Zhaoying (Natalie) Han

Palo Alto, CA •(615) 483-6742 https://github.com/nataliehan23
zhan1@stanford.edu www.linkedin.com/in/nataliehan

Technical skills

- **Languages:** Python, MySQL, Perl, Matlab, Shell scripting, HTML, C++ (some experience), Java (some experience), JavaScript (some experience)
- **Tools:** Flask, Twitter Bootstrap, D3, NLTK, Scikit-learn, Pandas, Matplotlib, Patterns, SVN/git
- **Machine learning:** Linear and logistic regression, neural networks, support vector machine, principle component analysis, clustering, collaborative filtering, recommendation engine
- **Big Data:** Map Reduce, Hadoop, HDFS, Hive, Yelp MRJob, Amazon Web Service

Experience

Fellow, Insight Data Science, *Mountain View, CA* *01/2014–present*

- Created OldiO.net, a music album recommendation engine based on over 6 million album ratings.
- Merged album data from scraping Amazon, Google/Bing/Youtube APIs, and Stanford SNAP lab.
- Leveraged Hive on top of AWS Hadoop to preprocess data, then applied collaborative filtering with Map Reduce to calculate pairwise Pearson and Cosine similarity between albums.
- Calculated top ranked albums, filtered duplicate titles with NLTK, and stored data into MySQL.
- Validated the recommendation algorithm with five-fold cross validation, resulting in RMSE of 1.1.
- Deployed app with Flask on AWS, designed front end using HTML, Bootstrap CSS and JavaScript.

Postdoctoral Fellow, Stanford University*, Stanford, CA* *10/2011-present*

- Developed a clinical data process pipeline with Perl and C++ to automatically segment brain MRI images, compare features against normal subject database, and generate quantitative reports.
- Integrated the data analysis pipeline into current Stanford hospital clinical workflow and deployed on 8 Stanford MRI scanners. The system has processed over 100 patient cases.
- Evaluated different imaging protocols and test-retest reliability with ANOVA regression model.

Research Assistant, Vanderbilt University, *Nashville, TN* *01/2011-08/2011*

- Investigated the effect of non-rigid registration algorithms on brain morphometry to detect group differences for subjects with Williams Syndrome and children with math disabilities.
- Built software pipelines integrating MATLAB, C and Bash to process data in parallel utilizing hundreds of nodes in a Linux cluster to detect statistically significant group differences.
- Proposed a Support Vector Machine (SVM) classifier with brain features to classify children with low and normal math performances. The leave-one-out classification rate reached 92.5%.
- Prototyped a unit test in Java for an image processing plugin, deployed at a Hudson server.

Education

Vanderbilt University, *Nashville, TN* *12/2011*
Ph.D. in Electrical Engineering
Educational Stipends Awards for ISMRM Scientific Meeting in 2008, 2009, 2011

Chinese Academy of Sciences*, Beijing, China* *06/2005*
M.S. in Electrical Engineering
Honors: Entrance Scholarship; Liu Yongling Award in 2004 (top 1%); Graduated with honor.

Inner Mongolia University*, Hohhot, China* *06/2002*
B.S. in Electrical Engineering
Honors: Bao Steel Award in 2001 (No. 1 out of 900 students); Graduated with honor (top 1%).