# Da Cao

214 Price Avenue, Narberth, PA 19072 Tel: 6087706029 E-mail: caoda0711@gmail.com

## PROFESSIONAL SUMMARY

An engineer passionate in solving real world problems in fields of big data applications, with strong background and working experience in machine learning and digital signal processing.

## SKILLS

- Python, NumPy, Pandas, SciPy, Scikit-learn, Tensorflow
- MySQL, MongoDB, AWS
- C++, Matlab
- Git
- Languages: English, Chinese

## WORK EXPERIENCE

**ZOOMI, INC** Princeton, NJ
Research Software Engineer / Data Engineer Jun 2015 – present

- Implemented a Multiple EM for Motif Elicitation algorithm for identifying motifs in user behavior sequences.
- Apply machine learning algorithms for student performance and completion prediction.
- Perform clickstream data processing and feature engineering for user behavior data
- Perform content topic analysis on courses using LDA and create data visualization tool with RShiny
- Built and optimized discussion forum based social network model through convex optimization.
- Automated data reporting tool that generates feature matrices from measurements in MySQL and MongoDB
- Building and testing neural network models for prediction and classification using python and Tensor flow
- Perform research projects and writing papers for publication in journals and conferences.
- Image feature extraction based on wavelet transform and classification with Convolutional Neural Networks.

**PHILADELPHIA GAME LAB** Philadelphia, PA
Software Engineer, Digital Signal Processing Jul 2014 – May 2015
Sonic project: 3D binaural audio game engine for mobile devices.

- Improved efficiency of Fast Fourier Transform algorithm by applying Fast Hartley Transform technique and Radix-4 FFT technique; Decreased the computing time of old FFT algorithm by at least 50%
- Created 3D effect for mono-channel audio by using Head Related Transfer Function (HRTF) libraries
- Modeling and implementing real-time reverberation and Doppler effect on audio mixer using IIR filter, FIR filter and convolution;
- Helped to develop and debug 3D audio mixer and circular buffer in C++;

Phonic project: English pronunciation error-detecting software.

- Applied Linear Predictive Coding technique to find formants of human speech
- Wrote an audio-based gender detection function using pitch detection methods
- Developed an intensity analysis algorithm to mark human speech in an audio and to locate exact syllables (vowel region) in a word of speech.
- Mapping formant ranges for general male, female and kid speaker

## RESEARCH & PUBLICATIONS

**D. Cao**, A.S.Lan, W.Chen, C. Brinton, M. Chiang, "Learner Behavioral Feature Refinement and Augmentation using GANs", International Conference on Artificial Intelligence in Education (AIED), June 2018

W.Chen, A.S.Lan, **D. Cao**, C. Brinton, M. Chiang, "Behavioral Analysis at Scale: Learning Course Prerequisite Structures from Learner Clickstreams". International Conference on Educational Data Mining (EDM), July 2018

W. Chen, C. Brinton, **D. Cao**, M. Chiang. "Early Detection Prediction of Learning Outcomes in Online Short-Courses via Learning Behaviors." IEEE Transactions on Learning Technologies, 2017.

W. Chen, C. Brinton, **D. Cao** and M. Chiang, "Behavior in social learning networks: Early detection for online short-courses", Proc. of IEEE INFOCOM, Atlanta, GA, May 2017.

C. Brinton, S. Buccapatnam, L. Zheng, **D. Cao**, A. Lan, F. Wong, S. Ha, M. Chiang, H. Vincent Poor, "On efficiency of Online Social Learning Networks" IEEE/ACM Transactions on Networking, 2018.

W. Chen, C. Joe-Wong, C. Brinton, L. Zheng and **D. Cao**, "Principles for Assessing Adaptive Online Courses". International Conference on Educational Data Mining (EDM), July 2018.

## EDUCATION

**University of Pennsylvania, Philadelphia, PA**
*Master of Science in Electrical Engineering, May 2014*

**University of Wisconsin - Madison, Madison, WI**
*Bachelor of Science in Electrical Engineering, May 2012*
*Certificate of Biology in Engineering, May 2012*

## CLASS PROJECT

**MUSIC GENRE CLASSIFIER**
University of Pennsylvania, Oct 2012 – Dec 2012

- Wrote a Matlab program that categorizes songs into music genres based on lyrics and audio features of the song
- Developed learning algorithms including SVM, Kernel Regression, Naïve Bayes and Boosting; Experimented the effect of different combinations of feature sets on the accuracy of prediction through incorporating unigram, bigram, chorus count, and audio features.
- Analyzed over 30 audio features of songs and over 17 lyrics features of songs to find out the characteristics of different genres of music; Used a dataset of more than 5000 songs to test our music classifier and achieved a low error rate of 18%

**CONVEX OPTIMIZATION**
University of Pennsylvania, Mar 2013 – May 2013

- Using graph theory and GAMS software, helped retail companies to design their logistics from source to distributor such that the transportation cost is minimized. Optimization included 10 clothing categories and 3 transportation methods; Analyzed the shortest path for distributor to travel among 9 cities in order to maintain minimized cost.
- Using linear programming and GAMS software, optimized (Team work) the classroom assignment schedule of two buildings on University of Pennsylvania campus in order to reduce budget and power consumption. Designed the allocation of 141 courses to 55 classrooms while maintaining the minimum cost of electricity.

**DIGITAL SIGNAL PROCESSING**
University of Pennsylvania, Jan 2014 – Mar 2014

- Developed an Automatic Speaker Recognition System on Matlab using techniques in DSP and Machine Learning.
- Extract speech features using frame blocking, windowing, fast fourier transform and process speech signals into Mel-Frequency Ceptrum Coefficients (MFCC).
- Distinguish specific speaker by using LBG algorithm and vector quantization techniques to recognize patterns in the extracted MFCCs

## AWARDS AND HONORS

- Recipient of Henry Steenbock Academic Merit Award for 2009/2010 school year (UW-Madison)
- Recipient of Claude and Dora Richardson Engineering Scholarship for 2010/2011 academic year (UW-Madison)
- Dean's Honor List for 6 continuous semesters (UW-Madison)

## ACTIVITIES

- Member of Eastern United States Kendo Federation
- University of Pennsylvania Kendo Club / Represented Upenn in Eastern United States Kendo Tournament
- University of Wisconsin - Madison Kendo Club / Represented UW-Madison twice in U.S. Mid-West Kendo Tournament

214 Price Avenue, Narberth, PA 19072 Tel: 6087706029 E-mail: caoda0711@gmail.com