# Ying-Yu (Christine) Chen

Website: http://steggie3.github.io
E-mail: yingyu.christine.chen@gmail.com
Telephone: +1-217-819-0752
Address: 4113 Branson Dr., San Mateo, CA, 94403

## PROFILE

- **Seeking full-time software engineering positions on machine-learning and/or data science**.
- Software engineer (**C++**) with experiences in **data science, algorithms,** and **compilers**. 2.5 yrs industry experience.
- Ph.D. in electrical and computer engineering with a focus on **data modeling** and **algorithms**, 14 publications in top conferences and journals including *International Conference on Data Mining*.
- Worked on a wide variety of projects including **machine learning**, **optimization**, data modeling, data mining, cloud computing, simulation, embedded systems, signal processing, design automation, etc.

## SKILLS

- **C/C++, Python, Keras, Tensorflow**, **scikit-learn**, Perl, MATLAB, Visual C++/C#, Tcl, LaTeX

## CERTIFIED ONLINE COURSES AND TRAINING PROGRAMS

**AI Bootcamp for Machine Learning Engineers, deeplearning.ai** Completed July 2018

- Built and deployed a real-time object detection system with the You-Look-Only-Once deep learning model. **(Python/Tensorflow)**

**CSMM.102x: Machine Learning, edX/Columbia University** Completed May 2018

- Regression, classification, active learning, clustering, topic modeling, recommendation systems, matrix factorization, hidden Markov models, Kalman filters, association analysis.
- Projects: ridge regression, active learning, Bayes classifier with Gaussian mixture models, K-means and EM algorithm on Gaussian mixture models, recommendation with probabilistic matrix factorization. (PMF) **(Python)**

**Deep Learning Specialization, Coursera/deeplearning.ai** Completed March 2018

- Neural networks, regularization, optimization, convolutional neural networks (CNN), ResNet, recurrent neural networks, long short-term memory (LSTM) models.
- Projects: object detection, word embedding, sentiment analysis, music synthesis. **(Python/Keras/Tensorflow)**

## EDUCATION

**M.S./Ph.D., Electrical and Computer Engineering** 2008 – 2015
**University of Illinois at Urbana-Champaign**, Urbana, IL, USA GPA: 3.71/4
*Data Mining Principles* (A+), *Random Processes* (A), *Fundamental Algorithms* (A-), *Engr. Software Systems* (A), *Cloud Computing Infrastructure* (A), *Graph Theory* (A-)

**B.S., Electrical Engineering** 2004 – 2008
**National Taiwan University**, Taipei, Taiwan GPA: 3.79/4
*Algorithms* (97/100), *Data Structures and Programming* (97/100), *Digital Speech Processing* (95/100), *Networking and Multimedia Lab* (91/100), *Computer Programming* (100/100)

## SELECTED PROJECTS

**F-RankClass** (https://steggie3.github.io/projects/f-rankclass.html) **(C++/MATLAB)** Spring 2013

- Proposed and implemented a novel heterogeneous information network mining algorithm and experimented on Wikipedia documents for joint text-image classification. Text/image classification accuracy improved up to 27.3% and 26.4%, respectively, from baseline SVM. **Published in *International Conference on Data Mining (ICDM)***.

**Loan Default Prediction** (https://steggie3.github.io/projects/loan-default.html) (**Python/Keras)** May 2018

- Applied ML models: logistic regression, random forests, gradient-boosted decision trees (XGBoost), SVM, nearest neighbors, neural networks on real-world loan application data.
- Applied techniques: one-hot encoding, imputation, normalization, feature expansion, feature hashing, feature selection, PCA, feature discretization with random forest.

**The GANder Project** (https://steggie3.github.io/projects/gander.html) (**Python/Keras/Tensorflow**) July 2018

- Applied various Generative Adversarial Network (GAN) architectures, e.g. GAN, Wasserstein GAN, Variational

Autoencoder (VAE) GAN to generate artificial Canada goose faces.

**Algorithmic Stock Trading System (C++/Python/Tensorflow/MATLAB)** 2011 - 2016

- Applied SVM, logistic regression, CNN, LSTM on real-world stock market data to predict buy events and achieved 60-80% of precision at 10% of recall on a subset of stock symbols.

**Social Network Graph Partitioning (C++)** Fall 2010

- Designed and implemented an algorithm to solve the large-scale social network graph partitioning problem and implemented algorithms based on group migration and simulated annealing for comparison.

## WORK EXPERIENCE

**R&D Engineer (Software), Sr. II, Synopsys, Inc.**, Mountain View, CA November 2015 – February 2018

Implemented a mixed-language (SystemVerilog/VHDL) compiler for the ZeBu emulation system, including:

- Optimization and transformation on computation graphs and circuit data structures, achieving 27% running time improvement, 11% peak memory usage reduction, 35% emulation resource reduction. **(C++)**
- Automatic test environment generation for time-series simulation and formal verification, creating a seamless testing and debugging framework, greatly accelerating the work flow for R&D and application engineers. **(C++/Tcl)**

**Research Intern (Software), Intel Corporation**, Hillsboro, OR August – December 2011

- Performed numerical simulation and mathematical analysis of the most probable point theory. **(C++/MATLAB)**
- Performed experiments on discrete gate-sizing optimization. **(Perl/Tcl)**

## RESEARCH EXPERIENCE

**Modeling and simulation of novel transistors and circuits** 2010 – 2015

- Compact modeling of emerging non-silicon-based transistors; statistical modeling of process variation.
- Monte Carlo numerical simulation based on above model; explorations on advantages over traditional transistors.
- Released three open-source transistor models on nanohub.org. (1200+ users since 2014.)

**Joint text-image document classification and mining** 2013 – 2014

- Proposed F-RankClass, extending a well-known ranking-based classification algorithm to support multiple modalities by including feature information in heterogeneous information networks.
- Experimented on Wikipedia, a real-world and user-generated dataset.

**Clock tree routing algorithm for circuit design automation** 2008 – 2011

- Proposed and implemented an algorithm focusing on robust control of clock accuracy under extreme conditions.

**Distributed speech recognition** 2007 – 2008

- Improved a histogram-based quantization method for distributed speech recognition systems for edge computing.

## ACADEMIC EXPERIENCE

**Research Assistant** (Advisor: Prof. Deming Chen of University of Illinois at Urbana-Champaign) 2008 – 2015

**Teaching Assistant**, University of Illinois at Urbana-Champaign 2009 – 2014

ECE 527 System-on-Chip Design, ECE 412 Microcomputer Laboratory, ECE 385 Digital Systems Laboratory
ECE 298 Digital System Design Laboratory

- Developed two labs on system-on-a-chip and embedded systems, including a USB (Universal Serial Bus) driver and an AES (Advanced Encryption Standard) decryption system.

## SELECED PUBLICATIONS

- **Y.-Y. Chen**, A. Sangai, A. Rogachev, M. Gholipour, and D. Chen, "A SPICE-compatible model of graphene nano-ribbon field-effect transistors with circuit-level delay and power analysis of process variation," IEEE *Transactions on Nanotechnology*, November 2015.
- **Y.-Y. Chen**, Z. Sun, and D. Chen, "A SPICE model of flexible transition metal dichalcogenide field-effect transistors," IEEE/ACM Design Automation Conference, June 2015.
- S. Chen, **Y.-Y. Chen**, J. Han, and P. Moulin, "A feature-enhanced ranking-based classifier for multimodal data and heterogeneous information networks," *IEEE* ***International Conference on Data Mining***, December 2013.
- **Y.-Y. Chen**, C. Dong, and D. Chen, "Clock tree synthesis under aggressive buffer insertion," *IEEE/ACM Design Automation Conference*, June 2010.
- **More at Google Scholar:** https://scholar.google.com/citations?user=_TA5iKMAAAAJ