Cheng Liu, PhD

Department of Statistics
Purdue University (509)339-5609
West Lafayette, IN 47907-2066 *liu105@purdue.edu*

OBJECTIVE

Highly motivated PhD with strong experience in statistical research, data analytics, data mining, machine learning, and statistical consulting seeking data scientist position to refine statistical methodologies for big data analysis.

EDUCATION

Ph.D., Statistics, **Purdue University**, West Lafayette, IN Aug 2008 - Oct 2013

Dissertation: Non-parametric Spatial Modeling

M.S., Statistics, **Washington State University**, Pullman, WA Aug 2006 - May 2008

Thesis: Process Capability Indices for Multivariate Quality Control

B.S., Information and Computing Science, HuaZhong University of Science and Technology, Wuhan, China Aug 2002 - May 2006

INTERNSHIP

Data Scientist, eBay Inc., CA Jun 2013 - Aug 2013

Developed machine learnt propensity models for eBay Inc..

- Utilized machine learning algorithms such as logistic regression and stochastic gradient descent to analyze user transaction data at eBay Marketplaces.
- Evaluated the impact of item facts (e.g.: sale type and item condition) and time dependence to propensity modeling.
- Scaled machine learning algorithms for parallel computing.
- *Improved the predictive accuracy over the baseline models by up to 10%*.

Quantitative Analyst, Statistics Bureau of Shenzhen Municipality, China May 2007 - Jul 2007

COMPUTER SKILLS

Extensive experience in R, MATLAB, SAS, Latex, C, SPSS, Minitab, and etc.

Knowledgeable in Python, SQL, and Hadoop/RHIPE.

Familiar with both PC and UNIX operation systems.

RESEARCH AND TEACHING EXPERIENCE

Research Assistant, Purdue University Aug 2008 – June 2013

Develop data-driven predictive methods for massive spatial data:

- Proposed a novel method via Lagrange interpolation to construct low rank models which are more flexible in modeling complex spatial process in a large area.
- Implemented machine learning and data mining algorithms for spatial modeling.
- *Developed parallel computing algorithms using Hadoop/RHIPE for traditional spatial models (e.g., Kriging with compact covariance structure) to accelerate the computation. Provide empirical guide line for data division.*
- Utilized space deformation techniques to visualize and simplify spatial structures of climate patterns and assess their connections to international agricultural trading.
- Participate in grant proposal writing (10%) and preliminary data analysis (50%).

Teaching Assistant, Purdue University Aug 2008 – May 2013

- Appointed as the head TA, recitation instructor, and lecture assistant for the course 'Statistics and Society'.
- Independently instructed the laboratory course *'Introduction to Computational Statistics'*.
- Instructed the laboratory for course ‘Elementary Statistical Methods’.

Teaching Assistant, Washington State University Jan 2007 - May 2008

CONSULTING EXPERIENCE

Statistical Consultant, Purdue University Aug 2010 - Aug 2011

Responsible for statistical tasks with a high level of independence:

- Helped clients to assess research objectives, identify possible statistical issues and propose appropriate models (e.g., generalized linear models, PCA, and Bayesian models) to ensure the validity of statistical methods.
- Actively involved in experimental design, making progress schedule, carrying out pilot studies, and performing power calculation to ensure the quality of statistical modeling.
- Cleaned data, provided software support, and evaluated the results for the clients.
- Assisted clients on report writing and presentation preparation.

ACHIEVEMENTS AND AWARDS

Ross Fellowship, Purdue University Aug 2008 - present

Excellence in Teaching Award, Committee for the Education of Teaching Assistants, Purdue University May 2012

Graduate School Summer Research Grants, Purdue University May 2012

Range Line Pioneer Award, City of Carmel, IN Nov 1992

PRESENTATIONS AND PUBLICATIONS

Presentations

- Teleconnection and Its Impact on Spatial Modeling (Oral presentation). Joint Statistical Meetings, Montréal, Aug 2013
- An Application of Lagrange Interpolation to Spatial-temporal Processes (Poster presentation). Joint Statistical Meetings, San Diego, Aug 2012
- A Non-Parametric Spatial Low Rank Model via Lagrange Polynomials (Oral presentation). Joint Statistical Meetings, Miami, Aug 2011
- Generalized Linear Regression in R: Tricks and Treats (STAT 598HZ guest lecture). Purdue University, Feb 2013
- Spatial Deformation and Space Folding (STAT 598S guest lecture). Purdue University, Oct 2012
- Non-parametric Spatial Low Rank Models (Invited talk). Graduate Student Organization, Purdue University, Apr 2012

Publications

- Mallory L. Chua, **Cheng Liu**, Jorge A. Alfaro-Murillo, Wu Han, Qiming Huang, Kyle D. Bemis, Robert D. Ness, and Xiaosu Tong, R/RHIPE/Hadoop for D&R, Technical Report. (In preparation)
- **Cheng Liu** and Hao Zhang. Non-Parametric Spatial Low Rank Models for Massive Data. (To be submitted)
- **Cheng Liu**, Hao Zhang, and Nelson Villoria, Spatial Covariance Function for Teleconnection. (In preparation)
- Nelson Villoria, **Cheng Liu**, and Hao Zhang, Implications of spatially correlated trade shocks for food price stability. (In preparation)

LEADERSHIP AND SERVICE

Member, Statistics in the Community (StatCom), Purdue University 2009 - present

- Provide free statistical consulting service for local nonprofit organizations.

Graduate student mentor, Purdue University 2009 - 2012

- Help new students to become familiarized with the department and the Purdue campus

Vice president, Chinese Student and Scholar Association, Washington State University 2007 - 2008

- Participated in major events organizing and coordinated resources for the association

President, Student union of the Dept. of Math., HuaZhong University of Science and Technology 2003 - 2004

- Supervised the departments of student union. Made budget and raised funding. Coordinated university activities.

PROFESSIONAL AFFILIATION

Member of the American Statistical Association