Zheye Yuan (Tetsuya Hara)

Pennsylvania State University   ztyh0121@gmail.com or zxy124@psu.edu
Department of Statistics   personal.psu.edu/zxy124
325 Thomas Building, University Park, PA 16802   +1 (814) 883-6639

Education

Pennsylvania State University, University Park, Ph.D., Statistics, expected May 2019,
Advisor: Bing Li, GPA: 3.96/4.00
Hitotsubashi University, Kunitachi, B.A., Economics, 2014,
Advisors: Naoyuki Ishimura, Kenta Kobayashi, GPA: 3.94/4.00
University of California, Berkeley, Exchange, Mathematics, 2013, GPA: 3.814/4.00

Work & Service Experience

Intern, Research & Development, Ford Global Data Insights & Analytics, Summer 2018
Clustered car nameplates using hierarchical clustering and Poincaré embedding.
Intern, Nomura Research Institute, Summer 2013
Analyzed interface & data storage optimization for portable device used by insurance sales person.
Consultant, Statistical Consulting Center, Pennsylvania State University, 2015
Provided statistical consulting for graduate students from other departments.
ASA Datafest Volunteer, Department of Statistics, Pennsylvania State University, 2017, 2018
Dorm Management Volunteer, International Students Dormitory Association of Kodaira, 2012

Awards, Scholarships & Certification

Best Poster Award, Penn State Statistics 50th Anniversary Conference, May 2018
Graduate Assistantship, Department of Statistics, Pennsylvania State University, 2014 to present
Graded Ph.D. level and taught Junior and Senior level courses.
Excellent Academic Achievement, Department of Economics, Hitotsubashi University, 2012
Japan Student Services Organization Scholarship, for exchange to UC Berkeley, 2013
Actuarial Exam P, 11/17/2011, ID: 82124, *Actuarial Exam FM*, 4/9/2012, ID: 66205

Selected Research

Asymptotic Distribution of Neural Network Estimator
Devised methods to find variance of estimators using second derivative information of a neural network.
Reformulated the problem as a likelihood problem to obtain variance from just the first derivative.
Dimension Reduction with Deep Learning
Capitalized on the similarity between minimum average variance estimation (MAVE) and deep learning.
Found that the first matrix of weights, as well as other quantities should reside in the dimension reduction space. Simulation study gave results superior to MAVE.
Nonlinear Support Vector Machine (SVM) on Multiple Functional Data
Each predictor was a vector of functions, for which we demonstrated how to use the reproducing kernel Hilbert space (RKHS) to catch nonlinear features. Both simulation and application gave superior results.
Presented at Joint Statistical Meeting, Vancouver, 2018.
Combining Global and Local Dimension Reduction Methods
MAVE looks at local structure, while contour regression looks at global structure. Attempted to combine the two for robust performance.

Selected Practical Experiences

Poincaré Embedding of Chinese Characters in a Novel
Reduced relationship storage space by embedding a graph of proximity of Chinese characters by embedding them in a Hyperbolic space. This allowed hierarchical structure to be captured more effectively. The algorithm was modified suitably to be implemented in car nameplate classification as well.
Other Miscellany
Semiprametric Copula Estimation, Quantile Regression, Non-stationary Time Series Analysis,
Discrete Cosine Transform Portfolio Construction, Linear & Quadratic Programming,
Numerical Calculation of Derivative Prices, Substitute Charge Method, Finite Element Method,
Reinforcement Learning, Gradient Boosting

Notable Coursework

Statistical Computing, High Dimensional Modelling & Applications, Nonparametric Methods,
Stochastic Processes & Monte Carlo Methods, Categorical Data, Regression Models, Multivariate Analysis

---

Languages

{python} (used in internship and research), {julia, r} (used in research),
{c, c++, matlab} (previously used in research), {LaTeX, html, css, git, UNIX} (used daily),
{sql} (have some experience), {Japanese > English > Chinese > French > Spanish}.
Packages I am familiar with in python include numpy, pandas, tensorflow, pytorch, chainer, keras, gensim, etc.