Jie Liu

Mail Address South Foege Building, S-220, 3720 15th Ave NE, Seattle WA 98195
E-Mail liu6@uw.edu
Home Page www.cs.wisc.edu/~jieliu
Phone (608) 698-8016

Research Interests

Machine learning: massive scale data analysis, graphical models, statistical relational learning
Statistics: large-scale statistical inference, Bayesian statistics, nonparametric statistics
Bioinformatics and medical informatics: personalized medicine, genome-wide association studies, population genetics, biological sequence analysis

Education

PhD in Computer Science, University of Wisconsin-Madison, 07/2014
    Advisor: Prof David Page
    Thesis: Statistical Methods for Genome-wide Association Studies and Personalized Medicine
Master in Statistics, University of Wisconsin-Madison, 05/2011
Master in Computer Science, University of Wisconsin-Madison, 05/2009
Master in Signal and Information Processing, Peking University, 07/2007
Bachelor in Info. Management & Info. Systems, Beijing Univ. of Posts and Telecom, 07/2004

Employment

Research Associate, Department of Genome Sciences and eScience Institute, UW-Seattle
Washington Research Foundation Innovation Postdoctoral Fellow in Data Science
Moore/Sloan Data Science Postdoctoral Fellow
09/2014-present, mentors: Prof William S. Noble, Prof Jeffrey A. Bilmes

Develop machine learning methods for analyzing big genomic and proteomic data (supported by Moore Foundation/Sloan Foundation Data Science Environments Project, Washington Research Foundation Fund for Innovation in Data Science, and NIH's 4D Nucleome Project)

Research Assistant, Department of Biostatistics & Medical Informatics, UW-Madison
01/2008-07/2014, PI: Prof David Page, Prof Elizabeth Burnside

Develop statistical machine learning methodologies, and apply them to real-world genome-wide association studies on different diseases and clinical data, and translate our methods into a health care decision making system for personalized care (supported by several grants from NCI, NIGMS, NLM, NIEHS)

Project Assistant, Department of Radiology, UW-Madison
09/2007-12/2007, PI: Prof Elizabeth Burnside

Develop a health care decision-making system for breast cancer risk prediction and personalized health care

Research Assistant, Center for Information Science, Peking University
09/2004-07/2007, PI: Prof Shaohua Tan

Develop a web-based intelligent financial analysis platform to support large-scale financial data analysis

Researcher (part-time), Department of Neuroscience, Univ. of Minnesota-Twin Cities
07/2004-12/2006, PI: Prof Tongbin Li

Develop meta-prediction methods applied in kinase recognition and protein subcellular localization

Peer-reviewed Publications

Machine Learning and Statistics

- Zhaobin Kuang*, Sinong Geng*, **Jie Liu**, Stephen Wright, and David Page. Stochastic Learning for Sparse Discrete Markov Random Fields with Controlled Gradient Approximation Error. *Uncertainty in Artificial Intelligence (UAI),* 2018. (* indicates equal contribution)
- **Jie Liu**, Chunming Zhang, and David Page. Multiple Testing under Dependence via Graphical Models. *Annals of Applied Statistics*, 2016.
- Jun Fan, Yirong Wu, Ming Yuan, David Page, **Jie Liu**, Irene Ong, Peggy Peissig and Elizabeth Burnside. Structure-leveraged Methods in Breast Cancer Risk Prediction. *Journal of Machine Learning Research*, 17, 1-15, 2016.
- **Jie Liu**, Chunming Zhang, Elizabeth Burnside and David Page. Learning a Semiparametric Graphical Model for Multiple Testing under Dependence. The *31st International Conference on Machine Learning (ICML),* 2014.
- **Jie Liu**, Chunming Zhang, Elizabeth Burnside and David Page. Learning Heterogeneous Hidden Markov Random Fields. The 17*th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2014.
- **Jie Liu** and David Page. Bayesian Estimation of Latently-grouped Parameters in Undirected Graphical Models. *Neural Information Processing Systems (NIPS),* 2013.
- **Jie Liu**, Chunming Zhang, Catherine McCarty, Peggy Peissig, Elizabeth Burnside and David Page. Graphical-model Based Multiple Testing under Dependence, with Applications to Genome-wide Association Studies. *Uncertainty in Artificial Intelligence (UAI)* (oral presentation), 2012.
- **Jie Liu**, Chunming Zhang, Catherine McCarty, Peggy Peissig, Elizabeth Burnside and David Page. High-Dimensional Structured Feature Screening Using Binary Markov Random Fields. The 15*th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2012.

Medical Informatics, Bioinformatics and Other Machine Learning Applications

- **Jie Liu**, Dejun Lin, Gurkan Yardimci, and William S. Noble. Unsupervised embedding of single-cell Hi-C data. *The 26th Annual Conference on Intelligent Systems for Molecular Biology (ISMB),* 2018.
- Shara I. Feld, Kaitlin M. Woo, Roxana Alexandridis, Yirong Wu, **Jie Liu**, Peggy Peissig, Adedayo A. Onitilo, Jennifer Cox, C. David Page, Elizabeth S. Burnside. Improving breast cancer risk prediction by using demographic risk factors, abnormality features on mammograms and genetic variants. *AMIA Annual Symposium*, 2018.
- **Jie Liu**, John T. Halloran, Jeffrey A. Bilmes, Riza M. Daza, Choli Lee, Elisabeth M. Mahen, Donna Prunkard, Chaozhong Song, Sibel Blau, Michael O. Dorschner, Vijayakrishna K. Gadi, Jay Shendure, C. Anthony Blau, and William S. Noble. Comprehensive statistical inference of the clonal structure of cancer from multiple biopsies. *Scientific Reports*, 2017.
- Yirong Wu, Craig K. Abbey, **Jie Liu**, Irene Ong, Peggy Peissig, Adedayo A. Onitilo, Jun Fan, Ming Yuan and Elizabeth S. Burnside. Discriminatory Power of Common Genetic Variants in Personalized Breast Cancer Diagnosis. *SPIE Medical Imaging conference* (oral presentation), 2016.

- Elizabeth S. Burnside, **Jie Liu**, Yirong Wu, Adedayo A. Onitilo, Catherine A. McCarty, C. David Page, Peggy L. Peissig, Amy Trentham-Dietz, Terrie Kitchner, Jun Fan and Ming Yuan. Comparing Mammography Abnormality Features and Genetic Variants in the Prediction of Breast Cancer in Women Recommended for Breast Biopsy. *Academic Radiology* 23(1), 2016.
- Yirong Wu, Craig K. Abbey, Xianqiao Chen, **Jie Liu**, David C. Page, Oguzhan Alagoz, Peggy Peissig, Adedayo A. Onitilo, Elizabeth S. Burnside. Developing a Utility Decision Framework to Evaluate Predictive Models in Breast Cancer Risk Estimation. *Journal of Medical Imaging*, 2015.
- Jeremy Weiss, Finn Kuusisto, Kendrick Boyd, **Jie Liu**, David Page. Machine Learning for Treatment Assignment: Improving Individualized Risk Attribution. *American Medical Informatics Association Symposium (AMIA)*, 2015.
- **Jie Liu**, Yirong Wu, Irene Ong, David Page, Peggy Peissig, Catherine McCarty, Adedayo A. Onitilo and Elizabeth Burnside. Leveraging Interaction between Genetic Variants and Mammographic Findings for Personalized Breast Cancer Diagnosis. *AMIA Joint Summits on Translational Science* (regular paper, podium presentation), 2015.
- Yirong Wu, **Jie Liu**, Alejandro Munoz del Rio, David C. Page, Oguzhan Alagoz, Peggy Peissig, Adedayo A. Onitilo and Elizabeth S. Burnside. Developing a Clinical Utility Framework to Evaluate Prediction Models in Radiogenomics. *SPIE Medical Imaging conference* (oral presentation), 2015.
- Yirong Wu, **Jie Liu**, David Page, Peggy Peissig, Catherine McCarty, Adedayo A. Onitilo and Elizabeth S. Burnside. Comparing the Value of Mammographic Features and Genetic Variants in Breast Cancer Risk Prediction. *American Medical Informatics Association Symposium (AMIA)* (regular paper, podium presentation), 2014.
- **Jie Liu**, David Page, Peggy Peissig, Catherine McCarty, Adedayo A. Onitilo, Amy Trentham-Dietz and Elizabeth Burnside. New Genetic Variants Improve Personalized Breast Cancer Diagnosis. *AMIA Joint Summits on Translational Science* (regular paper, podium presentation, **Marco Ramoni Distinguished Paper Award**), 2014.
- **Jie Liu**, David Page, Houssam Nassif, Jude Shavlik, Peggy Peissig, Catherine McCarty, Adedayo A. Onitilo and Elizabeth Burnside. Genetic Variants Improve Breast Cancer Risk Prediction on Mammograms. *American Medical Informatics Association Symposium (AMIA)* (regular paper, podium presentation), 2013.
- **Jie Liu**, Elizabeth Burnside and David Page. Predicting Breast Cancer and Prostate Cancer Susceptibility from Single Nucleotide Polymorphisms. *ICML 2013 Workshop on Role of Machine Learning in Transforming Healthcare*, 2013.
- **Jie Liu** and David Page. Structure Learning of Undirected Graphical Models with Contrastive Divergence. *ICML 2013 Workshop on Structured Learning: Inferring Graphs from Structured and Unstructured Inputs*, 2013.
- **Jie Liu**, Humberto Vidaillet, Elizabeth Burnside and David Page. A Collective Ranking Method for Genome-wide Association Studies. *ACM Conference on Bioinformatics, Computational Biology and Biomedicine (ACM BCB)*, 2012.
- **Jie Liu***, Zaixia Hu* and Shaohua Tan. Cross-Sectional Stock Return Analysis Using Support Vector Regression. *Applied Economics Letters* 17(1), 2009. (* indicates equal contribution)

- Ji Wan, Shuli Kang, Chuanning Tang, Jianhua Yan, Yongliang Ren, **Jie Liu**, Xiaolian Gao, Arindam Banerjee, Lynda B.M.Ellis and Tongbin Li. Meta-prediction of Phosphorylation Sites with Weighted Voting and Restricted Grid Search Parameter Selection. *Nucleic Acids Research* 36(4), 2008.
- **Jie Liu**, Shuli Kang, Chuanning Tang, Lynda B.M. Ellis and Tongbin Li. Meta-prediction of Protein Subcellular Localization with Reduced Voting. *Nucleic Acids Research* 35(15), 2007.

Preprints

- Jung-Sik Kim, Xiaoyuan He, **Jie Liu**, Zhijun Duan, Taeyeon Kim, Julia Gerard, Brian Kim, William S. Lane, William S. Noble, Bogdan Budnik, Todd Waldman. Cohesin interacts with a panoply of splicing factors required for cell cycle progression and genomic organization. *Biorxiv*, 2018.
- Nao Hiranuma, **Jie Liu**, Chaozhong Song, Jacob Goldsmith, Michael Dorschner, Colin Pritchard, Kimberly Burton, Elisabeth Mahen, Sibel Blau, Francis Senecal, Wayne Monsky, Stephanie Parker, Stephen Schmechel, Stephen Allison, VK Gadi, Sofie Salama, Amie Radenbaugh, Mary Goldman, Jill Johnsen, Shelly Heimfeld, Vitalina Komashko, Marissa LaMadrid-Hermannsfeldt, Zhijun Duan, Steven Benz, Patrick Soon-Shiong, David Haussler, Jingchun Zhu, Walter Ruzzo, William Noble, and C. Anthony Blau. Cis-compound mutations are prevalent in triple negative breast cancer and can drive tumor progression. *Biorxiv*, 2016.
- Yirong Wu, **Jie Liu**, Peggy Peissig, Adedayo A. Onitilo, Jun Fan, Ming Yuan, Jennifer Cox and Elizabeth S. Burnside. Improving Breast Cancer Risk Prediction in Women Recommended for Biopsy using Abnormality Features on Mammograms. *Under Review.*

Professional Service

Journal Referee: Journal of the American Statistical Association, Data Mining and Knowledge Discovery, Machine Learning, Pattern Recognition Letters, Genome Research, Cell Systems, Bioinformatics, Journal of Bioinformatics and Computational Biology, Journal of Digital Imaging, Economic Modelling, PLOS ONE

Conference Program Committee: ICLR 2019, AMIA Informatics Summit 2019, NIPS 2018, ISMB 2018, AMIA 2018, ICML 2018, RECOMB 2018, ICLR 2018, AMIA Informatics Summit 2018, NIPS-MLCB 2017, NIPS 2017, AMIA 2017, AMIA-TBI/CRI 2017, AAAI 2017, NIPS 2016, AMIA 2016, AAAI 2016, AMIA-TBI/CRI 2016, ACMBCB 2016, MLSB 2016, NIPS 2015, IJCAI 2015, AMIA 2015, AMIA-TBI/CRI 2015, AISTATS 2015, WABI2015, NIPS 2014, UAI 2014, AAAI 2014, AMIA 2014, IJCAI 2013, AMIA 2013, GENSIPS 2012

Honors and Awards

2015	Washington Research Foundation Innovation Postdoctoral Fellowship
2014	AMIA Marco Ramoni Distinguished Paper Award
2014	Moore/Sloan Data Science Postdoctoral Fellowship
2004-2007	PKU graduate student fellowship
2004	Merit Student of Beijing City (top 1%)
2001-2003	BUPT student fellowship (top 1%)

Teaching and Mentoring

Courses

Adjunct instructor, CS 760 Machine Learning, UW-Madison, 2014 Spring.
Guest instructor, STAT 992 Large-scale inference, UW-Madison, 2013 Spring.
Teaching assistant, Introduction to Artificial Intelligence, Peking University, 2005 Spring.

Mentees

Kaishu Mason, undergraduate student, Harvard University (NSF REU summer research program)
Charles Kwong, PhD student, Department of Computer Sciences, UW-Madison
Sinong Geng, PhD student, Department of Statistics, UW-Madison
John T Halloran, PhD student, Department of Electrical Engineering, University of Washington

Invited Talks

"Tumor heterogeneity analysis via an extensible modeling system"
eScience Institute, University of Washington, December 2016

"Genomic Medicine, from the Perspective of Data Science"
eScience Institute, University of Washington, January 2015

"The Promise and Computational Challenges of Personalized Medicine"
Lane Center for Computational Biology, Carnegie Mellon University, March 2014
Department of Computer Science, University of Michigan, Ann Arbor, March 2014
IBM T. J. Watson Research Center, March 2014
"Multiple Testing under Dependence with Applications to Genome-wide Association Studies"
Department of Computer Science, Colorado State University, March 2013

References

William Stafford Noble, Professor
Dept. of Genome Sciences, University of Washington
Email: wnoble@uw.edu

Jeffrey A. Bilmes, Professor
Dept. of Electrical Engineering, University of Washington
Email: bilmes@uw.edu

David Page, Professor
Dept. of Biostatistics and Medical Informatics, University of Wisconsin-Madison,
Email: page@biostat.wisc.edu

Elizabeth Burnside, Professor
Department of Radiology, University of Wisconsin-Madison
Email: eburnside@uwhealth.org