# Zhaonan Sun

1101 Route 134 Kitchawan Rd
Yorktown Heights, NY 10598.

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

---

## EDUCATION

*Ph.D in Statistics*, Purdue University                              2008 - 2014

*M. Sc.* in Statistics, Renmin University of China              2006 - 2008

*B. Sc.* in Statistics, Renmin University of China              2002 - 2006
  Magna Cum Laude with Honor

## PUBLICATIONS

1. Soumya Ghosh, Yu Cheng, **Zhaonan Sun**. (2016) Deep State Space Models for Computational Phenotyping. *IEEE International Conference on Health Informatics*. To Appear.
2. 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)*. To Appear.
3. **Zhaonan Sun**, Xu Liu, Ping Zhang, Jianying Hu, Juan Wisnivesky, and Fei Wang. (2016). Joint Modeling of Survival Events through Multi-task Learning Framework. *American Medical Informatics Association Annual Symposium (AMIA)*. Abstract. To Appear.
4. **Zhaonan Sun**, Yu Cheng, Amos Cahan, Fei Wang, and Jianying Hu. (2016): Modelling the Progression of CKD with EMR Data: a Partially Hidden Markov Approach. *AMIA Summit on Clinical Research Informatics (CRI)*. Abstract.
5. 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.
6. 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.
7. **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.
8. 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.
9. **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.
10. **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.

11. **Zhaonan Sun** and Yu Zhu. (2012). Systematic Comparison of RNA-Seq Normalization Methods Using Measurement Error Models. *Bioinformatics*. 28:20. Pages 2584-2591.

12. S.V.N. Vishwanathan, **Zhaonan Sun.**, Theera-Ampornpunt, N. and Varma, M. (2010). Multiple Kernel Learning and the SMO Algorithm. *NIPS*. Pages 3311-3325.

13. 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. **Zhaonan Sun**, Ping Zhang, Fei Wang, and Jianying Hu. Evidence Boosting In Rational Drug Design And Indication Expansion By Leveraging Disease Association. *United States Patent Application 14/929995*.

2. 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.

## PROFESSIONAL EXPERIENCE

*Research Staff Member, IBM T. J. Watson Research Center*              Aug 2015 - present
- Developing disease progression model with EHR systems
- Developing machine learning method for drug repositioning

*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

*Statistical Consultant, Purdue University*              Aug 2013 - May 2014
- Provided data analysis and statistical software support for Purdue community.

*Research Assistant, Purdue University*              June 2011 - July 2013
- Developed statistical method for assessing and improving measurement errors in RNA-Seq gene expression data
- Developed statistical method for identify differential gene expression using RNA-Seq data
- Involved in extending SMO algorithm in multiple kernal learning

## TALKS

1. *The Second Statistical Forum on Huntington's Disease*, Princeton, NJ, 2016. Machine learning for disease progression models.

2. *American Medical Informatics Association Annual Symposium*, San Francisco, CA, 2015. A graph based methodology for temporal signature identification from EHR.

3. *Joint Statistical Meetings*, Seattle, WA, 2015. Comprehensive Risk Prediction Using Interactive Graph-Guided Fussed Lasso Penalty.

4. *IBM Research Health Informatics PICs*, Yorktown Heights, NY, 2015. Multi-task learning approach for comprehensive risk prediction.

5. *Eastern North American Region Meetings*, Baltimore, MD, 2014. Statistical calibration of qRT-PCR, microarray and RNA-Seq gene expression data with measurement error models.

6. *Purdue Bioinformatics Seminar*, West Lafayette, IN, 2014. Statistical calibration of high-throughput gene expression data using measurement error models.

7. *Joint Statistical Meetings*, San Diego, CA, 2012. Differential gene expression pattern analysis using exon-level RNA-Seq data.

8. *Joint Statistical Meetings*, Miami, FL, 2011. An integrative approach to comparing and normalizing gene expression data generated from RNA-Seq, Microarray and RT-PCR technologies.

## PROGRAM COMMITTEE

1. 
2. IEEE International Conference on Healthcare Informatics 2016
3. International Joint Conference on Artificial Intelligence(IJCAI) 2016
4. ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD) 2016
5. International Joint Conference on Artificial Intelligence 2015
6. IEEE International Conference on Healthcare Informatics 2015
7. KDD 2015 Workshop on "BigCHat: Connected Health at Big Data"
8. 1st Workshop on Matrix Computations for Biomedical Informatics 2015

## AWARD

1. IBM BlueThx Award for Client Success (2016)
2. IBM Excellent Invention Award (2015)
3. IBM Manager's Choice Award (2015, 2016)

## PROGRAMMING SKILLS

python, R, SQL, SAS, MATLAB, LaTeX