

SHUDONG SUN

▪ 213-709-0926 ▪ shudongsun@email.arizona.com ▪ Office: ENR2 S370BB UA, Tucson, AZ
▪ <https://shudongsun.github.io/>

EDUCATION

The University of Arizona

Ph.D. in Statistics and Data Science, GPA: 4.0/4.0

Tucson, AZ

Sep 2021 – Present

University of Southern California

M.S. in Computer Science, GPA: 3.96/4.0

Los Angeles, CA

Sep 2019 – Jun 2021

Beihang University

B.S in Electronic Information Engineering, GPA: 90.78/100, Top 5%

Beijing, China

Sep 2015 – Jun 2019

Honors: China National Scholarship (2018), Excellent Graduates in Beijing (2019), Grand Prize of Academic Excellence Award (2018), Excellent Student Cadre (2018), Merit Student Award (2016)

RESEARCH EXPERIENCE

Sample Size Determination in Pilot Data Study web

Los Angeles, CA

Supervisor: Prof. Xin Tong, Prof. Jingyi Li

Jun 2020 – June 2021

- Develop statistical models to generate training data according to small size pilot data
- Construct classifiers based on basic classification methods such as logistic regression, random forest, support vector machine, and XgBoost using generated data
- Aim to determine sample size according to classification results

Construction of Knowledge Graph for Specific Domains

Beijing, China

Supervisor: Dr. Xiaoming Zhang

Oct 2018 – Jun 2019

- Employed supervised learning method to identified knowledge entities and relationships amongst them from research paper abstracts and keywords
- Introduced author information as the entity attribute to construct the knowledge graph; realized visualization via Neo4j graph database platform
- Performed data fusion and entity disambiguation to enhance graph accuracy and quality

Application of Neyman-Pearson Classification Algorithms in Facial Recognition

Los Angeles, CA

Supervisor: Prof. Xin Tong

Jul 2018 – Sep 2018

- Set up facial recognition database by collecting online images and installing deep learning frameworks, such as Tensorflow and Keras, on Linux system
- Classified images by using R programming language to train convolution neural networks
- Optimized the umbrella algorithm to process classification results and to further control type I errors

INTERNSHIPS

Business Big Data

Chengdu, China

Mathematical Model Building Intern

Jun 2019 – Aug 2019

- Used SQL to access data and implemented data cleaning for local companies
- Engaged in teamwork to build statistical models through Python; examined and encapsulated programs

Chinese Academy of Sciences Institute of Automation

Beijing, China

Acoustic Model Training Intern

Jul 2017 – Aug 2017

- Extracted two audio spectrum features (PLP and FBank) according to given speech signals for neural network training and related models
- Adopted Gaussian mixture model, deep neural network and long short-term memory to construct acoustic models; applied these models to speech recognition systems and analyze recognition error rates

TEACHING EXPERIENCE

Teaching Assistant at University of Southern California

Los Angeles, CA

TA of Applied Modern Statistical Learning Methods (DSO530)

Jan 2020 – Jun 2020

- Established eight informative Python tutorials that served as the course foundation, including methods of feature engineering, implementation of resampling techniques, subset selection, shrinkage methods, and methods of supervised learning and unsupervised learning
- Graded all homework and midterm exam; fostered interactive and supportive environment for students during class and in office hours
- Assisted the professor to create the course project for students

PROJECTS

Classification on Breast Cancer Data Set

Jul 2020 – Aug 2020

- Preprocessed the Breast Cancer Wisconsin (Diagnostic) Data Set
- Implemented Random Monte-Carlo Simulation using supervised (k-means algorithms as basis), semi-supervised (L1-penalized SVM as basis) and unsupervised learning technique (L1-penalized SVM as basis) to train models on the processed data separately

Research on Google Matrix and PageRank Algorithm Simulation

May 2018 – Jun 2018

- Explored PageRank algorithm Google used, especially the relationship between Google matrix in the algorithm and Markov process in the random process theory
- Simulated PageRank algorithm on small networks and created a visual analysis with MATLAB

Digital Image Processing and Machine Learning

May 2018 – Jun 2018

- Accomplished image segmentation via inter-class variance thresholding algorithm and implemented morphological image processing
- Completed edge detection using three kinds of operators (Robert, Sobel and Laplacian); extracted parameters based on Hough transformation for image reconstruction
- Developed a learning algorithm for neural networks to recognize the handwritten numbers in the database

COMPETITIONS AND AWARDS

Mathematical Contest in Modeling - Meritorious Winner

Feb 2018 – Mar 2018

Multi-Hop Model for HF Radio Waves Propagation with Calculation of Radio Strength Distribution

- Constructed a fundamental model via combining ocean surface models (calm and turbulent) to determine the strength distribution and angles of reflected rays off the screen
- Proposed a multi-hop model to simulate the propagation of the HF radio waves between land and ocean based on the reflections off the land, ocean, and ionosphere
- Tested the model's validity and sensitivity through altering the amplitude, frequency of dropping stones and the incident angles

China Undergrad Mathematical Contest in Modeling – 1st Prize Beijing Region

Sep 2017 – Oct 2017

Parameter Calibration and Imaging of Computed Tomography (CT) System

- Determined a CT system's parameters based on the medium with known absorptivity as the calibration template
- Calculated and filtered the values reflecting relative absorptivity of each point to reconstruct CT images; verified the result accuracy through curve fitting
- Established the unified quantification standard of absorptivity to help compare

SKILLS

- Programming Language: Python, R, C, C++, JAVA, MATLAB, SQL