

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

Massachusetts Institute of Technology (MIT)

Cambridge, MA

PH.D. IN COMPUTER SCIENCE AND ARTIFICIAL INTELLIGENCE LABORATORY (CSAIL)

2017 - now

- Advisor: Dina Katabi.

National Taiwan University (NTU)

Taipei, Taiwan

M.S./B.S. IN DEPT. OF COMPUTER SCIENCE AND INFORMATION ENGINEERING (CSIE)

2016 - 2017/2011 - 2015

- (M.S.) Advisor: Yun-Nung (Vivian) Chen.
- (B.S.) 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 and Yun-Nung Chen](#), “MUSE: Modularizing Unsupervised Sense Embeddings” in *Proc. of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (acceptance rate: 22.8%), 2017. [[PDF](#)]
- [2] [Guang-He Lee](#), Shao-Wen Yang, and Shou-De Lin, “Toward Implicit Sample Noise Modeling: Deviation-driven Matrix Factorization,” arXiv preprint. [[PDF](#)]
- [3] [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

- [4] [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*.
- [5] [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*.
- [6] [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*.
- [7] 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

- 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 [3]

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

Research Experience

Machine Intelligence and Understanding Lab, National Taiwan University

Taipei, Taiwan

MASTER'S STUDENT (ADVISOR: PROF. YUN-NUNG (VIVIAN) CHEN)

Oct. 2016 - Jul. 2017

1. *MUSE: Modularizing Unsupervised Sense Embeddings* [1]
 - MUSE is the first system to achieve pure sense-level representation while maintaining linear time sense selection.
 - MUSE is the first system to exploit reinforcement learning to model the decision process for unsupervised sense embeddings.
 - Achieved state-of-the-art performance in the benchmark contextual word similarity task.

Intel Labs, Intel Corporation

Taipei, Taiwan

RESEARCH INTERN (MANAGER: DR. YEN-KUANG CHEN; MENTOR: DR. SHAO-WEN YANG)

Jun. 2015 - Sep. 2015

1. *Robust Machine Learning for Heteroscedasticity* [2], [4], [5]:
 - Proved that the noises upon data are almost surely diverse given continuous and stochastic noises.
 - Proposed a novel low-rank noise structure for factorization models to model the heteroscedasticity, c.f. a sparse structure.
 - Proposed a novel uncertainty modeling framework for tree-based model using gradient boosting.
 - 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* [6]:
 - 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* [7]:
 - 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)* [3]:
 - 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 $O(N \log N)$ 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

- Algorithm Design and Analysis, Fall 2013 (82 students), instructed by Prof. Hsin-Mu Tsai.
- Machine Discovery, Fall 2016 (90 students), instructed by Prof. Shou-De Lin.
- Intelligent Conversational Bot, Spring 2017 (88 students), instructed by Prof. Yun-Nung Chen.

2013, 2016, and 2017

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.