Xinyu Zhao

Cell: +1 (404) 408-7663 Email: xinyu.zhao@gao.io Web: http://xinyu.zhao.gao.io/
Mailing Address:359 Broun Hall, Auburn University, Auburn, AL 36849, U.S.

Objective and Qualifications

Objective: Seeking a software engineer position related to machine learning.

Highlight of Qualifications: positive attitude; fast learner; desire to pursue perfection; passion in software.

Skills

WORKING KNOWLEDGE

Programming Languages: C++11 (solid), MatLab (solid), Python (experienced), R (intermediate)
Unsupervised Learning Algorithms: Clustering, i.e., Center-based [K-Means], Hierarchical, Model-based [Expectation-Maximization], Density-based [OPTICS, DBSCAN, DPC]; Independent Component Analysis [ICA]

Supervised Learning Algorithms: Linear Regression, K Nearest Neighbors, Logistic Regression, Artificial Neural Networks, Naive Bayes, Support Vector Machine [SVM]
Other: Numpy/Pandas [for data mining], scikit-learn [for machine learning], Linux/Unix, HPCC, Statistical Analysis, Image Processing [registration, filtering, edge detection], Optimization, Simulation, Genetic Algorithms

BASIC KNOWLEDGE

MapReduce, MPI, Shell

LANGUAGES

English (full professional proficiency), Chinese (native or bilingual proficiency)

Experience

GRADUATE RESEARCH ASSISTANT | AUBURN UNIVERSITY | JAN. 2012 - PRESENT

Machine Learning

- **Investigate the correspondence of clinical diagnostic grouping with underling neurobiological clusters using unsupervised learning**
  - **Goal**: improve clinical diagnosis for different mental disorders (e.g., Autism, PTSD, etc.)
  - **Approach**: 1. Apply three different clustering methods, i.e., hierarchical, OPTICS, and DPC on functional MRI data. 2. Determine optimal subset of features and clustering result based on a genetic algorithm.
  - **Achievement**: a) high clustering accuracy (the average accuracy is 80% and the maximum accuracy is 100%); b) reduced number of features (below 100)
- **Airbnb challenge posted on Kaggle**
  - **Goal**: predict in which country a new user registered on Airbnb will make his or her first booking.

- **Approach**: 1. apply regularized logistic regression classifier on cleaned data with original features, 2. analyze learning curve to decide how to improve model, 3. Increase model complexity by adding polynomial features, 4. Adjust regularization parameter to achieve best performance.
- **Achievement**: modest accuracy (around 70%)
- **Source code: https://github.com/xinyuzhao/machine_learning_airbnb_newbooking.git**

Other Areas

- **Simulation and optimization of line edge roughness and critical dimension error in electron-beam lithography**
  - Improved simulation speed by significantly reducing number (about 0.1% of original) of point spread function (PSF) used for simulating exposure using statistical analysis and stochastic procedure.
  - Proposed two methods to determine the optimal dose required in e-beam lithography:
    - Adapted an iterative procedure which is time-consuming but can achieve high accuracy,
    - Proposed an non-iterative method by convert non-linear relationship between exposure and developing rate to piece-wise linear, which reduces computation time while slightly scarifying accuracy.
  - Simulated scanning electron microscope (SEM) images using Genetic Algorithm (GA). The simulated profile accurately matched to the profile measured from real SEM images with less than 5% error.
- **Image registration based on image moment**
  - Used image moment to calculate image transformation parameters from geometric distortion.
  - Applied point-to-point mapping to locate deformed part (e.g., tumor) in medical images.

GRADUATE TEACHING ASSISTANT | AUBURN UNIVERSITY | AUG. 2011 – DEC. 2011

- **ELEC3800 Random Signals and Systems**
  - Grader and after-class helper

Education

- **Auburn University | Auburn, AL, U.S.**

  PhD in *Electrical and Computer Engineering* | Jun. 2016 (expected) | GPA: 3.82/4.0
- **Auburn University | Auburn, AL, U.S.**

  Master in *Electrical and Computer Engineering* | Dec. 2012 | GPA: 3.82/4.0
- **East China University of Science and Technology, P.R. China**

  B.E. in *Information Engineering* | Jun. 2010

Publications

- [2016] X Zhao, et. al.: *Investigating the correspondence of clinical diagnostic grouping with underlying neurobiological and phenotypic clusters using unsupervised learning: An application to the Alzheimer's spectrum* (abstract), ISMRM 24th Annual Meeting (submitted).
- [2015] X Zhao, SY Lee, J Choi, SH Lee, IK Shin, CU Jeon: *Dependency analysis of line edge roughness in electron-beam lithography*, Microelectronic Engineering (Journal).
- [2014] X Zhao, Q Dai, SY Lee, SH Lee, BG Kim, HK Cho: *Determination and analysis of minimum dose for achieving vertical side-wall in electron-beam lithography*, Journal of Vacuum Science & Technology B 32 (6), 06F508.
- [2014] X Zhao, SY Lee, J Choi, SH Lee, IK Shin, CU Jeon, BG Kim, HK Cho: *Minimization of line edge roughness and critical dimension error in electron-beam lithography*, Journal of Vacuum Science & Technology B 32 (6), 06F505.
- [2012] X Zhao, SY Lee, SH Lee, BG Kim, HK Cho: *Fast simulation of stochastic exposure distribution in electron-beam lithography*, Journal of Vacuum Science & Technology B 30 (6), 06F308.