Jamal Ching-Chuan Chen
陳慶全

Data Engineer / Data Analyst / R

CONTACT

No. 114, Guangfeng St., Pingzhen Dist.
Taoyuan City, 32452 Taiwan
(private) +886-966-676-326
(work) +886-963-855-707

zw12356@gmail.com

www.facebook.com/celestial0230

www.linkedin.com/in/celestial0230

github.com/ChingChuan-Chen

EDUCATION

2012.09 – 2014.09 **National Cheng Kung University, Tainan, TW**

Master

GPA: 4.0

Thesis:

A Classification Approach Based on Density Ratio Estimation with Subspace Projection

Advisor:

Ray-Bing Chen

Abstract:

For imbalanced data, the density ratio estimation (Kanamori et al. (2009)) is good solution to solve it. However, the performance of density ratio is poor when data is sparse in the high dimension. Therefore, we propose using projection to perform dimension reduction. Our result shows that the proposed method is better than the original method.

2008.09 – 2012.06 **National Cheng Kung University, Tainan, TW**

Bachelor

GPA: 3.5

LANGUAGES

ABOUT

My name is Jamal Chen and I am a data engineer and data analyst with 3+ years of experience in big data infrastructure, data preprocessing and modeling. I am an experienced R programmer in data preprocessing and modeling, also a experienced Linux maintainer in automated process and system service management. I am familiar with packaging codes and reusing for efficient and fast development of applications. Also, I can give insights from data and provides pricture for making decisions.

PROJECTS

Automatically Generated Resume

https://github.com/ChingChuan-Chen/python-yaml-resume

A tool for automatically generated resume written in Python by YAML and Jinja2.

Highlights

- Easily maintain resume by modifying the YAML file.
- Simply changing Jinja template for different themes.

R package RcppBlaze

https://github.com/ChingChuan-Chen/RcppBlaze

Blaze is an open-source, high-performance C++ math library for dense and sparse arithmetic. This package provides the header files for linking Blaze library in Rcpp.

Highlights

- Full API from R to Blaze under the RcppArmadillo-like framework.

R package milr

https://github.com/PingYangChen/milr

This package performs maximum likelihood estimation for multiple-instance logistic regression utilizing EM algorithm with LASSO penalty.

Highlights

- A first R package address the analysis of the multiple instance data.
- This package provides a MLE with EM algorithm under the framework of logistic regression.
- Providing not only prediction, but also variable selection with L1 panalty.
- The performance issues are addressed by using RcppArmadillo.

AWARDS

December 2017 **TSMC Kaggle Competition for the Defect Recognition**

Third Place

A internal competition in TSMC. Its purpose is to make classification of defects able to judge automatically by machine for lessening human cost. They provides 3000 pictures of 4 types of defects and let employees fit a deep learning model to classify. Then send the model to the platform for get the accuracy rate of testing set (1200 pictures.).

Chinese	Native speaker
English	Conversant
Japanese	Basic Knowledge

REFERENCES

Ray-Bing Chen
Professor
Department of Statistics
National Cheng Kung University
+886-6-275-7575 ext. 53645
rbchen@mail.ncku.edu.tw

Sheng-Mao Chang
Associate Professor
Department of Statistics
National Cheng Kung University
+886-6-275-7575 ext. 53632
smchang@mail.ncku.edu.tw

Jeng-Min Chiou
Research Fellow
Institute of Statistical Science
Academia Sinica
+886-2-2783-5611 ext 312
jmchiou@stat.sinica.edu.tw

August 2014

Competition for Data Analysis with R in Taiwan

Honorable Mention

A national competition in Taiwan. Its purpose is to let participants find their own topic in given data and try to explain by data. The whole analysis need to be done by R. The data is collected from a registering system created by Taiwan governmnet of the actual selling price of real estate. Our team chose to predict the price of house from a messy data. Each team had the times of a day to finish their report. We used half a day to clean data and visualize the data. Other half a day is used in modeling and writing report.

JOURNALS

milr: Multiple-Instance Logistic Regression with Lasso Penalty

Ping-Yang Chen, Ching-Chuan Chen, Chun-Hao Yang, Sheng-Mao Chang and Kuo-Jung Lee
The R Journal (2017) 9 :1 , pages 446-457 .
https://journal.r-project.org/archive/2017/RJ-2017-013/index.html