

AREA OF INTEREST

- Machine Learning and Deep Learning, Natural Language Processing, Computer Vision

EDUCATION

Carnegie Mellon University PITTSBURGH, PA 2018 – 2020

- M.S. in Language Technologies

National Taiwan University TAIPEI, TAIWAN 2013 – 2015

- M.S. in Computer Science and Information Engineering; Ranked 1st in CS dept.; GPA: 4.3/4.3

National Chiao Tung University HSINCHU, TAIWAN 2009 – 2013

- B.S. in Computer Science; Ranked 1st in CS dept.; GPA: 4.26/4.3

EXPERIENCE

Machine Learning Scientist **Appier Inc.** - TAIPEI, TAIWAN Aug. '15 – April '18

Cross-screen analysis

- Improved critical components for cross-screen prediction, realizing 2.2x speed improvement. * Spark
- Developed a new architecture for cross-screen analysis report, realizing 5x speed improvement. * Spark
- Developed semi-supervised learning for device representations to aid cross-screen prediction.

User profiling

- Increased coverage to 1.5x while maintaining accuracy in gender prediction for online users by employing unsupervised learning to deal with label and feature sparsity. * Python, Spark, Scala
- Developed interest detection and age prediction for users.

Product recommendation

- Devised novel methods to recommend e-commerce products for ad targeting.
- Developed novel algorithms for clothing recommendation and visualization [1]. * TensorFlow, GANs

Graduate Research Assistant **National Taiwan University** Sep. '13 - Aug. '15
TAIPEI, TAIWAN

- Chinese discourse analysis [2]
- Entity recognition & disambiguation [3]
- Question answering system development

SDE Intern, Bing **Microsoft** - BELLEVUE, WA June '14 – Sep. '14

- Implemented backend data processing and frontend UX to show nearby parking lots on a map when relevant queries are issued. * C#, Cosmos, Sangam

Software Engineering Intern **IBM** - TAIPEI, TAIWAN July '13 – Aug. '13

- Unified different test frameworks for DataPower while keeping compatibility. * STAF, Python, Java

Software Engineering Intern **Google** - TAIPEI, TAIWAN July '11 – Sep. '11

- Improved template file management for iGoogle to ease development process. * Python, C++

SELECTED PUBLICATIONS

1. **Yong-Siang Shih**, Kai-Yueh Chang, Hsuan-Tien Lin, and Min Sun. Compatibility family learning for item recommendation and generation. In *Proceedings of AAAI*, February 2018
2. **Yong-Siang Shih** and Hsin-Hsi Chen. Detection, disambiguation and argument identification of discourse connectives in Chinese discourse parsing. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, pages 1891–1902, 2016

3. Yen-Pin Chiu, **Yong-Siang Shih**, Yang-Yin Lee, Chih-Chieh Shao, Ming-Lun Cai, Sheng-Lun Wei, and Hsin-Hsi Chen. NTUNLP approaches to recognizing and disambiguating entities in long and short text at the ERD challenge 2014. In *Proceedings of the First International Workshop on Entity Recognition & Disambiguation, ERD' 14*, pages 3–12, New York, NY, USA, 2014. ACM

HONORS AND AWARDS

Academic

- **Excellent Master Thesis Award** (2015) – ACLCLP – Taipei, Taiwan
- **1st Prize for Short Track** (2014) – ERD Challenge, SIGIR Workshop – Gold Coast, Australia
- **Excellent Teaching Assistant Award** (2014) – NTU CSIE Dept. – Taipei, Taiwan
- **Phi Tau Phi Award** (2013) – The Phi Tau Phi Scholastic Honor Society – Taipei, Taiwan
It was awarded to top-ranked students in CS department among all classes.
- **Presidential Award** (Fall '09, Spring '10, Fall '10, Spring '11, Fall '11, Spring '12, Fall '12) – NCTU CS Dept. – Hsinchu, Taiwan

Problem Solving & Software Development

- **1st Place** (2014) – IBM Master the Mainframe World Championship – New York, NY, U.S.A
- **Linpus Award** (2010) – Open Source Innovative Application Competition – Taipei, Taiwan
- **Honorable Mentioned** (2010) – ACM ICPC Kuala Lumpur Regional – Kuala Lumpur, Malaysia
- **16th Place** (2010) – ACM ICPC Kaohsiung Regional – Kaohsiung, Taiwan
- **3rd Place** (2010) – Taiwan National Collegiate Programming Contest – Taipei, Taiwan
- **1st Place** (2009) – IBM Taiwan z Master Contest – Taipei, Taiwan
- **11st Place** (2007) – National Problem Solving Contest – Taipei, Taiwan
- **1st Place** (2004) – Acer Digital Creation Award – Taipei, Taiwan

ADDITIONAL PROJECTS

Projects

- DeepNetworks - implementations of GANs. (2017)
– <https://github.com/shaform/DeepNetworks>
- py3track - tracking GitHub repositories that don't support Python 3 yet. (2016)
– <https://github.com/shaform/py3tracker>
- Opinion word analysis on public discussion boards. (2016)
– <http://city.shaform.com/blog/2016/02/28/scrapy.html>
- Chronological order prediction and source identification for news articles. (2015)
- Sentiment analysis and aspect term extraction for hotel reviews. (2014)
- A minimax search based player for Dala board games. (2012)
– <https://github.com/shaform/school-projects/tree/master/100-ai/pj2>
- FreeArray - an intelligent Chinese input method. (2010)
– <https://github.com/shaform/ibus-freearray>

Books

- CS Dream - an open book that introduces computer science to the general audience. (2014)
– <http://shaform.com/csdream/>

Open Source Contributions

- s3tools/s3cmd - commit several patches for Python 3 compatibility. (2016)
- dmlc/xgboost - add S3 filesystem support for model persistence. (2015)
- google/vim-codefmt - add yapf support. (2015)
- GitbookIO/gitbook - add reference links support. (2014)

LANGUAGES AND TECHNOLOGIES

- Python; C++; Scala; Spark; TensorFlow