Hui Lin

✉️ longqiman@gmail.com   happyrabbit   🌐 scientistcafe.com

Computer Skill:

Programming: expertise in R, Python, SQL, CSS, JavaScript, HTML5
Software and system: Linux/Unix, Windows, AWS, Spark

Working and Research Experience

May.2018 - present, Data Science @Netlify, San Francisco, CA, USA

I am building the data science department @Netlify from scratch

Sep.2017 - Apr.2018, Data Scientist, DowDuPont, Johnston, IA, USA

Provide data science leadership for a broad range of analytics in North America Marketing:
1. Predictive Analytics: *Bayesian Regression Model, Super Learner ensembles: Elastic Network Regression, AIC Stepwise Regression, Random Forest, Neural Network, Support Vector Machine, Gradient Boosting Tree, Partial Least Square Regression*
2. Marketing programs analysis under the observational scenario: *Provide statistical training to other analysts on experimental design and propensity score analysis*

Lead the project of building marketing data pipeline:
1. Build marketing database in the cloud
2. Automate models and dashboard that track business performance
    (1) Identify key marketing metrics with the marketing analytics team
    (2) Develop and implement Shiny dashboard including those metrics
3. Social media analytics
    (1) Construct automatic social media data pipeline with data engineer: scrape data from Ag Forum or get data through API, build database in the cloud, develop dashboard for marketing users
    (2) Network Analysis, Natural Language Processing (POS tagging, entity recognition, sentimental analysis, topic modeling)

May.2013 - Aug.2017, Data Scientist, DuPont Pioneer, Johnston, IA, USA

Provide data science leadership for a broad range of analytics in North America Marketing focused on improving Sales and Marketing productivity

Analyze market research survey data using psychometric models and natural language processing; quantitatively study customer perception and identify variables that are predictive of patterns and/or are leading indicators
    Factor analysis, Topic Model, Longitudinal Analysis

Work with data engineers to design and build marketing database in the cloud to support current and future marketing analytics
    It evolves to the current project on marketing data pipeline

Jun.2015 - Present, Co-founder, Central Iowa R User Group

Hui Lin - CV   1/5

Jun.2012 - Apr.2013, Consultant, Statistical Consulting for Business College, Iowa State University

Sep.2009 - Apr.2013, Statistician, Production Animal Disease Risk Assessment Program

Aug.2011 - May.2012, Consultant, General University Statistical Consulting and Statistics Consulting for College of Veterinary Medicine, Iowa State University

Book

Machine Learning using R: By Hui Lin
  Publisher: Publishing House of Electronics industry
  Status: Published in Oct 2017
  ISBN : 9787121326585
Introduction to Data Science: By Hui Lin and Ming Li
  Publisher: CRC Press
  Status: 2019 Dec 31 (Expected)
  Github: https://github.com/happyrabbit/IntroDataScience

Book Translation

Applied Predictive Modeling: By Max Kuhn and Kjell Johnson
  Chinese translator for chapters 1-4, 12-14, 16, 18-20
  Pubisher in China: Huazhang Publishing Inc.
  Status: Published in May 2016
R for Marketing Research and Analytics: By Chris Chapman and Elea McDonnell Feit
  Chinese translator for the book
  Pubisher in China: Huazhang Publishing Inc.
  Status: Published in Oct 2016
Statistical Rethinking - A Bayesian Course with Examples in R and Stan: By Richard McElreath
  Chinese translator for the book
  Pubisher in China: Huazhang Publishing Inc.
  Status: Finished Sep 2018

Journal Publications

Hui Lin (2017.11), How to be a data scientist?, Programmer, China Software Developer Network

Krysta Deitz, Kelly Makielski, Jackie M. Williams, **Hui Lin**, Jo Ann Morrison (2015). Effect of 6-8 weeks of oral ursodeoxycholic acid administration on serum concentrations of fasting and postprandial bile acids and biochemical analytes in healthy dogs,Veterinary Clinical Pathology, DOI: 10.1111/vcp.12275.

Hui Lin (2013), Statistical methods in disease risk analysis, disease testing and nutrition epidemiology, Graduate Theses and Dissertations. Paper 13380.

Holtkamp D.J., **Lin H.**, Wang C, Polson D.D. (2013), Evaluation of an objective risk scoring system for assessing the likelihood of virus introduction in porcine reproductive and respiratory

syndrome virus-free breed-to-wean sow herds in the U.S. Open Journal of Veterinary Medicine. 3 168-175.

Hui Lin, Chong Wang, Peng Liu, Derald J. Holtkamp (2012), Construction of Disease Risk Scoring Systems using Logistic Group Lasso: Application to Porcine Reproductive and Respiratory Syndrome Survey Data, Journal of Applied Statistics. DOI:10.1080/02664763.2012.752449

Derald J. Holtkamp, **Hui Lin**, Chong Wang (2012), Identifying questions in the American Association of Swine Veterinarian's PRRS risk assessment survey that are important for retrospectively classifying swine herds according to whether they reported clinical PRRS outbreaks in the previous 3 years., Preventive Veterinary Medicine 106(2012) 42-52

Allen, Kelli; Coetzee, Johann; Edwards-Callaway, Lily; Glynn, Hayley; Dockweiler, Jenna; KuKanich, Butch; **Lin, Hui**; Wang, Chong; Fraccaro, Elena; Jones, Meredyth; Bergamasco, Luciana (2012), The impact and effect of timing of oral meloxicam administration on physiological responses in calves after dehorning with local anesthesia, Journal of Dairy Science, DOI: http://dx.doi.org/10.3168/jds.2012-6251

Hui Lin, Cheng Hao, Xiao Wang, Qiang Yuan, Tong-Jie Zhang, and Bao-Quan Wang (2008), Observational H(z) Data as a complementary to other cosmological probes, Modern Physics Letter A (No. MPLA-D-08-00168R1)

Certifications

Convolutional Neural Networks by deeplearning.ai on Coursera. Certificate earned on January 5, 2018

Social and Economic Networks: Models and Analysis by Stanford University on Coursera. Certificate earned on January 3, 2018

Programming Foundations with JavaScript, HTML and CSS by Duke University on Coursera. Certificate earned on November 18, 2017

Structuring Machine Learning Projects by deeplearning.ai on Coursera. Certificate earned on October 20, 2017

Improving Deep Neural Networks: Hyperparameter tuning, Regularization and Optimization by deeplearning.ai on Coursera. Certificate earned on October 10, 2017

Neural Networks and Deep Learning by deeplearning.ai on Coursera. Certificate earned on September 26, 2017

Pioneer Young Professionals Leadership Exploration and Development Program

DuPont Pioneer
January 2016 – December 2016

Tutorials and Invited Talks

Big Data, Data Science and Deep Learning for Statistician

2019-02-14, 2019 ASA Conference on Statistical Practice, New Orleans, LA
2019-03-24, ENAR 2019 Spring Meeting, Philadelphia

2019-04-25, Innovations in Design, Analysis, and Dissemination: Frontiers in Biostatistical Methods, Kansas City

2019-05-18, ASA Council of Chapters Traveling Course [Snake River Chapter]

Bridging Statistics and Data Science, 2018-10-03, Fall Technical Conference, West Palm Beach, FL

Preparing Statistician to be Successful Big Data Scientist, 2018-06-14, 2018 ICSA Applied Statistics Symposium, New Brunswick, NJ

Bridging Statistics and Data Science, 2018-06-11, Joint Research Conference, Santa Fe, NM

Marketing Data Science, 2018-3-23, The University of Iowa, Slides

How to be a data scientist, 2017-12-27, DATAGURU, Slides

Introduction to Data Science, 2017-11-29, American Statistical Association, Slides

Preparing Statistician/Statistics Graduates to Be Data Scientist, 2017-07-30, Joint Statistical Meeting 2017, Course Website

Interactive document and dashboard using R and shiny, 2017-07-12, Central Iowa R User Group, Blog

Data Scientist Using R - Data Wrangling, 2017-05-02, Central Iowa R User Group, Blog

Data Science Introduction, 2017-03-29, Central Iowa R User Group, Slides

Unlock Unstructured Data, 2016-12-29, Beijing Institute of Technology, Slides, Rcode

Market Research Customer Perception Analytics, 2016-7-21, Video, Slides, Rcode

Data Wrangling in R and Model Technique, 2016-6-8, Capital of Statistics(COS)-North America, Video, Slides, Rcode1, Rcode2

General Process of Data Analysis and Data Preprocessing, 2016-6-2, Capital of Statistics(COS) - North America, Video, Slides, Rcode

Jewels Academy Math Camp, 2016.2.20, Jewels Academy, Des Moines, Iowa

What is data scientist, 2015.12.12, Capital of Statistics(COS), Beijing, CHINA, Video, Slides

Data Pre-processing Using R, 2015.11, Central Iowa R User Group, Johnston, Iowa, USA

Analytics to Support Business Decision in Traditional Industry, 2014, 7th R conference, Beijing, CHINA

Services and Professional Membership

Program Committee: Annual Data Institute Conference 2019
2018 Program Chair: ASA Section on Statistics in Marketing
2017 Program Chair-Elect: ASA Section on Statistics in Marketing
2016-2017, Web developer and maintainer,Johnston Alternative Gift Market
2015-Present, ASA Section on Statistics in Marketing
2012-Present, Capital of Statitics, administrator

2012-Present, Referee, Journal of Applied Statistics
2011-Present, American Statistical Association (ASA), member
May 24-25, 2014, Committee member, 7th R conference, Beijing, CHINA
2013-2014, Chinese teacher, Iowa Chinese Language School (ICLS), Johnston, Iowa
    ICLS is a non-profit and volunteer-based organization established on July 17, 1999.

Education

2011 - 2013, PhD in Statistics, Department of Statistics, Iowa State University, USA

2009 - 2011, MS in Statistics, Department of Statistics, Iowa State University, USA

2005 - 2009, BS in Statistics, School of Mathematical Science, Beijing Normal University, CHINA