Guang-He **Lee**

+886-975-156-835 | guanghe.lee0304@gmail.com | guanghelee.github.io

Research Interests

Machine Intelligence, Language Understanding, Natural Language Processing, and Machine Learning.

Education

National Taiwan University (NTU) *Taipei, Taiwan*

B.S. in Dept. of Computer Science and Information Engineering (CSIE) *2011 - 2015*

- GPA: 4.2/4.3; Rank: 2/111; Straight A+ in undergraduate CSIE courses. Straight A in graduate CSIE courses.
- Selected Courses (All A+): Machine Learning, Machine Learning: Theory and Practice (KDD Cup), Machine Learning and Having it Deep and Structured, Natural Language Processing, Social Network Analysis, Web Retrieval and Mining, Parallel Programming.
- Programming Language: C/C++, Python, Matlab; Tool: Scikit-Learn, OpenMP, Pthread, TensorFlow, Caffe, Latex.

Conference Papers and Patents

Conference Papers

[1] Guang-He Lee, Shao-Wen Yang, and Shou-De Lin, "Toward Implicit Sample Noise Modeling: Deviation-driven Matrix Factorization," *arXiv preprint*. [PDF]

[2] Guang-He Lee and Shou-De Lin, "LambdaMF: Learning Nonsmooth Ranking Functions in Matrix Factorization Using Lambda," in *Proc. of the 15th IEEE International Conference on Data Mining (ICDM) (acceptance rate: 18.2%)*, IEEE, Nov. 2015, p. 823-828. [PDF]

Pending Patent Applications

[3] Guang-He Lee and Shao-Wen Yang, "Observing the Unobserved: A Multi-modal Approach using Missing Data Tensor Factorization," international patent filing application under Patent Cooperation Treaty (PCT): *PCT/US2015/049110*.

[4] Guang-He Lee and Shao-Wen Yang, "Interactive Sampling Monitoring for Low-power IOT using Discriminative Probabilistic Tensor Factorization," international patent filing application under PCT: *PCT/US2015/000390*.

[5] Guang-He Lee, Kalpana Algotar, Shao-Wen Yang, and Addicam Sanjay, "Ultra-reliable Indoor Positioning using Random Forests with Temporal Bagging," international patent filing application under PCT: *PCT/US2015/067244*.

[6] Yu-An Chung, Guang-He Lee, and Shao-Wen Yang, "Cost-Sensitive Classification with Deep Learning using Cost Aware Pre-Training," U.S. patent filing application: *P88497/14/757,959*.

Honors and Awards

Presidential Awards (8 times) *NTU*

- Recognizes students with top 5% GPA in each department in each semester. *2011-2015 (every semester)*

The Honorary Member of the Phi Tau Phi Scholastic Honor Society *Phi Tau Phi Scholastic Honor Society, Taiwan*

- Honors top 1% of undergraduate graduands in academic performance and moral conduct among about 300 graduands in the College of EECS at NTU. *2015*

Undergraduate Research Project Exhibition Awards [2] *NTU CSIE, Appier Inc., and Delta Inc.*

- 2 First Place and 1 Second Place from the 3 organizers among undergraduate research projects recommended by faculty (17 projects in total). *2015*

Microsoft-IEEE Young Fellowship *Microsoft Research Asia and IEEE*

- Recognizes prominent young researchers in Asia (3 recipients in Taiwan). *2014*

Irving T. Ho Memorial Scholarship *Irving T. Ho Memorial Foundation*

- Awards to top undergraduate students in the College of EECS at NTU (4 recipients in 2013 and 2 recipients in 2014). *2013 and 2014*

Excellent Performance Award *TSMC Big Data Competition*

- The largest annual data analytic competition in Taiwan with 124 teams in 2014.
- Developed an Genetic Algorithm for feature selection in high dimensional data (advised by Prof. Hsuan-Tien Lin). *2014*

Research Experience

Machine Intelligence and Understanding Lab, National Taiwan University

Taipei, Taiwan

Research Assistant (Advisor: Prof. Yun-Nung (Vivian) Chen) *Oct. 2016 - present*

1. *Semantically Grounded Multi-sense Word Representation*
   - Proposed to learn multi-sense word representation from dictionary definitions and examples, cf. a corpus or lexical ontologies.
   - Proposed to infer sense embeddings from attentive contextual word embedding and sentence-level definition embeddings.

Intel Labs, Intel Corporation

Taipei, Taiwan

Research Intern (Manager: Dr. Yen-Kuang Chen; Mentor: Dr. Shao-Wen Yang) *Jun. 2015 - Sep. 2015*

1. *Robust Factorization Model* [1], [3], [4]:
   - Proved that the noises upon data are almost surely diverse given continuous and stochastic noises.
   - Proposed a novel low-rank noise structure to model the diverse characteristics of noise while combating overfitting.
   - Achieved 7 times faster training time and significantly lower error than the state-of-the-art deep learning model.
2. *Passive RFID tracking in retail stores* [5]:
   - Invited to visit Hillsboro (Oregon), Chandler (Arizona) and Santa Clara (California) offices to cooperate with the IoT Group.
   - Passive RFID data are unreliable due to lack of internal power source; thus precise tracking of thousands of clothes is hard.
   - Proposed a temporally smoothed random forest model with an empirical 90% accuracy using multi-store real data.
3. *Cost-sensitive Deep Learning* [6]:
   - Designed a deep learning model to embed cost information in the pre-training and training stage.
   - Achieved superior performance to the Bayes, one-sided regression, and standard deep learning methods in 7 out of 8 datasets.

Machine Discovery and Social Network Mining Lab, National Taiwan University

Taipei, Taiwan

Research Assistant (Advisor: Prof. Shou-De Lin) *Feb. 2014 - Jun. 2015*

1. *Learning-to-Rank (LTR) Matrix Factorization (MF)* [2]:
   - Proposed to optimize ranking in MF directly by lambda gradient, compared with existing approximation and bounding methods.
   - Proved that there is a divergent effect on directly combining lambda gradient with MF.
   - Proposed a stable formulation for lambda gradient on MF and a faster training algorithm from O(NlogN) to O(1) for a pair of data.
2. *Heterogeneous Transfer Learning for Convolutional Neural Network (CNN) in Super Resolution (SR):*
   - Designed a transfer learning procedure for CNN among heterogeneous tasks.
   - Achieved 11.54 times faster training time by transferring CNN from object recognition to super resolution.
3. *KDD Cup 2014 for predicting promising projects for DonorsChoose.org (Rank 12/472), jointly advised by Prof. Chih-Jen Lin:*
   - Proposed the most accurate validation set in NTU team by analyzing the temporal relationship in data.
   - Created the best single model in NTU team by designing a joint feature weighting and selection procedure for random forest.
4. *Multi-round Multi-party Influence Maximization (IM) in Social Network:*
   - Proposed a Genetic Algorithm model for IM, which is NP-hard.
   - Achieved superior performance to the Greedy Algorithm, which holds currently the best theoretical approximation factor in PTIME.

Professional and Extracurricular Activity

Teaching Assistant

Dept. of CSIE, NTU

Instructor: Prof. Hsin-Mu Tsai (2013); Prof. Shou-De Lin (2016) *2013 and 2016*

- Algorithm Design and Analysis, Fall 2013 (82 students) and Machine Discovery, Fall 2016 (90 students).

Student Volunteer

Ukulele Club, NTU

Charity camp for countryside elementary school *2014*

- Raised 40 ukuleles and taught disadvantaged children to play ukulele.

Director of Team Mentors

Dept. of CSIE, NTU

NTU CSIE camp for senior high school students *2012*

- Led 20 mentors (college students) to train 100 participated senior high school students on basic computer science.

Director of Public Relation

The Affiliated Senior High School of National Taiwan Normal University

Sign language club *2009*

- Coordinated with club representatives from other senior high schools.
- Organized a joint performance with 31 senior high schools for more than 1,300 audience.