

TECHNICAL STRENGTHS

Programming Skills

- R, Python, SAS, SQL, MATLAB, Microsoft Excel.

Expertise

- Machine Learning, A/B testing, R Package Building, Time Series, Experimental Design, Linear (Mixed) Models, Generalized Linear Model, Bayesian Analysis

Research Interest

- Nonparametric and Semiparametric Methods, Data Mining, Spatial Data Analysis

EDUCATION

Iowa State University, Ames, IA

08/2014-07/2019

- Ph.D. Statistics and Applied Mathematics (Co-major), GPA: 3.73/4.0

University of Wisconsin Madison, Madison, WI

08/2010-05/2014

- B.B.A. Mathematics and Actuarial Science, GPA: 3.74/4.0

AWARD

Bancroft Statistics Award of year 2017, Statistical Laboratory, Iowa State University

08/2016-07/2017

An award for the most outstanding Ph.D candidate with a joint major in Statistics

Dean's list of School of Business, UW-Madison

02/2012-11/2013

INDUSTRY WORKING EXPERIENCE

Merck & Co., Inc., Optimization Analytics Intern

06/2018-08/2018

Data Science, Machine Learning, Sequential Mining, and Optimization

North Wales, PA

- Built predictive models and conducted feature importance ranking analysis to find optimized sequences of different promotion channels of products by training machine learning methods based on cross validation error.
- Manipulated promotion and sales data by using R (dplyr, caret, lubridate, and h2o) and SAS.
- Obtained an ensemble model applied to the dataset to discover business impacts and modify business decisions.

PROFESSIONAL EXPERIENCE

R Packages Building Applied in Geospatial Analysis

01/2018-Present

Software, Algorithms

Ames, IA

- "Triangulation", a package creating triangles mesh by using Delaunay triangulation in 2D domain, available on Github, and to be released in CRAN soon.
- "GGAM", a package for Generalized Geospatial Additive Models, to be released in CRAN soon.
- Paper in Progress: Wang G., **Wang J.**, and Wang, L. (2019+), GGAM: An R package for Generalized Geospatial Additive Models.

Variable Selection in Semiparametric Spatial Regression Models (Research Project)

01/2018-Present

R Package Building and Spatial Data Analysis

Ames, IA

- Studied the effects of socioeconomic, safety and spatial factors on U.S. mortality of year 1998-2002 by using a doubly penalized estimator in semiparametric spatial model.
- **Referred Publication:** Wang, G., & **Wang, J.** (2019). On selection of semiparametric spatial regression models. *Stat*, 8(1), e221.

Kaggle Competition on Ames Housing Prices, Obtained Top 13%

02/2017-05/2017

Data mining, Machine Learning and Deep Learning

Ames, IA

- Mainly used R (caret, keras) and Python (scikit-learn) to predict housing Prices and get Kaggle score based on test error.
- Did exploratory data analysis and data visualization (dplyr, ggplot2, seaborn).
- Cleaned datasets by filling missing values using different methods (median, random forest regression and etc.).
- Applied feature engineering by creating new features and converting some categorical variables to numerical variables.
- Tuned different models (Random Forest, Gradient Boosting, Elastic Net, XGBoost and Deep Neural Network (tensorflow in Python)) on the training dataset by using cross validation error.
- Stacked predictions from different models by using different machine learning methods.

Quantile Regression (Research Project)

08/2017-01/2018

Nonparametric Study and Constrained Optimization Problem Solving

Ames, IA

- Quantile regression for spatially varying coefficient models based on Bivariate Penalized Spline over Triangulation (BPST), applied in U.S. county level mortality dataset.
- Constrained optimization problems solved using R and MATLAB.
- Hypothesis testing on non-stationarity of the coefficients.

Data Analysis of Larval of Striped Bass (Class Project)

01/2017-02/2017

Data Analysis and Monte Carlo Study

Ames, IA

- Built the model of survival probability of larval of striped bass.
- Did data exploratory analysis by aggravating parts of data and by data visualization of the correlations between potential covariates (seasonal factors, PH levels of the stream, ambient temperature and so on) and response variable.
- Fitted different models (generalized linear model, beta-binomial model and hierarchical model).
- Assessed models based on whether the models could capture the characteristics of the original data by Monte Carlo simulation.

ACADEMIC WORKING EXPERIENCE

Teaching Assistant, Department of Mathematics, Iowa State University

08/2014-Present

- Led recitation lectures of undergraduate math classes.
- Held office hours for answering questions.

Research Assistant, Department of Statistics, Iowa State University

05/2016-07/2016

- Investigated datasets from Census Bureau available in Research Data Center (RDC), and studied publications based on those data.
- Conducted research to develop materials to support interested groups of ISU researchers of RDC data.

CERTIFICATION

Passed SOA Exams Exam FM, Exam P, Exam MFE and Exam MLC

02/2012-11/2013