

# Resume ZHENG ZHAO

---

SAS Headquarters, S6085, 600 Research Drive, Cary, NC, 27513, USA  
Voice: (919) 889 7779 Fax: (919) 677 4444 E-mail: zheng.zhao@sas.com

**Career Interests:** Statistical Predictive Modeling, Machine Learning, Data Mining, High Performance Computing, Database and Data Warehousing, Knowledge and Data Integration

## Education:

- **(Aug-2004 – July-2010) Ph.D.**, Computer Science and Engineering  
Arizona State University, Tempe, AZ, USA.,  
Thesis topic: Handling Data of Extremely High Dimensionality in Data Mining  
Area of Study: Data Mining, Machine Learning
- **(Sept-1996 – Jun-2002) M.S., B.S.** Computer Science and Technology  
Harbin Institute of Technology, Harbin, Heilong Jiang, China,  
Area of Study: Relational Database System

## Research Interests:

- Data mining and machine learning
- Dimension reduction, feature selection
- Text mining, bioinformatics
- High performance parallel processing

## Awards:

- **Best Paper Award:** Biological relevance detection via network dynamic analysis  
In the 2nd International Conference on Bioinformatics and Computational Biology (BICoB), 2010
- **Graduate & Professional Student Association (GPSA) Research Grant Award,**  
Graduate Research Support Program, ASU, 2009-2010.
- **Outstanding PhD Student,**  
The School of Computing, Informatics, and Decision Systems Engineering, ASU, 2008.

## Work Experience:

- **Research Statistician**, SAS Institute, Inc. NC, USA, July 2010 – present  
**Responsibility:** Developing SAS Procedures for large scale data mining and text mining
- **Software Engineer**, Core R&D Dept. of Alcatel Shanghai Bell (ASB), Shanghai, China. 2002-2004.  
**Responsibility:** Maintain the software control component of the telephone exchange system A1000 S12 CDMA.

## Professional Services

- **Co-chair** for the fourth international workshop on feature selection in data mining (FSDM 2010). Hyderabad, India, 21-24 June, 2010
- **Program committee member** of the 24th AAAI conference on artificial intelligence (AAAI 2010) AIW Track. Atlanta, Georgia, USA, July 11–15, 2010.
- **Program committee member** of the 22nd International Joint Conferences on Artificial Intelligence (IJCAI 2011). Barcelona, Spain, 16-22, July, 2011

- **Journal paper review:** Bioinformatics (**BIOINF**), Data Mining and Knowledge Discovery (**DMKD**), Data & Knowledge Engineering (**DKE**), International Journal of Pattern Recognition and Artificial Intelligence (**IJPRAI**), Journal of Machine Learning (**JMLR**), Pattern Recognition (**PR**), Applied Intelligence (**APIN**), Information Fusion (**IF**), International Journal of Functional Informatics and Personalized Medicine (**IJFIPM**), Transactions on Knowledge and Data Engineering (**TKDE**), Information Sciences (**INS**), Neurocomputing (**NEUCOM**)
- **Conference paper review:** International Conference on Web-Age Information Management (**WAIM**), AAAI Conference on Artificial Intelligence (**AAAI**), ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (**KDD**), European Conference on Computer Vision (**ECCV**), Workshop on Feature Selection in Data Mining (**FSDM**), International Joint Conferences on Artificial Intelligence (**IJCAI**)

## Publication

- **Book**

1. Zheng Zhao, Huan Liu. **Spectral Feature Selection for Data Mining**. CRC Press, January 2012.

- **Book Chapters**

2. Huan Liu and Zheng Zhao. Encyclopedia of Complexity and Systems Science, chapter **Manipulating Data and Dimensionality Reduction Methods: Feature Selection**, 5348-5359. Springer, 2009.
3. Zheng Zhao and Huan Liu. Encyclopedia of Data Warehousing and Mining, chapter **Interacting Features in Subset Selection**, 1079-1084. Idea Group, Inc, 2008.
4. Sai Moturu, Lance Parsons, Zheng Zhao, and Huan Liu. Encyclopedia of Data Warehousing and Mining, chapter **Integrative Data Analysis for Biological Discovery**, page 1058-1065. Idea Group, Inc, 2008.

- **Thesis and Dissertation**

5. Zheng Zhao. **Spectral feature selection for mining ultrahigh dimensional data**, Doctoral Dissertation, Arizona State University, 2010.
6. Zheng Zhao. **Algorithms for materialized view design in data warehousing environment**. Master's Thesis, Harbin Institute of Technology, 2002.

- **Proceeding Edited**

7. Huan Liu, Hiroshi Motoda, Rudy Setiono, Zheng Zhao. **Proceedings of the Fourth International Workshop on Feature Selection in Data Mining**, JMLR Workshop and Conference Proceedings Volume 10, 2010

- **Journal Articles**

8. Zheng Zhao, Lei Wang, Huan Liu, and Jieping Ye. **On similarity preserving feature selection**. IEEE Transactions on Knowledge and Data Engineering, to appear 2011.
9. Gang Zhao, Jiawei Chen, Yanqiu Deng, Feng Gao, Jiwei Zhu, Zhenzhong Feng, Xiuhong Lv, Zheng Zhao. **Identification of NDRG1-regulated genes associated with invasive potential in cervical and ovarian cancer cells**. Biochemical and Biophysical Research Communications, 408(1), Pages 154-159, 2011.
10. Jiangxin Wang, Xiaoshan Zhu, Xuezhi Zhang, Zheng Zhao, Huan Liu, Rajani George, Jeanne Wilson-Rawls, Yung Chang, and Yongsheng Chen. **Disruption of zebrafish (Danio rerio) reproduction upon chronic exposure to TiO<sub>2</sub> nanoparticles**. Chemosphere, 83(4), page 461-7, 2011.
11. Zheng Zhao and Huan Liu. **Searching for interacting features in subset selection**. Intelligent Data Analysis (IDA), 13:207-228, 2009.
12. Huan Liu, Edward R. Dougherty, Jennifer G. Dy, Kari Torkkola, Hanchuan Peng, Fuhui Long, Michael Berens, Lance Parsons, Zheng Zhao, Lei Yu, and George Forman. **Evolving feature selection**. IEEE Intelligent Systems, pages 64-76, 2005.

13. Ni min, Xiaofei Xu, Deng Shengchun, Zhao Zheng. **Freq\_Apriori: An Association Rule Mining Algorithm Based on Frequent Itempair Support Matrix**, Journal of Mini-Micro Systems, Vol. 25, 1-4, 2004.
14. Yanqiu Deng, Gang Zhao, and Zheng Zhao. **Expression of immediate early gene c-jun in rat with focal cerebral ischemic tolerance**. Chinese Journal of Clinical Rehabilitation, 8(13):2474-2475, 2004.
15. Deng Yanqiu, Chen Yingzhun, and Zhao Zheng. **The study on the apoptosis-related protein dynamic expression after focal cerebral ischemic preconditioning**. Chinese Journal of experiment surgery. 21(5):636-8, 2004.
16. Zheng Zhao, Xiaofei Xu, and Shengchun Deng. **Materialized view selection using improved genetic algorithms**. Journal of Computer Science and Technology, 29(8):139-142, 2002.

- **Conference Papers**

17. Zheng Zhao, Lei Wang, and Huan Liu. **Efficient Spectral Feature Selection with Minimum Redundancy**. In Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence (AAAI), 2010
18. Zheng Zhao, Jiangxin Wang, Shashvata Sharma, Nitin Agarwal, Huan Liu, and Yung Chang. **An Integrative Approach to Identifying Biologically Relevant Genes**. In Proceedings of SIAM International Conference on Data Mining (SDM), 2010.
19. Zheng Zhao, Jiangxin Wang, Huan Liu, and Yung Chang. **Biological relevance detection via network dynamic analysis**. In Proceedings of the 2nd International Conference on Bioinformatics and Computational Biology (BICoB), 2010. **(was granted the best paper award)**
20. Zheng Zhao, Liang Sun, Shipeng Yu, Huan Liu, and Jieping Ye. **Multiclass probabilistic kernel discriminant analysis**. In Proceedings of the 21th International Joint Conference on AI (IJCAI), 2009.
21. Zheng Zhao, Jiangxin Wang, Shashvata Sharma, Nitin Agarwal, Huan Liu, and Yung Chang. **A knowledge-oriented framework for gene selection**. In 13th Annual International Conference on Research in Computational Molecular Biology (RECOMB), 2009. Poster.
22. Zheng Zhao, Jiangxin Wang, Huan Liu, Jieping Ye, and Yun Chang. **Identifying biologically relevant genes via multiple heterogeneous data sources**. In Proceedings of The 14th ACM SIGKDD International Conference On Knowledge Discovery and Data Mining (KDD), 2008.
23. Jieping Ye, Kewei Chen, Teresa Wu, Jing Li, Zheng Zhao, Rinkal Patel, Min Bae, Ravi Janardan, Huan Liu, Gene Alex Ander, and E. Reiman. **Heterogeneous data fusion for alzheimer's disease study**. In Proceedings of The 14th ACM International Conference On Knowledge Discovery and Data Mining (KDD), 2008.
24. Zheng Zhao and Huan Liu. **Semi-supervised feature selection via spectral analysis**. In Proceedings of SIAM International Conference on Data Mining (SDM), 2007.
25. Zheng Zhao and Huan Liu. **Spectral feature selection for supervised and unsupervised learning**. In Proceedings of The 24th Annual International Conference on Machine Learning (ICML), 2007.
26. Zheng Zhao and Huan Liu. **Searching for interacting features**. In Proceedings of the 20th International Joint Conference on AI (IJCAI), January 2007.
27. Jieping Ye, Zheng Zhao, and Huan Liu. **Adaptive distance metric learning for clustering**. In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2007.
28. Jieping Ye, Zheng Zhao, and Mingrui Wu. **Discriminative k-means for clustering**. In Proceedings of 21th Advances in Neural Information Processing Systems (NIPS), 2007.
29. Jianhu Chen, Zheng Zhao, Jieping Ye, and Huan Liu. **Nonlinear adaptive distance metric learning for clustering**. In Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2007.
30. He Zengyou, Deng ShengChun, Xiaofei Xu, and Zhao Zheng. **Induction of Global Rules Considering Schema Conflicts in Multidatabase System**, in the proceedings of 19th National Conference on Databases (NDBC), 2001.

- **Workshop Papers**

31. Salem Alelyani, Zheng Zhao, and Huan Liu. **A Dilemma in Assessing Stability of Feature Selection Algorithms**. In the 13th International Workshop on Future Trends of Distributed Computing Systems (FTDCS 2011), 2011.

32. Huan Liu, Hiroshi Motoda, Rudy Setiono, and Zheng Zhao. **Feature Selection: An Ever Evolving Frontier in Data Mining**. Journal of Machine Learning Research - Proceedings Track 10: 4-13 (2010)
33. Zheng Zhao, Jiangxin Wang, Shashvata Sharma, Huan Liu, and Yung Chang. **Integrating knowledge in search of biologically relevant genes**. In the 3rd International Workshop on Domain Driven Data Mining (DDDM), 2009.
34. Zheng Zhao and Huan Liu. **Multi-source feature selection via geometry-dependent covariance analysis**. Journal of Machine Learning Research (JMLR), Workshop and Conference Proceedings Volume 4: New challenges for feature selection in data mining and knowledge discovery, 4:36-47, 2008.
35. Keshu Zhang, Kari Torkkola, Haifeng Li, Christopher Schreiner, Harry Zhang, Mike Gardner, and Zheng Zhao. **A context aware automatic traffic notification system for cell phones**. In Proceedings of The 7th International Workshop on Smart Appliances and Wearable Computing (IWSAWC'07), 2007.

#### **Technical Skills:**

- SAS System: SAS programming language, SAS PROC Programming, High Performance Computing **(highly proficient)**
- Statistics Modeling, Matrix Computation, Convex and Nonconvex Optimization, Parallel Computing, Software Engineering **(highly proficient)**
- Supervised/Unsupervised/Semi-supervised Learning, Ensemble Learning, Bayesian Inference, Feature Selection/Extraction, Sparse Learning, Distributed and Parallel Predictive Modeling **(6 years of daily experience)**
- Programming: C/C++, SAS, Java, Matlab, R, Python, PHP, MPI/OpenMPI, MapReduce **(10 years of daily experience)**
- Databases: SQL Server, MySQL, Oracle **(proficient)**
- Operating Systems: Windows, LINUX/Unix, MAC **(proficient)**