Wei-Ying Wang, Ph.D.

wei-ying_wang@brown.edu wayinone.github.io +1-401-556-9066
Providence, RI (willing to relocate)

SUMMARY

- Seeking a statistician/data scientist position in which I would contribute to the success of a business
- Applied Mathematics Ph.D. with 10+ years experience in programming and statistical modeling
- Analyzing large dataset to develop an image compression algorithm with the optimal compression rate

TECHNICAL SKILLS

Statistics	Mathematical statistics, regression, Baysian analysis, MCMC, information theory
Machine Learning	Classification, decision tree, random forest, deep learning, SVM, gradient boost
Image analysis	Compression, denoising, 3D reconstruction, pattern recognition
Programming	Proficient with: Python(numpy, keras, sklearn, pandas), Matlab, Latex
	Experienced with: SQL(MySQL), C/C++, Amazon EC2, Mathematica
Operating system	Windows, Linux(light)

EDUCATION

Brown University, Providence, RI Sep 2010 - May 2017
Ph.D. Applied Mathematics (GPA: 3.9/4.0)

· Dissertation: Image Compression and Data Clustering: New Takes on Some Old Problems
· Advisor: Stuart Geman

National Taiwan University, Taiwan Sep 2004 - May 2006
M.Sc. Mathematics/Track of Statistics (GPA: 3.8/4.0)

National Taiwan University, Taiwan Sep 2000 - May 2004
B.A. Economics (GPA: 3.8/4.0)

PAPERS

- **W.-Y. Wang** and S. Geman, "Comparison Based Image Compression." In progress.
  - A novel lossless image compression scheme with analytic performance guarantees
  - Outperforming state-of-the-art algorithms in bit-per-pixel results
  - Implemented with Amazon cloud (EC2) parallel computing (in C and Python) on 80 million image patches (∼3.6GB)
- **W.-Y. Wang** and S. Geman, "Robust Generalized Clustering." In progress.
  - A highly robust unsupervised data clustering algorithm which fits multiple structures (even when data is 70% corrupted)
  - Approximating an NP-hard problem with a modified backward selection procedure
  - Applied to traffic lane detection and tone perception data
- **W.-Y. Wang** and S. Geman, "Clustering to shapes" In progress.
  - An iterative PCA method for clustering high dimensional data into descriptive manifolds
  - Capable of obtaining complicated structures, like spiral-shaped data, in a short amount of time
  - A novel dimension reduction tool, capable of obtaining complicated structures, like spiral-shaped data, in a short amount of time

INDEPENDENT PROJECTS

Kaggle Competition: Digit Recognizer

· Applying convolution neural network to build a digit classifier from MNIST data (37,800 training data)
· Achieving 99.21% correction rate on 28,000 test data (Code available at Kaggle)

Kaggle Competition: New York City Taxi Trip Duration

· Using Gradient Boost model to predict a taxi trip duration (in second) in New York city from 1.5 million data, where 8 features, includes GPS information, pickup date, and time, were provided
· Acquiring additional 16 features, like travel distance, weather, and holiday information
· Achieving RMSLE score 0.379 (top 6% of 600 teams)

Text Message Spam Detector

· Using NLP model to analyze 5,572 text message (13% of are labeled as spam)
· Comparing multiple machine learning tools: SVM, Naive Bayes, Random Forest, and Gradient Boost, etc.
· Achieved accuracy 98.9% (precision and recall are 95.7%) on 1,115 test set (code available at Kaggle)

Text Generator with Recurrent Neural Network

· Using LSTM on Python (Keras on Tensorflow backend) to build a predictive context model
· Generating Shakespeare-like article by analyzing Shakespeare's work

Improving 3D Stereo Data with Markov Random Field

· Reducing the mismatching problem when reconstructing 3D images from stereo data
· Applying conjugate gradient to speed up the procedure and obtaining a smooth reconstruction

3D Reconstruction with Structured Light

· Reconstructing a 3D image with a camera and structured light from a projector
· Obtaining a high resolution 3D image in a split second

ACADEMIC EXPERIENCE

Brown university Sep 2016 - Dec 2016
Teaching Assistant on: Probabilities in Quantum Mechanics

Brown university Sep 2011 - May 2012
Teaching Assistant on: Statistical Inference

· Topics: statistical models, point estimator, ANOVA, hypothesis test, and regression

EMPLOYMENT

Brown University, Providence, RI Jun 2017 - Present
Postdoc

Academia Sinica, Institute of Mathematics, Taiwan Nov 2008 - Aug 2010
Research Assistant

· Utilizing PCA to build an image prior for denoising. Implemented with convolution operations on GPU (speed up by 300%) in Matlab, made it possible to estimate parameters

Military Service, Taiwan Jan 2007 - Jan 2008
Coastal Patrol Corporal

· Leading about 50 troopers patrolling coastal areas

HONORS AND AWARDS

- Sigma Xi National Nomination, Brown University, 2016-2017
- University Fellowship, Brown University, 2010-2017