陳 慶全

Curriculum Vitae

"Follow your own course, and let people talk." – Alghieri Dante

Education

2012.09–2014.06 **Master of Statistics**, *National Cheng Kung University (國立成功大學)*, Tainan, *GPA – 4.00*. Supervisor: Prof. Ray-Bing Chen with thesis title:*A Classification Approach Based on Density Ratio Estimation with Subspace Projection*.

2008.09–2012.06 **Bachelor of Economics and Statistics (double major)**, *National Cheng Kung University*, Tainan, *GPA – 3.50*.

Experience

Vocational

2016.07–Now **CIM Engineer**, Taiwan Semiconductor Manufacturing Company Limited, Taichung, Taiwan.

- Design a big data solution for semiconductor manufacturing. It covers the several fields, setting up highly distribued file system with data warehouse storing data in wide-column format, distribued computing environment and query server for back-end data analysis.

2015.09–2016.06 **Research Assistant for Research Fellow Jeng-Min Chiou**, Institute of Statistical Science, Academia Sinica, Taipei, Taiwan.

- Download open data from websites with R.
- Development of the data analyzing procedure with MatLab.
- Construction of the statistical method with MatLab and R.
- Data analysis about estimating travel time and predicting flow rate with functional PCA for highway data.

2014.10–2015.09 **Substitute Military Service**, National Immigration Agency (移民署), Taichung, Taiwan.

2013.10–2014.08 **Research Assistant for Prof. Ray-Bing Chen**, National Cheng Kung University, Tainan, Taiwan.

- Organization of journals about estimation of importance and dimension reduction.
- Construction of the statistical model with MatLab.

2012.06–2013.09 **Research Assistant for Prof. Yun-Chan Chi**, National Cheng Kung University, Tainan, Taiwan.

- Organization of journals about regression model and GEE for bivariate zero-inflated Poisson distribution.
- Construction of the statistical method with R.
- Data analysis about factor of accidents at railroad crossing - Taking Taiwan railways for example.

32452 桃園市平鎮區 – 廣仁里 11 鄰廣豐街 114 號
0961 131 126 • (03) 495 1040 • zw12356@gmail.com
chingchuan-chen.github.io

Miscellaneous

2015.08–Now **Moderator**, *Statistics*, PTT.cc.
Manage day-to-day affairs of the board Statistics. For instance, gathering and classifying the useful information and questions into the best posts for users.

2013.08–Now **Moderator**, *R_Language*, PTT.cc.
Manage day-to-day affairs of the board R_Language. For example, gathering and classifying the useful information and questions into the best posts for users and helping users to solve their problems in R programming.

Awards

2014 2014 Competition for Data Analysis with R in Taiwan (2014 年 R 資料分析競賽) – Honorable Mention

2014 Cathay Charity Foundation Scholarship (蔡萬霖先生紀念獎學金)

Communication Skills

2010 Oral Presentation at the 23rd South Taiwan Statistics Conference

Languages

Chinese **Mothertongue**

English **Intermediate**

Familiar Statistical Methods or Machine Learnings

Linear Model Multivariate Linear Model, Generalized Linear Model (GLM) like Logistic Regression, Linear Mixed Effect Model (LMEM), Regression Model with LASSO Penalty, Ridge Regression, Generalized Estimation Equation (GEE)

Machine Learning Support Vector Machine (SVM), Kernel SVM, Generalized Additive Model (GAM), Multivariate Adaptive Regression Splines (MARS), Tree, Boosting Trees, Random Forest

Functional Data Analysis Functional Principal Component Analysis, Multivariate Functional Principal Component Analysis, Functional Linear Response Model, Functional Clustering

Miscellaneous Methods of Dimension Reduction like PCA and SIR, Clustering Methods like k-means and Fisher Discriminant Analysis, Smoothing Techniques like Kernal Smoothing Methods, Locally Weighted Regression (LOWESS)

Computer skills

Basic SQL, HTML, CSS, JAVA, PERL

Intermediate MONGODB, C/C++, , shell script, HADOOP, SPARK, SCALA, PYTHON, LaTeX, Linux, Microsoft Windows, Microsoft Office

32452 桃園市平鎮區 – 廣仁里 11 鄰廣豐街 114 號
0961 131 126 • (03) 495 1040 • zw12356@gmail.com
chingchuan-chen.github.io

Advanced

MatLab:

- Familiar with vectorizing programming skills and parallel computing skills.
- Using MEX to call C/C++, CUDA.
- Data manipulation and data visualization.

R:

- Familiar with vectorizing programming skills and naive R parallel computation with parallel, snow and foreach.
- Data cleaning and summarization with plyr, data.table, dplyr, tidyr and etc.
- Data visualization with basic plot commands, lattice, ggplot2, plotly and shiny. The graphs, interactive graphs and web UI are all covered.
- Link with other programming languages (C/C++, Java, MatLab).

Selected Programming works

R packages

rfda:

- PACE is a MatLab package which is based on the principal analysis by conditional estimation (PACE) algorithm to perform the functional principal component analysis (FPCA). It is useful to analyze the random trajectories, stochastic processes and longitudinal data.
- PACE focus on univariate functional data, but rfda will be able to analysis multivariate functional data.
- This package is under construction. It provides R functions to use PACE and a faster implementation with C++ via Rcpp. With the power of C++, it is at least 10 times faster than the MatLab package in the smoothing part.
- This is open at https://github.com/ChingChuan-Chen/rfda.

milr:

- This package performs maximum likelihood estimation for multiple-instance logistic regression utilizing EM algorithm with LASSO penalty.
- This is open at https://github.com/ChingChuan-Chen/milr.

32452 桃園市平鎮區 – 廣仁里 11 鄰廣豐街 114 號
0961 131 126 • ☎ (03) 495 1040 • ✉ zw12356@gmail.com
chingchuan-chen.github.io

R works Introduction to the R packages about data analysis: It was originally published at ptt.cc and republished at my blogger.

- magrittr (http://chingchuan-chen.github.io/r/2015/07/01/magrittr/): introduce the pipe operators in magrittr.
- data.table (http://chingchuan-chen.github.io/r/2015/07/02/data-table/): introduce the class data.table and relative usages.
- dplyr (http://chingchuan-chen.github.io/r/2015/07/03/dplyr/): introduce the basic methods to manipulate data.frame and summarize those methods.
- tidyr (http://chingchuan-chen.github.io/r/2015/07/04/tidyr/): introduce the techniques to tidy up tables.

Shiny app for demonstrating the Ising model:

- It was originally published at ptt.cc and republished at my blogger. (The blogger link: http://chingchuan-chen.github.io/r/2015/04/20/shiny-app-for-ising-model/) It is based on the R package shiny to develop. This app provides the simple simulation about the change of ferromagnetism. The Metropolis-Hasting algorithm and Gibbs sampling algorithm are provided. Also, the relative parameter is tunable. Figure 1 is the presentation of work.

Shiny app for quality tool:

- It is a private case. I provide some advises such as the usage of web app built on shiny, the interactive graph with rChart and table presentation. I also support some programming problem about R and javascript. Figure 2 is the presentation of work.

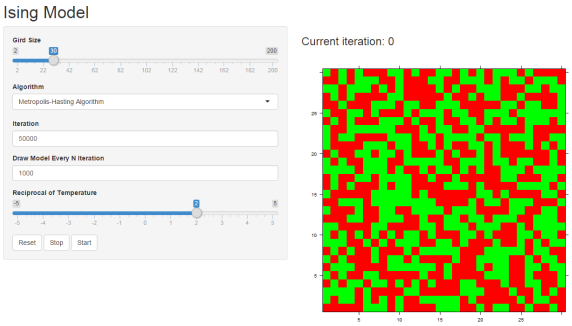

Figure 1: The shiny app for the Ising model

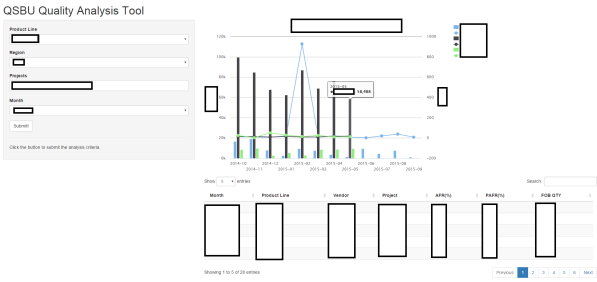

Figure 2: The shiny app for the quality tool

32452 桃園市平鎮區 – 廣仁里 11 鄰廣豐街 114 號
0961 131 126 • ☎ (03) 495 1040 • ✉ zw12356@gmail.com
chingchuan-chen.github.io