

Cornell University — New York, US
- Postdoctoral Associate — Dec. 2016 - present
- Mentor: Fei Wang

Noah's Ark Lab, Huawei — Shenzhen, China
- Research Scientist — Aug. 2015 - Nov. 2016
- Mentors: Hang Li, Xin Jiang

Education

Institute of Automation, Chinese Academy of Sciences — Beijing, China
- Ph.D., Computer Science — Sep. 2011 - Jul. 2015
- Advisors: Hanqing Lu, Jian Cheng
- Studied at National Laboratory of Pattern Recognition

National Space Science Center, Chinese Academy of Sciences — Beijing, China
- M.S., Computer Science — Sep. 2008 - Jul. 2011
- Advisor: Bo Liu

Sichuan University — Chengdu, China
- B.E., Electrical Engineering — Sep. 2004 - Jul. 2008
- Rank: 3/173
- Recommended to Chinese Academy of Sciences

Selected Publication

Xi Zhang, Fengyi Tang, Hiroko Dodge, Jiayu Zhou, Fei Wang, "MetaPred: Meta-Learning for Clinical Risk Prediction with Limited Patient Electronic Health Records". *SIGKDD'19*: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2019. **AMIA Year-in-Review Highlight Paper**

Xi Zhang, Jingyuan Chou, Jian Liang, Cao Xiao, Yize Zhao, Harini Sarva, Claire Henchcliffe, Fei Wang, "Data-Driven Subtyping of Parkinson's Disease Using Longitudinal Clinical Records: A Cohort Study". Scientific Reports, Nature. **Blue Ribbon Highlights in Movement Disorder Society**

Xi Zhang, Jingyuan Chou, Fei Wang, "Integrative Analysis of Patient Health Records and Neuroimages via Memory-based Graph Convolutional Network". *ICDM'18*: IEEE International Conference on Data Mining, 2018.

Xi Zhang, Lifang He, Kun Chen, Yuan Luo, Jiayu Zhou, Fei Wang, "Multi-View Graph Convolutional Network and Its Applications on Neuroimage Analysis for Parkinson's Disease". *AMIA'18*: American Medical Informatics Association Annual Symposium, 2018.

Xi Zhang, Dandi Chen, Yongjun Zhu, Chao Che, Chang Su, Sendong Zhao, Xu Min, Fei Wang, "Multi-View Ensemble Classification for Clinically Actionable Genetic Mutations". *NeurIPS'17*:

Thirty-first Conference on Neural Information Processing Systems. **NeurIPS Competition: 1st Place Prize**

Xi Zhang, Jian Liang, Cao Xiao, Yize Zhao, Fei Wang, "Subtyping Parkinson's Disease with Recurrent Neural Network Models". *AMIA'17*: American Medical Informatics Association Annual Symposium, Abstract, 2017.

Inci M. Baytas, Cao Xiao, **Xi Zhang**, Fei Wang, Anil K. Jain, Jiayu Zhou, "Patient Subtyping via Time-Aware LSTM Networks". *SIGKDD'17*: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2017.

Jing Liu, Yu Jiang, Zechao Li, **Xi Zhang**, Hanqing Lu, "Domain-Sensitive Recommendation with User-Item Subgroup Analysis". *TKDE*: IEEE Transactions on Knowledge and Data Engineering, Vol. 28, Issue 4, pp.939-950, October, 2016.

Xi Zhang, Jian Cheng, Shuang Qiu, Hanqing Lu, "When Personalization Meets Conformity: Collective Similarity based Multi-Domain Recommendation". *SIGIR'15*: Annual ACM SIGIR Conference, 2015.

Xi Zhang, Jian Cheng, Shuang Qiu, Guibo Zhu, Hanqing Lu, "DualDS: A Dual Discriminative Rating Elicitation Framework for Cold Start Recommendation".*KBS*: Knowledge-based Systems, Vol.73, pp.161-172, January, 2015.

Ting Yuan, Jian Cheng, **Xi Zhang**, Qingshan Liu, Hanqing Lu, "How friends affect user behaviors? An exploration of social relation analysis for recommendation". *KBS*: Knowledge-based Systems, Vol.88, pp.70-84, November, 2015.

Cong Leng, Jiaxiang Wu, Jian Cheng, **Xi Zhang**, Hanqing Lu, "Hashing for Distributed Data", *ICML'15*: International Conference on Machine Learning, 2015.

Shuang Qiu, Jian Cheng, **Xi Zhang**, Hanqing Lu, "Exploring Heterogeneity for Multi-Domain Recommendation with Decisive Factors Selection". *WWW'15*: ACM International World Wide Web Conference, Poster, 2015.

Ting Yuan, Jian Cheng, **Xi Zhang**, Shuang Qiu, Hanqing Lu, "Recommendation by Mining Multiple User Behaviors with Group Sparsity". *AAAI'14*: AAAI Conference on Artificial Intelligence, 2014.

Cong Leng, Jian Cheng, Jiaxiang Wu, **Xi Zhang**, Hanqing Lu, "Supervised Hashing with Soft Constraints". *CIKM'14*: ACM International Conference on Information and Knowledge Management, 2014.

Xi Zhang, Jian Cheng, Ting Yuan, Biao Niu, Hanqing Lu, "Semi-Supervised Discriminative Preference Elicitation for Cold-Start Recommendation". *CIKM'13*: ACM International Conference on Information and Knowledge Management, 2013.

Xi Zhang, Jian Cheng, Ting Yuan, Biao Niu, Hanqing Lu, "TopRec: Domain-Specific Recommendation through Community Topic Mining in Social Network". *WWW'13*: ACM International World Wide Web Conference, 2013.

Biao Niu, Bin Li, Peng Li, **Xi Zhang**, Jian Cheng, Hanqing Lu, "Attribute Expansion with Sequential Learning for Object Classification", *ICME'13*: IEEE International Conference on Multimedia & Expo, 2013. Best Paper Runner-Up.

PATENT

Xi Zhang, Lin Ma, Xin Jiang, Hang Li, "Community Question Answering-based Article Recom-

mendation Method, System, and User Device", US Patent App. 16/444,618

Research Experience

Meta-Learning Algorithms and Applications
Postdoctoral Associate, Cornell University Dec. 2018 - present

- Proposed a meta-transfer embedding method for adversarial multisource domain adaptation
- Provided an effective knowledge transfer with limited clinical records through meta-learning

Deep Learning Algorithms for Heterogeneous Clinical Data Modeling
Postdoctoral Associate, Cornell University Dec. 2016 - Dec. 2018

- Identified progressive subtypes from longitudinal clinical records using neural networks for representation learning and statistical methods for subtype characterization
- Designed Memory-based Graph Convolutional Networks to integrating longitudinal clinical records and neuroimages, where the learned representation can be well interpreted
- Introduced a multiview Graph Convolutional Network for disease classification. In order to get stable results, pairwise training strategy is utilized

Semantic Models with Deep Architecture for Search and Recommendation
Research Scientist, Noah Ark's Lab, Huawei Mar. 2016 - Nov. 2016

- Explored neural probabilistic language models to learn concept embedding for appstore
- Captured ambiguous meanings of general query via neural network to improve user experiences
- Conducted semantic matching between long queries and apps with multi-type descriptions

Sparse Linear Method for Multiple User Behaviors Prediction
Research Assistant, National Lab of Pattern Recognition, CASIA Sep. 2013 - Feb. 2015

- Proposed a *collective similarity* to embed conformity prior knowledge into user modeling
- Provided an optimization algorithm by Alternating Direction Method of Multipliers (ADMM)
- Worked as a collaborator on a collective matrix factorization method with group sparsity

Automated Annotation Algorithm to Address Cold-Start Problems
Research Assistant, National Lab of Pattern Recognition, CASIA Dep. 2012 - Jul. 2014

- Modeled *representative* ability to select query set for annotation via sparse learning algorithm
- Proposed a semi-supervised rating elicitation method to select query set for interview
- Introduced a unified framework with co-clustering to jointly select interview sets

User Profiling with Probabilistic Topic Model
Research Assistant, National Lab of Pattern Recognition, CASIA Sep. 2011 - Feb. 2013

- Designed a novel recommendation framework with social relations to alleviate sparsity problem
- Analyzed user preference over explicit topics via a semi-supervised probabilistic topic model

Project Experience

Classifying Clinically Genetic Mutations via Natural Language Processing
Xi Zhang et al. Jul. 2017 - Oct. 2017

- Led the team in winning the 1st Place Prize in NeurIPS 2017 Challenge
- Designed a novel multi-view machine learning framework with ensemble classification models
- Solved the problem as text and relation classification simultaneously using neural networks

App Search System with Question & Answering
Xi Zhang, Xin Jiang, Shaokeng Zhu Jan. 2016 - Nov. 2016

- Primary member of app search team. Alleviated the critical miss-match problem
- Builded an architecture based on semantic matching with the tag representation and expand the description for both apps and queries
- Designed general query expansion strategy for full query stream in a real-time manner to recall miss-matched apps

Mobile App Tagging
Xi Zhang Sep. 2015 - Feb. 2016

- Sole contributor app tagging algorithms that supports other products including search, browsing, recommendation, and categorization, etc.
- Improved the tag prediction accuracy for 120K apps at 85+% precision over 30K noisy tags.
- Invented a novel tagging framework and utilized learning to rank techniques to generate encouraging results

Lecast Data Mining Contest: Designing Shoe Recommender System
Xi Zhang, Shuang Qiu Sep. 2013 - Oct. 2013

- Constructed a system to recommend shoes on a real-world E-commerce website with sparse searching records
- Solved the sparsity problem successfully by creating a novel method to transfer co-occurrence knowledge and textual information from other e-commercial websites to our target website
- Integrated multiple machine learning techniques including support vector machine, topic models, feature selection, and transfer learning to implement the system

Teaching

Cornell University
Health Data Mining (*Graduate*), Summer 2017, 2018. Guest Lecturer on Deep Neural Networks

Invited Talk

Microsoft Research Lab, Redmond, Sep. 2019
Clinical Data Modeling with Deep Neural Network: Challenges and Solutions
The American Academy of Neurology, Philadelphia, May 2019
Subtypes in Parkinson's Disease and Neural Networks (*Identified as a critical advance in the field of Neuroscience by the Science Committee*)

Professional Service

Program Chair: SDM'18 Workshop on Data Mining for Medicine and Healthcare
Reviewer: ICML'20, ICLR'20, NeurIPS'19, KDD'17,18,19, AMIA'17,18,19, CIKM'17,18, IJCAI'16, Transaction on Neural Networks and Learning Systems'18,19, Transaction on Knowledge and Data Engineering'16, Pattern Recognition'16,17,18, Transaction on the Web'15, Information Retrieval Journal'15
Sub-Reviewer: ICDM'17, AAAI'14,15,16, SDM'15, ACL'15, WWW'15, Knowledge-based Systems'15, Transaction on Knowledge and Data Engineering'15, Transaction on Intelligent Systems and Technology'14

Honors and Awards

AMIA Year-in-Review Highlight paper, 2019 Annual Symposium
Blue Ribbon Highlights in Movement Disorder Society (*Top 11 out of 1700 Papers*)
NeurIPS Competition: Classifying Genetic Mutations, 1st Place Prize (*Top 1 out of 1386 Teams*)
Lecast Data Mining Contest, 1st Place Prize (*Top 1 out of 30 Teams*)
Excellent Student Award, Institution of Automation, Chinese Academy of Sciences
Excellent Student Award, National Space Science Center, Chinese Academy of Sciences
Excellent Thesis of Undergraduate Students, Sichuan University
Outstanding Undergraduate Student Award, Sichuan University
FIRA Robot Football Contest, 2nd Place Prize (*Top 10 out of 78 Teams*)
Top Scholarship (*Several Times*), Sichuan University

Personal Skills and Interests

Languages: Speaks, Reads and Writes in Mandarin and English
Professional Skills: Modeling and Inference in AI, Optimization Algorithms
Programming Languages: Experienced in Python, Java, Matlab, C, LATEX; Familiar with C++
Deep Learning Libraries: Theano, Tensorflow, Pytorch
Open Source Code: `https://github.com/sheryl-ai`
Hobbies: Hiking, Photography, Classical Music, Literature