YIDONG ZHOU

400 Russell Park, Davis, CA 95616

530-302-6078 ⋄ ydzhou@ucdavis.edu ⋄ https://yidongzhou.github.io/index.html

EDUCATION

University of California, Davis — Davis, USA
Ph.D. Candidate in Statistics — *Sep 2019 - Present*

University of Science and Technology of China — Hefei, China
B.S. in Statistics — *Aug 2015 - Jul 2019*

- 2016 Excellent Student Scholarship (2 out of 85)
- 2018 Excellent Student Scholarship (2 out of 83)
- Relevant Coursework: Bayesian Analysis (**99**), Regression Analysis (**90**), Complex Analysis (**90**), Real Analysis (**87**), Functional Analysis (**86**), Partial Differential Equations (**86**), Applied Stochastic Processes (**95**), Probability (**86**), Operations Research (**97**), Mathematical Statistics (**92**), Algebraic Structure (**91**), Applied Statistical Software (R & Python) (**95**), Data Structures and Database (C & SQL) (**91**), Computer Programming A (C programming language) (**88**), Introduce to Data Science (**A+**)

Harbin Institute of Technology — Harbin, China
International Summer School — *Jul 2017 - Aug 2017*

RESEARCH EXPERIENCE

Impact of Network Features on Correlation Structure — Cambridge, MA
Research Assistant, supervised by Prof. Rui Wang, Harvard University — *Jan 2019 - May 2019*

- Implemented **network modeling** and propagation epidemic processes
- Derived the relationship between core network features (various centralities including degree, eigenvector, PageRank, closeness, and betweenness; density) and correlation structure
- Proposed the definition of **hotspots** in contact networks and further implemented the prediction of hotspots

Interim Analyses Using Multivariate Repeated Confidence Bands (MRCBs) — Cambridge, MA
Research Assistant, supervised by Prof. Rui Wang, Harvard University — *Sep 2018 - Dec 2018*

- Implemented multivariate repeated confidence bands which can accommodate multivariate response and multiple treatment arms under three settings (repeated measures responses, recurrent events, survival data)
- Derived the mean and covariance function of treatment difference and further proved the correctness of the result
- Utilized **resampling algorithm** to approximate critical values and thus computed the border of confidence bands
- Proposed an algorithm which can place **confidence bands** and confidence regions adaptively to visualize MRCBs

R Package Development for RTREE, STREE, and MASAL Model — New Haven, CT
Research Assistant, supervised by Prof. Heping Zhang, Yale University — *Jun 2018 - Sep 2018*

- Constructed the macs package with **Rcpp**, including package architecture and late-stage testing & refinement
- Integrated **C/C++** into **R** to implement Classification Trees (RTREE), Survival Analysis Trees (STREE) and Multivariate Adaptive Splines for Analysis of Longitudinal Data (MASAL) model
- Programmed powerful functions in R, such as R interface function, plot function, print function, etc.
- Proposed an **adaptive tree drawing algorithm** which can calculate coordinates by recursive computation according to node distribution, and then arrange the position of each node in the plot

Data Analysis for Inventory Management — Hefei, China
Research Assistant, supervised by Prof. Yugang Yu, Univ. of Sci. & Tech. of China — *Dec 2017 - May 2018*

- Conducted case-driven survey of cross-border e-commerce to develop inventory management strategy

· Performed substantial cleaning of in-situ data from a top Amazon seller and fitted data in R and **Python**
· Used machine learning models to improve stock tracking accuracy, anticipate sales and customer traffic to forecast stock requirements

Text emotion analysis based on theme Hefei, China
Research Assistant, Supervised by Prof. Qi Liu, Univ. of Sci. & Tech. of China *Sept 2017 - Dec 2017*

· Preprocessed Taobao users review data with the Chinese text segmentation package (**jieba**) in Python
· Performed part-of-speech tagging, removed stop words and extracted text features based on the theme (price, express, etc.)
· Adopted the K-Nearest Neighbors (**KNN**) algorithm to classify the texts

MANUSCRIPT IN PREPARATION

Zhang, H., **Zhou, Y.**, Ma, M., Lu, Y., & Ma, H. (2018+). "Recursive Partitioning Based Multivariate Adaptive Regression Models, Classification Trees, and Survival Trees: The macs Package for **R**".

TEACHING

Computer Programming A (C Programming Language) Hefei, China
Teaching Assistant, University of Science and Technology of China *Sept 2017 - Jan 2018*

· Lectured over **90** UTSC students on Q&A sessions of Computer Programming A for four hours per week
· Managed lab guidance, graded student assignments and exams

LEADERSHIP

Student Union, Minister of Activity Hefei, China
Minister of Activity *Sept 2016 - Jun 2017*

· Led a team of **15** to draft policy memos, coordinated discussions, and finalize policy enactment
· Led the development and launch of a performance assessment system to improve operational excellence
· Chaired monthly meetings to collect feedback from student representatives to enhance the interaction between faculty and students

Freshman Seminar Hefei, China
Research Team Leader *Sept 2015 - Jun 2016*

· Summarized literature related to research topics, established a mathematical model about predicting global climate
· Completed the essay and **oral defense** with teammates, got an A with the joint efforts of team members

SKILLS AND OTHERS

Statistics & Computer

· Naive Bayesian Model, Decision Tree Model, Logistic Regression Model, Convolutional Neural Network, etc.
· C/C++, R, Python, LaTeX, HTML, CSS, JavaScript, MATLAB, Microsoft Office, PS, AE

Activities

· CCF Big Data & Computing Intelligence Contest (50 out of 751)
· National College Students Mathematical Modeling Competition (top 10%)
· Big Data Modeling Contest (top 10%)
· "Find the next madma" Competition (top 10%)