



# Xinyu Zhao

Cell: +1 (404) 408-7663

Email: [xinyu.zhao@gao.io](mailto:xinyu.zhao@gao.io)

Web: <http://xinyu.zhao.gao.io/>

Mailing Address: 359 Broun Hall, Auburn University, Auburn, AL 36849, U.S.

## Summary

I am a PhD candidate with a strong background on machine learning, and former experiences with lithography simulation and optimization as well as image processing. I'm passionate in machine learning, and I love solving problems with programming. I am looking for software engineering and data science positions beginning in 2016.

## Skills

**Programming Languages:** C++11, MatLab, Python3, R

**Unsupervised Learning Algorithms:** Center-based [K-Means], Hierarchical, Model-based [Expectation-Maximization], Density-based [OPTICS, DBSCAN, DPC]; Independent Component Analysis

**Supervised Learning Algorithms:** Linear Regression, K Nearest Neighbors, Logistic Regression, Artificial Neural Networks, Naive Bayes, Support Vector Machine, Decision Tree

**Tooling / Libraries:** Linux/Unix, HPC, IPython Notebook, numpy/scipy/pandas, scikit-learn

**Limited Experience:** Deep Learning, MapReduce, MPI, Genetic Algorithms, Image Processing, Statistical Analysis

## Experiences

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

#### *Machine Learning*

- ▶ **Investigate the correspondence of clinical diagnostic grouping with underlying neurobiological clusters using unsupervised learning**
  - **Goal:** improve clinical diagnosis for different mental disorders (e.g., Autism, PTSD, etc.)
  - **Approach:** 1. Apply different clustering methods on fMRI data. 2. Determine optimal subset of features and clustering result based on a genetic algorithm.
  - **Achievement:** a) high clustering accuracy (average: 80%; best-case: 100%); b) reduced number of features (from 80,000 to below 100)
  - **Keywords:** Machine Learning, Unsupervised Clustering, Big Data Processing, Hierarchical Clustering, OPTICS, Density Peak Clustering, Genetic Algorithm
- ▶ **Airbnb Kaggle Challenge**
  - **Goal:** predict which country a new Airbnb user will make his or her first booking in.
  - **Approach:** 1. Apply regularized logistic regression classifier on cleaned data; 2. analyze learning curve; 3. changing model from linear to non-linear decision tree; 4. adjust parameters to achieve best performance.
  - **Achievement:** accuracy around 85%
  - **Source code:** [https://github.com/xinyuzhao/machine\\_learning\\_airbnb\\_newbooking.git](https://github.com/xinyuzhao/machine_learning_airbnb_newbooking.git)

- **Keywords:** Machine Learning, Supervised Learning, Decision Tree, Logistic Regression

### 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 a non-iterative method by convert non-linear relationship between exposure and developing rate to piece-wise linear, which reduces computation time while slightly sacrificing 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* | 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.