

"The superior man makes the difficulty to be overcome his first interest; success only comes later." – William Jennings Bryan, Politician

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.
- Construct a big data solution for the data coming from semiconductor manufacturing system alone. It covers the several fields which includes to design data schema to store data, use Apache Hive to store massive and messy data, security and authorization for Apache Hive, using Apache Spark to synchronize Oracle database and Apache Hive and securing the Apache Hive with LDAP authorizations etc.
 - Build up R working environment, setup several RStudio servers for developers, build up a mini CRAN for those servers without network to get R packages installed, packing common function as several R packages like auto-install ROracle, Oracle SQL automatic build-up, graphic functions and Hive connector etc.
 - Analyze the semiconductor manufacturing data like offline / inline measurement, tool sensors or yield related issues.
 - Open R workshop to teach colleagues to use R more professionally and efficiently, introduce new R packages, use R to perform machine learning and data analysis.
 - Construct R web service for other colleagues easily using R to plot, perform machine learning without knowing R or machine learning knowledge.
 - Construct local git server powered by GitLab CE for the management of code, project in department.
 - Reaches an accuracy of 99.98% on identifying the wafer with defects by deep learning (residual net).
 - Identifies the bad tools or key process with statistical model.
 - Using statistical method to find out an optimal control limit to avoid the out of control of tools.

- 2015.09–2016.06 **Research Assistant for Research Fellow Jeng-Min Chiou**, INSTITUTE OF STATISTICAL SCIENCE, ACADEMIA SINICA, Taipei, Taiwan.
- A crawler to automatically download open data from websites with R.
 - The 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.
 - Construct interactive data visualization for the highway data with R shiny.
- 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.

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.

Journals

- 2017 Ping-Yang Chen, Ching-Chuan Chen, Chun-Hao Yang, Sheng-Mao Chang and Kuo-Jung Lee, milr: Multiple-Instance Logistic Regression with Lasso Penalty, The R Journal, 2017-013, <https://journal.r-project.org/archive/2017/RJ-2017-013/index.html>

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

Familiar Statistical Methods or Machine Learnings

- Linear Model Regression Model with L1/L2 Penalty, Generalized Linear Model (GLM) like Logistic Regression, Linear Mixed Effect Model (LMEM), 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, Neural Network for Image classification
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 Kernel Smoothing Methods, Locally Weighted Regression (LOWESS)

Computer skills

Basic Knowing	HTML, CSS, JAVASCRIPT
Intermediate	ORACLE SQL, MONGODB, C/C++, SCALA, shell script, SPARK, Linux
Advanced	MATLAB: <ul style="list-style-type: none"> Familiar with vectorizing programming skills and parallel computing skills. Using MEX to call C/C++, CUDA. Data manipulation and data visualization. R: <ul style="list-style-type: none"> Familiar with vectorizing programming skills and naive R parallel computation with PARALLEL and FOREACH. Data cleaning and summarization with PLYR, DATA.TABLE and etc. Data visualization with basic plot commands, LATTICE, GGLOT2, PLOTLY and SHINY. The graphs, interactive graphs and web UI are all covered. Modeling with R packages for those methods mentioned above. Link with other programming languages (C/C++, JAVA) for improving performance. Pack functions into R package for re-usage.

Selected Programming works

R packages	MILR: <ul style="list-style-type: none"> The package is now on CRAN. This package performs maximum likelihood estimation for multiple-instance logistic regression utilizing EM algorithm with LASSO penalty. The project url: https://github.com/ChingChuan-Chen/milr. RCPPBLAZE: <ul style="list-style-type: none"> The package is now on CRAN. The RcppBlaze package includes the header files from the Blaze library with disabling some functionalities related to link to the thread and system libraries which make RcppBlaze be a header-only library. Therefore, users do not need to install Blaze and the dependency Boost. This project url: https://github.com/ChingChuan-Chen/RcppBlaze.
------------	---

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

- Introduce to magrittr: <http://chingchuan-chen.github.io/r/2015/07/01/magrittr/>.
- Introduce to pipeR: <http://chingchuan-chen.github.io/posts/2016/07/10/pipe-operators-in-R>.
- Introduce to data.table: <http://chingchuan-chen.github.io/posts/2015/07/02/data-table>.
- Introduce to dplyr: <http://chingchuan-chen.github.io/posts/2015/07/03/dplyr>.
- Introduce to tidyr: <http://chingchuan-chen.github.io/posts/2015/07/04/tidyr>.
- Introduce to foreach and iterators: <http://chingchuan-chen.github.io/posts/2017/04/23/introduce-to-foreach-and-iterators>.

SHINY APPLICATION 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 application 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 APPLICATION FOR QUALITY TOOL:

- It is a private case. I provide some advises such as the usage of web application 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.

Ising Model

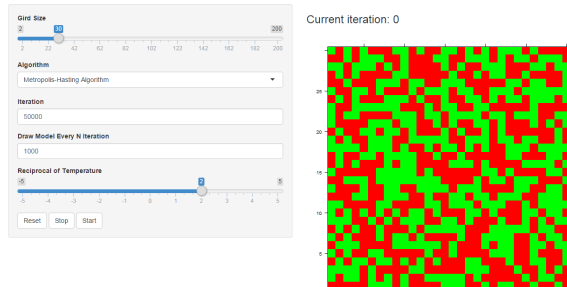


Figure 1: The shiny application for the Ising model tool

QSB Quality Analysis Tool



Figure 2: The shiny application for the quality