

# Zhaonan Sun

1101 Route 134 Kitchawan Rd  
Yorktown Heights, NY 10598.

Email: zhaonan.sun@gmail.com  
Phone: 765-464-4751

## EDUCATION

<i>Ph.D in Statistics</i> , Purdue University	2008 - 2014
<i>M. Sc. in Statistics</i> , Renmin University of China	2006 - 2008
<i>B. Sc. in Statistics</i> , Renmin University of China Magna Cum Laude with Honor	2002 - 2006

## PROFESSIONAL EXPERIENCE

- Research Staff Member, IBM T. J. Watson Research Center* Aug 2015 - present
- Developing machine learning and deep learning methodologies for analyzing real-world healthcare data in a broad range of research topics, and lead to
    - Proven record of high-impact publications
    - IBM Research Outstanding Technical Achievement Award (2018, 2019, 2021)
  - Team lead in the IBM-CHDI collaboration on Huntington's Disease (HD) progression modeling
    - Develop a novel staging system for HD based on multi-modal medical data
    - IBM Research A-level Accomplishments Reward: Scalable and Reusable Impact Science for Huntington's Disease Research
  - Technical lead in Disease Progression Modeling research area
  - Co-Chair, IBM Health Informatics Professional Interest Community

- Postdoctoral Researcher, IBM T. J. Watson Research Center* Aug 2014 - Jul 2015
- Developed large-scale risk screening method for comprehensive geriatric care
  - Developed multi-task learning method for evidence boost in drug repositioning

## PUBLICATIONS

1. Bum Chul Kwon, Vibha Anand, Kristen A Severson, Soumya Ghosh, **Zhaonan Sun**, Brigitte I Frohnert, Markus Lundgren, Kenney Ng. (2020) DPVis: Visual analytics with hidden markov models for disease progression pathways. *IEEE transactions on visualization and computer graphics*.
2. Bin Liu, Ying Li, Soumya Ghosh, **Zhaonan Sun**, Kenney Ng, Jianying Hu. (2019) Complication risk profiling in diabetes care: A bayesian multi-task and feature relationship learning approach. *IEEE Transactions on Knowledge and Data Engineering*. 32 (7), 1276-1289.

3. **Zhaonan Sun**, Soumya Ghosh, Ying Li, Yu Cheng, Amrita Mohan, Cristina Sampaio, Jianying Hu. (2019) A probabilistic disease progression modeling approach and its application to integrated Huntington's disease observational data. *Journal of American Medical Informatics Association Open*. 1(1), 123-130.
4. Zach Shahn, Ying Li, **Zhaonan Sun**, Amrita Mohan, Cristina Sampaio, Jianying Hu. (2019) G-Computation and Hierarchical Models for Estimating Multiple Causal Effects From Observational Disease Registries With Irregular Visits. *AMIA Summits on Translational Science Proceedings*. 789.
5. Sandra Liu, Jie Chen, **Zhaonan Sun**, Yu Zhu. (2018) From good to great: nonlinear improvement of healthcare service. *International Journal of Pharmaceutical and Healthcare Marketing*. 12(4), 391-408.
6. Jaehee Shim, **Zhaonan Sun**, Amos Cahan. (2018) Patient specific Vancomycin Dose Recommendation with Baseline Information at the Time of the First Dose. *Journal of Pharmacokinetics and Pharmacodynamics*. 45, S48-S49.
7. Bin Liu, Ying Li, **Zhaonan Sun**, Soumya Ghosh, Kenney Ng. (2018) Early Prediction of Diabetes Complications from Electronic Health Records: A Multi-task Survival Analysis Approach. *AAAI Conference on Artificial Intelligence(AAAI)*. 32 (1).
8. Zhengping Che, Yu Cheng, Shuangfei Zhai, **Zhaonan Sun**, Yan Liu. (2017) Boosting Deep Learning Risk Prediction with Generative Adversarial Networks for Electronic Health Records. *The IEEE International Conference on Data Mining(ICDM)*. 787-792.
9. Xiang Li, **Zhaonan Sun**, Xin Du, Haifeng Liu, Gang Hu, Guotong Xie. (2017) Bootstrap-based Feature Selection to Balance Model Discrimination and Predictor Significance: A Study of Stroke Prediction in Atrial Fibrillation. *American Medical Informatics Association Annual Symposium(AMIA)*. 1130-1139.
10. **Zhaonan Sun**, Ying Li, Soumya Ghosh, Yu Cheng, Amrita Mohan, Cristina Sampaio, Jianying Hu. (2017) A Data-Driven Method for Generating Robust Symptom Onset Indicators in Huntington's Disease Registry Data. *American Medical Informatics Association Annual Symposium(AMIA)*. 1635-1644.
11. Soumya Ghosh, Yu Cheng, **Zhaonan Sun**. (2017) Deep State Space Models for Computational Phenotyping. *IEEE International Conference on Health Informatics*. 399-402.
12. Ying Li, Ping Zhang, **Zhaonan Sun**, and Jianying Hu. (2016) Data-Driven Prediction of Beneficial Drug Combinations in Spontaneous Reporting Systems. *American Medical Informatics Association Annual Symposium (AMIA)*. 808-817.
13. Ben Li, **Zhaonan Sun**, Qing He, Yu Zhu, and Zhaohui Qin. (2016): Bayesian inference with historical data-based informative priors improves detection of differentially expressed genes. *Bioinformatics*. 32 (5): 682-689.
14. Ping Zhang, **Zhaonan Sun**, Fei Wang, and Jianying Hu. (2015). Towards computational drug repositioning: a comparative study of single-task and multi-task learning. *American Medical Informatics Association Annual Symposium (AMIA)*. Abstract.
15. **Zhaonan Sun**, Fei Wang, and Jianying Hu. (2015): LINKAGE: An Approach for Comprehensive Risk Prediction for Care Management. *21st ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*. Pages 1145-1154.
16. Jiang, Y., Frankenberger, J. R., Bowling, L. C., and **Sun, Z.** (2014). Quantification of Uncertainty in Estimated Nitrate-N Loads in Agricultural Watersheds. *Journal of Hydrology*. 519:106-116.

17. **Zhaonan Sun**, Thomas Kuczek, and Yu Zhu. (2014). Statistical calibration for qRT-PCR, microarray and RNA-Seq expression data with measurement error models. *The Annals of Applied Statistics*. 8(2):1022-1044.
18. **Zhaonan Sun**, Han Wu, Zhaohui Qin, and Yu Zhu. (2013). Model-Based Methods for Transcript Expression Level Quantification in RNA-Seq in *Advances in Statistical Bioinformatics: Models and Integrative Inference for High-Throughput Data*, edited by Do, K-A., Qin, S. and Vannucci, M. Cambridge University Press.
19. **Zhaonan Sun** and Yu Zhu. (2012). Systematic Comparison of RNA-Seq Normalization Methods Using Measurement Error Models. *Bioinformatics*. 28:20. Pages 2584-2591.
20. S.V.N. Vishwanathan, **Zhaonan Sun.**, Theera-Ampornpunt, N. and Varma, M. (2010). Multiple Kernel Learning and the SMO Algorithm. *NIPS*. Pages 3311-3325.
21. Xing Wang, Xi Wang and **Zhaonan Sun** (2009). Comparison on confidence bands of decision boundary between SVM and Logistic Regression. *Proceedings of fifth international joint conference on INC, IMS and IDC*.

## PATENTS

1. Yu Cheng, Soumya Ghosh, Jianying Hu, Ying Li, **Zhaonan Sun**. Identifying and indexing discriminative features for disease progression in observational data *US Patent Application 15/799664*
2. **Zhaonan Sun**, Zach Shahn, Ping Zhang, Fei Wang, Jianying Hu. Determining potential cancer therapeutic targets by joint modeling of survival events. *US Patent Application 16/011231*
3. Ying Li, Ping Zhang, **Zhaonan Sun** and Jianying Hu. Data-Driven Prediction of Drug Combinations That Mitigate Adverse Drug Reaction. *US Patent Application US 15/257535*.
4. **Zhaonan Sun**, Ping Zhang, Fei Wang, and Jianying Hu. Evidence Boosting In Rational Drug Design And Indication Expansion By Leveraging Disease Association. *US Patent Application 14/929995*.
5. **Zhaonan Sun**, Fei Wang, and Jianying Hu. A Method For Proactive Comprehensive Geriatric Risk Screening. *US Patent Application 15/048413*.
6. Xiang Li, **Zhaonan Sun**, Haifeng Liu, Jingjing Tao, Gang Hu, Guotong Xie. A Method to Identify Risk Factors with Both Predictive Power and Statistical Significance. Chinese Patent Application CN920160040.

## PROGRAM COMMITTEE

1. International Conference on Machine Learning (ICML) 2021
2. International Joint Conference on Artificial Intelligence(IJCAI) 2021
3. ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD) 2021
4. NIPS Machine Learning for Health (ML4H) workshop 2019
5. International Joint Conference on Artificial Intelligence(IJCAI) 2019
6. ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD) 2019
7. ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD) 2018

8. NIPS Machine Learning for Health (ML4H) workshop 2017
9. IEEE International Conference on Healthcare Informatics 2016
10. International Joint Conference on Artificial Intelligence(IJCAI) 2016
11. ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD) 2016
12. International Joint Conference on Artificial Intelligence 2015
13. IEEE International Conference on Healthcare Informatics 2015
14. KDD 2015 Workshop on "BigCHat: Connected Health at Big Data"
15. 1st Workshop on Matrix Computations for Biomedical Informatics 2015