Chia-Hui (Anita) Shen

Department of Statistics, University of California, Davis 95616
chshen@ucdavis.edu · linkedin.com/in/chia-hui-shen

Master of Science in Statistics at University of California Davis. Seeking **full-time data analytics related position**. Experienced in **genetic datasets, business strategies, social science research** and other forms of data analysis with over 5 years of experience in R programming and familiarity with Python and Unix.

EDUCATION

M.S. Statistics	University of California, Davis	GPA3.71/4.0	12/2017
Exchange Student	Engineering Science, Osaka University (OU), Japan	GPA3.60/4.0	09/2015 - 08/2016
B.B.A. Statistics	National Cheng Kung University (NCKU), Taiwan	GPA3.75/4.0	06/2016
B.S. Mathematics	National Cheng Kung University (NCKU), Taiwan	GPA3.50/4.0	06/2016

WORK EXPERIENCE

Bioinformatics Intern, Genomic Health, Redwood City, CA — 06/2017-09/2017
- Designed panel for tumor burden tracking by exploring different gene regions, cost functions and patients
- Developed optimization of algorithms to track tumor burden mutation with up to 90% patient coverage in less than 50,000 base pair panel size
- Improved cost efficiency by more than 50% for the optimization of algorithms in UNIX

Surveying Specialist, Survey and Statistics Research Center, NCKU — 09/2012 - 06/2014
- Assisted investigations with pool surveys by collaborating and consulting with industries

RESEARCH EXPERIENCE

Spam Email Detection, Department of Statistics, UC Davis — 04/2017 - 06/2017
- Filtered important information by collecting and cleaning over 100k text emails
- Applied Natural Language Processing, including tf-idf and Latent Semantic Analysis on feature engineering
- Compared and applied 6 machine learning methods including logistic regression, kNN, and others on classifications, improving accuracy to 97%

Housing Decision Making Criteria in California, Department of Statistics, UC Davis — 09/2016 - 12/2016
- Applied the Zillow housing dataset on making comparisons for every county in California
- Predicted risk and values to decide affordable housing prices and regions for customers
- Designed a front-end webpage to visualize the Zillow housing dataset and statistical results

Causal Inference in Boston Housing Data, SHIMODAIRA Laboratory, OU, Japan — 09/2015 - 08/2016
- Simulated the estimation of correlation for the causal effect in grouped variables
- Predicted strengths of relations between features by devising housing strategies for clients

Quality of Genetic Sequencer-MiSeq, Department of Statistics, NCKU, Taiwan — 09/2014 - 02/2015
- Pre-processed data of 100,000 genetic images to obtain matrix dataset and to visualize genetic images
- Predicted underlying factors by using multivariate analysis in improving the quality of sequencers

SKILL

- Experienced data analysis by using R for over 5 years
- Programming Languages: Python, SAS (Based Programmer 9.0), MATLAB
- Tools: SQL, Unix, Linux, Git, GGplot, Plotly, HTML, Excel, Word, PowerPoint

LEADERSHIP

Director of General Affairs, Student Association of the Department of Statistics, NCKU
- Reported and managed the yearly financials and reduced at least 10% of expenses

Leader of Communication Sections, Night of Stat, Student Campaign, NCKU
- Improved profitability by collaborating with other departments in designing marketing strategies