Chih-Jen Lin

- PERSONAL DATA

1. Address: Department of Computer Science and Information Engineering, National Taiwan University, Taipei 106, Taiwan
2. Phone: (886) 2-33664923
3. E-mail: cjlin@csie.ntu.edu.tw
4. Homepage: `http://www.csie.ntu.edu.tw/~cjlin`

- EDUCATION AND CURRENT POSITION:

1. Distinguished professor, Department of Computer Science and Information Engineering, National Taiwan University, Taipei, Taiwan, 2011 – present
2. Affiliated professor, MBZUAI, Abu Dhabi, UAE, December 2022 – present
3. Adjunct distinguished professor, Graduate Institute of Networking and Multimedia, National Taiwan University, Taipei, Taiwan, August 2011 – present
4. Ph.D., Industrial & Operations Engineering, University of Michigan, September 1995 – May 1998.
5. M.S.E., Industrial & Operations Engineering, University of Michigan, September 1995 – December 1996.
6. B.S., Mathematics, National Taiwan University, October 1989 – June 1993.

- RESEARCH INTERESTS:

1. Machine learning: support vector machines, large-scale data classification, and applications

   We develop popular machine learning software including LIBSVM (`http://www.csie.ntu.edu.tw/~cjlin/libsvm`) and LIBLINEAR (`http://www.csie.ntu.edu.tw/~cjlin/liblinear`).
2. Large-scale optimization and its applications

- AWARDS AND RECOGNITION:

– International:

1. Outstanding paper award, Annual Meeting of the Association for Computational Linguistics (ACL) 2023 (with students Yu-Chen Lin, Si-An Chen, Jie-Jyun Liu)

2. Best paper award, Asian Conference on Machine Learning (ACML) 2018 (with students Chih-Yang Hsia and Wei-Lin Chiang)
3. ACM fellow, 2015
4. AAAI fellow, 2014
5. Best paper award, ACM Recommender Systems 2013 (with students Yong Zhuang, Wei-Sheng Chin, and Yu-Chin Juan)
6. ACM Distinguished Scientist, 2011
7. IEEE fellow (class of 2011) for contributions to support vector machine algorithms and software.
8. Member of the NTU team to win the first prize of KDD cup 2010, 2011 and 2013.
9. Best research paper award, ACM KDD 2010 (with students Hsiang-Fu Yu, Cho-Jui Hsieh, and Kai-Wei Chang).
10. Supervising students Chia-Hua Ho and Ming-Hen Tsai to win the 2nd place of Active Learning Challenge 2010.
    `http://www.causality.inf.ethz.ch/activelearning.php`
11. Member of the NTU team to win the 3rd place of KDD cup 2009 (extended track) .
12. Winner of ICML 2008 large-scale learning challenge (linear SVM track; with students Hsiang-Fu Yu, Cho-Jui Hsieh, and Kai-Wei Chang).
    `http://largescale.first.fraunhofer.de/summary/`
13. Supervising student Yin-Wen Chang to win WCCI 2008 Causation and Prediction challenge.
    `http://www.causality.inf.ethz.ch/home.php`
14. Winner of WCCI 2002 competition on sequence recognition (with master students Ming-Wei Chang and Bo-Juen Chen)
15. Winner the EUNITE 2001 world wide competition (18 research groups) on electricity load prediction (`http://neuron-ai.tuke.sk/competition`). EUNITE is the European Network of Excellence on Intelligent Technology for smart adaptive systems (with master students Ming-Wei Chang and Bo-Juen Chen).
16. Winner of IJCNN Challenge 2001. IJCNN is one of the major Neural Networks conferences (with master student Chih-Chung Chang).
17. Winner of the OCR (Optical Character Recognition) competition organized by the University of Essex and the UK Post Office, December 2000. (with master student Chih-Chung Chang)
18. Second prize of the student paper competition, Fifth Copper Mountain conference on iterative methods, 1998.

19. Wallace J. Givens Research Associate (twice): competitive positions in Mathematics and Computer Science Division of Argonne National Laboratory which are intended to encourage graduate students who are beginning careers in computational science.

– Domestic:

1. Micron Chair Professorship, 2022
2. Academic Award, Ministry of Education, 2021
3. Science Chair Award, Far Eastern Y. Z. Hsu Science and Technology Memorial Foundation, 2019
4. Outstanding Scholar Award, Foundation for the Advancement of Outstanding Scholarship (FAOS), 2016
5. Outstanding research award of Pan Wen Yuan Foundation, Taiwan, 2016
6. Pegatron Chair Professorship, 2016
7. Teco Award, 2015
8. Macronix International Co. Chair Professorship, 2014
9. K. T. Li Breakthrough Award, Institute of Information & Computing Machinery, Taiwan, 2012
10. NTU EECS Academic Excellence Award, NTU College of EECS, 2011.
11. Ten outstanding young persons of Taiwan, 2011
12. Distinguished Scholar Research Project, National Science Council, Taiwan, 2009 – 2012.
13. Outstanding Research Award, National Science Council, Taiwan, 2007, 2010, and 2013.
14. Ta-You Wu Memorial Award, National Science Council, Taiwan, 2006.
15. Fu Ssu-Nien Award of National Taiwan University, 2005
16. Research award for young researchers from Pan Wen-Yuan Foundation, Taiwan, 2003.
17. K. T. Li award for young researchers from ACM Taipei/Taiwan chapter, July, 2002 (one awarded per year for young computer scientists in Taiwan)
18. Young investigator award from Academia Sinica, Taiwan, May, 2002 (15 awarded per year in all research areas)
19. Prize for Outstanding Performance, National Mathematics Contest, R.O.C. 1989.

- PROFESSIONAL EXPERIENCE:

1. Distinguished Professor (August 2011–present), Professor (August 2006–present), Associate Professor (August 2002–August 2006), Assistant Professor (August 1998–August

2002), Department of Computer Science and Information Engineering, National Taiwan University

2. Visiting professor, UCLA, January 2019 – February 2020.
3. Visiting professor, Alibaba, June – September 2017.
4. Visiting researcher, Microsoft, USA, January 2015 – September 2015, August 2016 – February 2017.
5. Visiting principal research scientist, eBay Research Labs, January 2012 – September 2012.
6. Visiting scientist, Google Research, February 2008 – September 2008.
7. Visiting scientist, Yahoo! Research, Burbank, California, August 2006 – February 2007.
8. Adjunct Associate Professor, Graduate Institute of Networking and Multimedia, National Taiwan University, August 2004– July 2006
9. Adjunct distinguished professor (August 2011–July 2023), Adjunct professor (August 2006–July 2011), Adjunct Associate Professor (August 2002–July 2006), Adjunct Assistant Professor (August 2001–July 2002), Graduate Institute of Industrial Engineering, National Taiwan University,
10. Visiting Scientist, Mathematics and Computer Science division, Argonne National Laboratory, January 1999–February 1999, May 1999–August 1999.
11. Research Associate, Mathematics and Computer Science division, Argonne National Laboratory, January 1997–April 1997, September 1997–September 1998.
12. Wallace J. Givens Research Associate, Mathematics and Computer Science division, Argonne National Laboratory, May 1996–August 1996 and May 1997–August 1997.
13. Research Assistant, Department of Industrial and Operations Engineering, University of Michigan, September 1995–August 1998.
14. Teaching Assistant, Department of Industrial and Operations Engineering, University of Michigan, September 1996–December 1996.
15. Second Lieutenant, R.O.C. Army, July 1993 – May 1995.

- JOURNAL PAPERS:

[1] Ching-Pei Lee, Po-Wei Wang, and Chih-Jen Lin. Limited-memory common-directions method for large-scale optimization: convergence, parallelization, and distributed optimization. *Mathematical Programming Computation*, 14:543–591, 2022. URL `https://www.csie.ntu.edu.tw/~cjlin/papers/l-commdir-journal/l-commdir.pdf`.

[2] Leonardo Galli and Chih-Jen Lin. A study on truncated Newton methods for linear classification. *IEEE Transactions on Neural Networks and Learning Systems*, 33(7):2828–2841, 2022. URL https://www.csie.ntu.edu.tw/~cjlin/papers/tncg/tncg.pdf.

[3] Giulio Galvan, Matteo Lapucci, Chih-Jen Lin, and Marco Sciandrone. A two-level decomposition framework exploiting first and second order information for SVM training problems. *Journal of Machine Learning Research*, 22(23):1–38, 2021. URL https://www.csie.ntu.edu.tw/~cjlin/papers/19-632.pdf.

[4] Jui-Yang Hsia and Chih-Jen Lin. Parameter selection for linear support vector regression. *IEEE Transactions on Neural Networks and Learning Systems*, 31:5639–5644, 2020. URL https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/warm-start/svr-param.pdf.

[5] Chien-Chih Wang, Kent Loong Tan, and Chih-Jen Lin. Newton methods for convolutional neural networks. *ACM Transactions on Intelligent Systems and Technology*, 11(2):19:1–19:30, 2020. URL https://www.csie.ntu.edu.tw/~cjlin/papers/cnn/newton-CNN.pdf.

[6] Po-Wei Wang, Ching-Pei Lee, and Chih-Jen Lin. The common-directions method for regularized empirical risk minimization. *Journal of Machine Learning Research*, 20(58):1–49, 2019. URL https://www.csie.ntu.edu.tw/~cjlin/papers/commdir/commdir.pdf.

[7] Wei-Sheng Chin, Bo-Wen Yuan, Meng-Yuan Yang, and Chih-Jen Lin. An efficient alternating Newton method for learning factorization machines. *ACM Transactions on Intelligent Systems and Technology*, 9(6):72:1–72:31, 2018. URL https://www.csie.ntu.edu.tw/~cjlin/papers/fm/scalefm.pdf.

[8] Chien-Chih Wang, Kent-Loong Tan, Chun-Ting Chen, Yu-Hsiang Lin, S. Sathiya Keerthi, Dhruv Mahajan, Sellamanickam Sundararajan, and Chih-Jen Lin. Distributed Newton methods for deep learning. *Neural Computation*, 30(6):1673–1724, 2018. URL http://www.csie.ntu.edu.tw/~cjlin/papers/dnn/dsh.pdf.

[9] Wei-Sheng Chin, Bo-Wen Yuan, Meng-Yuan Yang, Yong Zhuang, Yu-Chin Juan, and Chih-Jen Lin. LIBMF: A library for parallel matrix factorization in shared-memory systems. *Journal of Machine Learning Research*, 17(86):1–5, 2016. URL https://www.csie.ntu.edu.tw/~cjlin/papers/libmf/libmf_open_source.pdf.

[10] Wei-Sheng Chin, Yong Zhuang, Yu-Chin Juan, and Chih-Jen Lin. A fast parallel stochastic gradient method for matrix factorization in shared memory systems. *ACM*

Transactions on Intelligent Systems and Technology, 6:2:1–2:24, 2015. URL http://www.csie.ntu.edu.tw/~cjlin/papers/libmf/libmf_journal.pdf.

[11] Chien-Chih Wang, Chun-Heng Huang, and Chih-Jen Lin. Subsampled Hessian Newton methods for supervised learning. *Neural Computation*, 27(8):1766–1795, 2015. URL http://www.csie.ntu.edu.tw/~cjlin/papers/sub_hessian/sample_hessian.pdf.

[12] Po-Wei Wang and Chih-Jen Lin. Iteration complexity of feasible descent methods for convex optimization. *Journal of Machine Learning Research*, 15:1523–1548, 2014. URL http://www.csie.ntu.edu.tw/~cjlin/papers/cdlinear.pdf.

[13] Ching-Pei Lee and Chih-Jen Lin. Large-scale linear rankSVM. *Neural Computation*, 26(4):781–817, 2014. URL http://www.csie.ntu.edu.tw/~cjlin/papers/ranksvm/ranksvml2.pdf.

[14] Ching-Pei Lee and Chih-Jen Lin. A study on L2-loss (squared hinge-loss) multi-class SVM. *Neural Computation*, 25(5):1302–1323, 2013. URL http://www.csie.ntu.edu.tw/~cjlin/papers/l2mcsvm/l2mcsvm.pdf.

[15] Chia-Hua Ho and Chih-Jen Lin. Large-scale linear support vector regression. *Journal of Machine Learning Research*, 13:3323–3348, 2012. URL http://www.csie.ntu.edu.tw/~cjlin/papers/linear-svr.pdf.

[16] Guo-Xun Yuan, Chia-Hua Ho, and Chih-Jen Lin. An improved GLMNET for l1-regularized logistic regression. *Journal of Machine Learning Research*, 13:1999–2030, 2012. URL http://www.csie.ntu.edu.tw/~cjlin/papers/l1_glmnet/long-glmnet.pdf.

[17] Guo-Xun Yuan, Chia-Hua Ho, and Chih-Jen Lin. Recent advances of large-scale linear classification. *Proceedings of the IEEE*, 100(9):2584–2603, 2012. URL http://www.csie.ntu.edu.tw/~cjlin/papers/survey-linear.pdf.

[18] Hsiang-Fu Yu, Cho-Jui Hsieh, Kai-Wei Chang, and Chih-Jen Lin. Large linear classification when data cannot fit in memory. *ACM Transactions on Knowledge Discovery from Data*, 5(4):23:1–23:23, February 2012. URL http://www.csie.ntu.edu.tw/~cjlin/papers/kdd_disk_decomposition.pdf.

[19] Chih-Chung Chang and Chih-Jen Lin. LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3):27:1–27:27, 2011. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

[20] Wen-Yen Chen, Yangqiu Song, Hongjie Bai, Chih-Jen Lin, and Edward Y. Chang. Parallel spectral clustering in distributed systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(3):568–586, 2011.

[21] Ruby C. Weng and Chih-Jen Lin. A Bayesian approximation method for online ranking. *Journal of Machine Learning Research*, 12:267–300, 2011. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/online_ranking/online_journal.pdf`.

[22] Hsiang-Fu Yu, Fang-Lan Huang, and Chih-Jen Lin. Dual coordinate descent methods for logistic regression and maximum entropy models. *Machine Learning*, 85(1-2):41–75, October 2011. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/maxent_dual.pdf`.

[23] Guo-Xun Yuan, Kai-Wei Chang, Cho-Jui Hsieh, and Chih-Jen Lin. A comparison of optimization methods and software for large-scale l1-regularized linear classification. *Journal of Machine Learning Research*, 11:3183–3234, 2010. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/l1.pdf`.

[24] Yin-Wen Chang, Cho-Jui Hsieh, Kai-Wei Chang, Michael Ringgaard, and Chih-Jen Lin. Training and testing low-degree polynomial data mappings via linear SVM. *Journal of Machine Learning Research*, 11:1471–1490, 2010. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/lowpoly_journal.pdf`.

[25] Fang-Lan Huang, Cho-Jui Hsieh, Kai-Wei Chang, and Chih-Jen Lin. Iterative scaling and coordinate descent methods for maximum entropy. *Journal of Machine Learning Research*, 11:815–848, 2010. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/maxent_journal.pdf`.

[26] Chih-Jen Lin, Stefano Lucidi, Laura Palagi, Arnaldo Risi, and Marco Sciandrone. Decomposition algorithm model for singly linearly constrained problems subject to lower and upper bounds. *Journal of Optimization Theory and Applications*, 141:107–126, 2009.

[27] Tzu-Kuo Huang, Chih-Jen Lin, and Ruby C. Weng. Ranking individuals by group comparisons. *Journal of Machine Learning Research*, 9:2187–2216, 2008. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/genBTexp/genBTexp-jmlr.pdf`.

[28] Rong-En Fan, Kai-Wei Chang, Cho-Jui Hsieh, Xiang-Rui Wang, and Chih-Jen Lin. LIBLINEAR: a library for large linear classification. *Journal of Machine Learning Research*, 9:1871–1874, 2008. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/liblinear.pdf`.

[29] Kai-Wei Chang, Cho-Jui Hsieh, and Chih-Jen Lin. Coordinate descent method for large-scale L2-loss linear SVM. *Journal of Machine Learning Research*, 9:1369–1398, 2008. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/cdl2.pdf`.

[30] Chih-Jen Lin, Ruby C. Weng, and S. Sathiya Keerthi. Trust region Newton method for large-scale logistic regression. *Journal of Machine Learning Research*, 9:627–650, 2008. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/logistic.pdf`.

[31] Hsuan-Tien Lin, Chih-Jen Lin, and Ruby C. Weng. A note on Platt's probabilistic outputs for support vector machines. *Machine Learning*, 68:267–276, 2007. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/plattprob.pdf`.

[32] Chih-Jen Lin. On the convergence of multiplicative update algorithms for non-negative matrix factorization. *IEEE Transactions on Neural Networks*, 18(6):1589–1596, 2007. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/multconv.pdf`.

[33] Chih-Jen Lin. Projected gradient methods for non-negative matrix factorization. *Neural Computation*, 19:2756–2779, 2007. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/pgradnmf.pdf`.

[34] Tzu-Kuo Huang, Ruby C. Weng, and Chih-Jen Lin. Generalized Bradley-Terry models and multi-class probability estimates. *Journal of Machine Learning Research*, 7:85–115, 2006. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/generalBT.pdf`.

[35] Pai-Hsuen Chen, Rong-En Fan, and Chih-Jen Lin. A study on SMO-type decomposition methods for support vector machines. *IEEE Transactions on Neural Networks*, 17:893–908, July 2006. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/generalSMO.pdf`.

[36] Rong-En Fan, Pai-Hsuen Chen, and Chih-Jen Lin. Working set selection using second order information for training SVM. *Journal of Machine Learning Research*, 6:1889–1918, 2005. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/quadworkset.pdf`.

[37] Ming-Wei Chang and Chih-Jen Lin. Leave-one-out bounds for support vector regression model selection. *Neural Computation*, 17(5):1188–1222, 2005.

[38] Pai-Hsuen Chen, Chih-Jen Lin, and Bernhard Schölkopf. A tutorial on ν-support vector machines. *Applied Stochastic Models in Business and Industry*, 21:111–136, 2005. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/nusvmtoturial.pdf`.

[39] Ting-Fan Wu, Chih-Jen Lin, and Ruby C. Weng. Probability estimates for multi-class classification by pairwise coupling. *Journal of Machine Learning Research*, 5:975–1005, 2004. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/svmprob/svmprob.pdf`.

[40] Bo-Juen Chen, Ming-Wei Chang, and Chih-Jen Lin. Load forecasting using support vector machines: A study on EUNITE competition 2001. *IEEE Transactions on Power Systems*, 19(4):1821–1830, November 2004.

[41] Wei-Chun Kao, Kai-Min Chung, Chia-Liang Sun, and Chih-Jen Lin. Decomposition methods for linear support vector machines. *Neural Computation*, 16(8):1689–1704, 2004. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/linear.pdf`.

[42] Ming-Wei Chang, Chih-Jen Lin, and Ruby C. Weng. Analysis of switching dynamics with competing support vector machines. *IEEE Transactions on Neural Networks*, 15 (3):720–727, 2004.

[43] Chin-Sheng Yu, Chih-Jen Lin, and Jen-Kang Hwang. Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. *Protein Science*, 13:1402–1406, 2004.

[44] Kai-Min Chung, Wei-Chun Kao, Chia-Liang Sun, Li-Lun Wang, and Chih-Jen Lin. Radius margin bounds for support vector machines with the RBF kernel. *Neural Computation*, 15:2643–2681, 2003.

[45] S. Sathiya Keerthi and Chih-Jen Lin. Asymptotic behaviors of support vector machines with Gaussian kernel. *Neural Computation*, 15(7):1667–1689, 2003.

[46] Kuan-Min Lin and Chih-Jen Lin. A study on reduced support vector machines. *IEEE Transactions on Neural Networks*, 14(6):1449–1559, 2003. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/rsvmTEX.pdf`.

[47] Chin-Sheng Yu, Jung-Ying Wang, Jinn-Moon Yang, Ping-Chiang Lyu, Chih-Jen Lin, and Jen-Kang Hwang. Fine-grained protein fold assignment by support vector machines using generalize $n$peptide coding schemes and jury voting from multiple-parameter sets. *Proteins*, 50:531–536, 2003.

[48] Chih-Jen Lin. A formal analysis of stopping criteria of decomposition methods for support vector machines. *IEEE Transactions on Neural Networks*, 13(5):1045–1052, 2002. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/stop.ps.gz`.

[49] Chih-Jen Lin. Asymptotic convergence of an SMO algorithm without any assumptions. *IEEE Transactions on Neural Networks*, 13(1):248–250, 2002. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/q2conv.pdf`.

[50] Chih-Chung Chang and Chih-Jen Lin. Training ν-support vector regression: Theory and algorithms. *Neural Computation*, 14(8):1959–1977, 2002.

[51] Shuo-Peng Liao, Hsuan-Tien Lin, and Chih-Jen Lin. A note on the decomposition methods for support vector regression. *Neural Computation*, 14:1267–1281, 2002.

[52] Chih-Wei Hsu and Chih-Jen Lin. A comparison of methods for multi-class support vector machines. *IEEE Transactions on Neural Networks*, 13(2):415–425, 2002.

[53] Chih-Wei Hsu and Chih-Jen Lin. A simple decomposition method for support vector machines. *Machine Learning*, 46:291–314, 2002.

[54] Chih-Jen Lin. On the convergence of the decomposition method for support vector machines. *IEEE Transactions on Neural Networks*, 12(6):1288–1298, 2001. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/conv.ps.gz`.

[55] Jinn-Moon Yang, Jorng-Tzong Horng, Chih-Jen Lin, and Cheng-Yan Kao. Optical coating design using the family competition evolutionaty algorithm. *Evolutionary Computation*, 9(4):421–444, 2001.

[56] Chih-Chung Chang and Chih-Jen Lin. Training ν-support vector classifiers: Theory and algorithms. *Neural Computation*, 13(9):2119–2147, 2001.

[57] Chih-Jen Lin. Formulations of support vector machines: a note from an optimization point of view. *Neural Computation*, 13(2):307–317, 2001.

[58] Shu-Cherng Fang, Chih-Jen Lin, and Soon-Yi Wu. Solving quadratic semi-infinite programming problems by using relaxed cutting plane scheme. *Journal of Computational and Applied Mathematics*, 129:89–104, 2001.

[59] Soon-Yi Wu, Shu-Cherng Fang, and Chih-Jen Lin. Solving the general capacity problem. *Annals of Operations Research*, 103:193–211, 2001.

[60] Chih-Chung Chang, Chih-Wei Hsu, and Chih-Jen Lin. The analysis of decomposition methods for support vector machines. *IEEE Transactions on Neural Networks*, 11(4): 1003–1008, 2000.

[61] Chih-Jen Lin and Romesh Saigal. An incomplete Cholesky factorization for dense matrices. *BIT*, 40:536–558, 2000.

[62] Chih-Jen Lin and Jorge J. Moré. Newton's method for large-scale bound constrained problems. *SIAM Journal on Optimization*, 9:1100–1127, 1999.

[63] Chih-Jen Lin and Jorge J. Moré. Incomplete Cholesky factorizations with limited memory. *SIAM J. Sci. Comput.*, 21:24–45, 1999.

[64] Shu-Cherng Fang, Soon-Yi Wu, and Chih-Jen Lin. Relaxed cutting plane method for solving linear semi-infinite programming problems. *Journal of Optimization Theory and Applications*, 99:759–779, 1998.

[65] Chih-Jen Lin, Soon-Yi Wu, and Shu-Cherng Fang. An unconstrained convex programming approach for solving linear semi-infinite programming problems. *SIAM Journal on Optimization*, 8(2), 1998.

[66] Chih-Jen Lin, Soon-Yi Wu, and Shu-Cherng Fang. On the parametric linear semi–infinite optimization. *Applied Mathematics Letter*, 9:89–96, 1996.

[67] Chih-Jen Lin, E. K. Yang, Shu-Cherng Fang, and Soon-Yi Wu. Implementation of an inexact approach to solving linear semi-infinite programming problems. *Journal of Computational and Applied Mathematics*, 61:87–103, 1995.

[68] Shu-Cherng Fang, Chih-Jen Lin, and Soon-Yi Wu. On solving convex quadratic semi-infinite programming problems. *Optimization*, 37:107–125, 1994.

- REFEREED CONFERENCE PAPERS

  Some papers here are preliminary versions of journal papers.

[1] Yu-Jen Lin and Chih-Jen Lin. On the thresholding strategy for infrequent labels in multi-label classification. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management (CIKM)*, pages 1441–1450, 2023. doi: 10.1145/3583780.3614996. URL `https://www.csie.ntu.edu.tw/~cjlin/papers/thresholding/smooth_acm.pdf`.

[2] Yu-Chen Lin, Si-An Chen, Jie-Jyun Liu, and Chih-Jen Lin. Linear classifier: An often-forgotten baseline for text classification. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 1876–1888, 2023. URL `https://www.csie.ntu.edu.tw/~cjlin/papers/text_classification_baseline/text_classification_baseline.pdf`. Short paper.

[3] Yaxu Liu, Jui-Nan Yen, Bowen Yuan, Rundong Shi, Peng Yan, and Chih-Jen Lin. Practical counterfactual policy learning for top-k recommendations. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, 2022. URL `https://www.csie.ntu.edu.tw/~cjlin/papers/counterfactual_topk/xcf.pdf`.

[4] Si-An Chen, Jie-Jyun Liu, Tsung-Han Yang, Hsuan-Tien Lin, and Chih-Jen Lin. Even the simplest baseline needs careful re-investigation: A case study on XML-CNN. In *Proceedings of the Annual Conference of the North American Chapter of the Association*

for Computational Linguistics (NAACL), pages 1987–2000, 2022. URL https://www.csie.ntu.edu.tw/~cjlin/papers/xmlcnn/xml_cnn_study.pdf.

[5] Li-Chung Lin, Cheng-Hung Liu, Chih-Ming Chen, Kai-Chin Hsu, I-Feng Wu, Ming-Feng Tsai, and Chih-Jen Lin. On the use of unrealistic predictions in hundreds of papers evaluating graph representations. In *Proceedings of the Thirty-Sixth AAAI Conference on Artificial Intelligence (AAAI)*, pages 7479–7487, 2022. URL https://www.csie.ntu.edu.tw/~cjlin/papers/multilabel-embedding/multilabel_embedding.pdf.

[6] Jui-Nan Yen and Chih-Jen Lin. Limited-memory common-directions method with subsampled newton directions for large-scale linear classification. In *Proceedings of the IEEE International Conference on Data Mining (ICDM)*, 2021. URL https://www.csie.ntu.edu.tw/~cjlin/papers/commdir_subsampled/main.pdf.

[7] Bowen Yuan, Yu-Sheng Li, Pengrui Quan, and Chih-Jen Lin. Efficient optimization methods for extreme similarity learning with nonlinear embeddings. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2021. URL http://www.csie.ntu.edu.tw/~cjlin/papers/similarity_learning/pq.pdf.

[8] Jie-Jyun Liu, Tsung-Han Yang, Si-An Chen, and Chih-Jen Lin. Parameter selection: Why we should pay more attention to it. In *Proceedings of the 59th Annual Meeting of the Association of Computational Linguistics (ACL)*, pages 825–830, 2021. URL https://www.csie.ntu.edu.tw/~cjlin/papers/parameter_selection/acl2021_parameter_selection.pdf. Short paper.

[9] Bowen Yuan, Yaxu Liu, Jui-Yang Hsia, Zhenhua Dong, and Chih-Jen Lin. Unbiased ad click prediction for position-aware advertising systems. In *Proceedings of the 14th ACM Conference on Recommender Systems*, 2020. URL http://www.csie.ntu.edu.tw/~cjlin/papers/debiases/debiases.pdf.

[10] Chih-Yao Chang, Xing Tang, Bowen Yuan, Jui-Yang Hsia, Zhirong Liu, Zhenhua Dong, Xiuqiang He, and Chih-Jen Lin. AutoConjunction: adaptive model-based feature conjunction for CTR prediction. In *Proceedings of the 21st IEEE International Conference on Mobile Data Management (MDM)*, pages 202–209, 2020.

[11] Hung-Yi Chou, Pin-Yen Lin, and Chih-Jen Lin. Dual coordinate-descent methods for linear one-class SVM and SVDD. In *Proceedings of SIAM International Conference on Data Mining (SDM)*, 2020. URL http://www.csie.ntu.edu.tw/~cjlin/papers/linear_oneclass_SVM/siam.pdf.

[12] Chi-Cheng Chiu, Pin-Yen Lin, and Chih-Jen Lin. Two-variable block dual coordinate descent methods for large-scale linear support vector machines. In *Proceedings of SIAM International Conference on Data Mining (SDM)*, 2020. URL `https://www.csie.ntu.edu.tw/~cjlin/papers/2var_cd/twocddual.pdf`.

[13] Bowen Yuan, Jui-Yang Hsia, Meng-Yuan Yang, Hong Zhu, Chihyao Chang, Zhenhua Dong, and Chih-Jen Lin. Improving Ad click prediction by considering non-displayed events. In *Proceedings of the 28th ACM International Conference on Conference on Information and Knowledge Management (CIKM)*, 2019. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/occtr/ctr_oc.pdf`.

[14] Chih-Yang Hsia, Wei-Lin Chiang, and Chih-Jen Lin. Preconditioned conjugate gradient methods in truncated Newton frameworks for large-scale linear classification. In *Proceedings of the Asian Conference on Machine Learning (ACML)*, 2018. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/tron_pcg/precondition.pdf`.

[15] Yong Zhuang, Yuchin Juan, Guo-Xun Yuan, and Chih-Jen Lin. Naive parallelization of coordinate descent methods and an application on multi-core l1-regularized classification. In *Proceedings of the 27th ACM International Conference on Conference on Information and Knowledge Management (CIKM)*, 2018. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/l1_multicore_cikm.pdf`.

[16] Wei-Lin Chiang, Yu-Sheng Li, Ching-Pei Lee, and Chih-Jen Lin. Limited-memory common-directions method for distributed l1-regularized linear classification. In *Proceedings of SIAM International Conference on Data Mining (SDM)*, 2018. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/l-commdir-l1/OWL-commdir.pdf`.

[17] Chih-Yang Hsia, Ya Zhu, and Chih-Jen Lin. A study on trust region update rules in Newton methods for large-scale linear classification. In *Proceedings of the Asian Conference on Machine Learning (ACML)*, 2017. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/newtron/newtron.pdf`.

[18] Ching-Pei Lee, Po-Wei Wang, Weizhu Chen, and Chih-Jen Lin. Limited-memory common-directions method for distributed optimization and its application on empirical risk minimization. In *Proceedings of SIAM International Conference on Data Mining (SDM)*, 2017. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/l-commdir/l-commdir.pdf`.

[19] Hsiang-Fu Yu, Hsin-Yuan Huang, Inderjit S. Dihillon, and Chih-Jen Lin. A unified algorithm for one-class structured matrix factorization with side informa-

tion. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI)*, 2017. URL http://www.csie.ntu.edu.tw/~cjlin/papers/ocmf-side/biased-leml-aaai-with-supp.pdf.

[20] Hsiang-Fu Yu, Mikhail Bilenko, and Chih-Jen Lin. Selection of negative samples for one-class matrix factorization. In *Proceedings of SIAM International Conference on Data Mining (SDM)*, 2017. URL http://www.csie.ntu.edu.tw/~cjlin/papers/one-class-mf/biased-mf-sdm-with-supp.pdf.

[21] Yuchin Juan, Yong Zhuang, Wei-Sheng Chin, and Chih-Jen Lin. Field-aware factorization machines for CTR prediction. In *Proceedings of the ACM Recommender Systems Conference (RecSys)*, 2016. URL http://www.csie.ntu.edu.tw/~cjlin/papers/ffm.pdf.

[22] Wei-Lin Chiang, Mu-Chu Lee, and Chih-Jen Lin. Parallel dual coordinate descent method for large-scale linear classification in multi-core environments. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2016. URL http://www.csie.ntu.edu.tw/~cjlin/papers/multicore_cddual.pdf.

[23] Hsin-Yuan Huang and Chih-Jen Lin. Linear and kernel classification: When to use which? In *Proceedings of SIAM International Conference on Data Mining (SDM)*, 2016. URL http://www.csie.ntu.edu.tw/~cjlin/papers/kernel-check/kcheck.pdf.

[24] Mu-Chu Lee, Wei-Lin Chiang, and Chih-Jen Lin. Fast matrix-vector multiplications for large-scale logistic regression on shared-memory systems. In *Proceedings of the IEEE International Conference on Data Mining (ICDM)*, 2015. URL http://www.csie.ntu.edu.tw/~cjlin/papers/multicore_liblinear_icdm.pdf.

[25] Bo-Yu Chu, Chia-Hua Ho, Cheng-Hao Tsai, Chieh-Yen Lin, and Chih-Jen Lin. Warm start for parameter selection of linear classifiers. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2015. URL http://www.csie.ntu.edu.tw/~cjlin/libsvmtools/warm-start/warm-start.pdf.

[26] Wei-Sheng Chin, Yong Zhuang, Yu-Chin Juan, and Chih-Jen Lin. A learning-rate schedule for stochastic gradient methods to matrix factorization. In *Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD)*, 2015. URL http://www.csie.ntu.edu.tw/~cjlin/papers/libmf/mf_adaptive_pakdd.pdf.

[27] Yong Zhuang, Wei-Sheng Chin, Yu-Chin Juan, and Chih-Jen Lin. Distributed Newton method for regularized logistic regression. In *Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD)*, 2015.

[28] Chieh-Yen Lin, Cheng-Hao Tsai, Ching-Pei Lee, and Chih-Jen Lin. Large-scale logistic regression and linear support vector machines using Spark. In *Proceedings of the IEEE International Conference on Big Data*, pages 519–528, 2014. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/spark-liblinear/spark-liblinear.pdf`.

[29] Meng-Chieh Yu, Tong Yu, Shao-Chen Wang, Chih-Jen Lin, and Edward Y. Chang. Big data small footprint: The design of a low-power classifier for detecting transportation modes. *Proceedings of the VLDB Endowment*, 7:1429–1440, 2014.

[30] Cheng-Hao Tsai, Chieh-Yen Lin, and Chih-Jen Lin. Incremental and decremental training for linear classification. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2014. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/ws/inc-dec.pdf`.

[31] Tzu-Ming Kuo, Ching-Pei Lee, and Chih-Jen Lin. Large-scale kernel rankSVM. In *Proceedings of SIAM International Conference on Data Mining*, 2014. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/ranksvm/kernel.pdf`.

[32] Yong Zhuang, Wei-Sheng Chin, Yu-Chin Juan, and Chih-Jen Lin. A fast parallel SGD for matrix factorization in shared memory systems. In *Proceedings of the ACM Recommender Systems*, 2013. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/libmf.pdf`.

[33] Raffay Hamid, Dennis Decoste, and Chih-Jen Lin. Dense non-rigid point-matching using random projections. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.

[34] Aditya Khosla, Raffay Hamid, Chih-Jen Lin, and Neel Sundaresan. Large-scale video summarization using web-image priors. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.

[35] Guo-Xun Yuan, Chia-Hua Ho, and Chih-Jen Lin. An improved GLMNET for l1-regularized logistic regression. In *Proceedings of the Seventeenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 33–41, 2011.

[36] Chia-Hua Ho, Ming-Hen Tsai, and Chih-Jen Lin. Active learning and experimental design with SVMs. In *JMLR Workshop and Conference Proceedings: Workshop on Active Learning and Experimental Design*, volume 16, pages 71–84, 2011. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/activelearning/activelearning.pdf`.

[37] Hsiang-Fu Yu, Cho-Jui Hsieh, Kai-Wei Chang, and Chih-Jen Lin. Large linear classification when data cannot fit in memory. In *Proceedings of the Sixteenth ACM SIGKDD*

International Conference on Knowledge Discovery and Data Mining, pages 833–842, 2010. URL http://www.csie.ntu.edu.tw/~cjlin/papers/kdd_disk_decomposition.pdf.

[38] Fang-Lan Huang, Cho-Jui Hsieh, Kai-Wei Chang, and Chih-Jen Lin. Iterative scaling and coordinate descent methods for maximum entropy. In *Proceedings of the 47th Annual Meeting of the Association of Computational Linguistics (ACL)*, 2009. Short paper.

[39] Yin-Wen Chang and Chih-Jen Lin. Feature ranking using linear SVM. In *JMLR Workshop and Conference Proceedings: Causation and Prediction Challenge (WCCI 2008)*, volume 3, pages 53–64, 2008. URL http://www.csie.ntu.edu.tw/~cjlin/papers/causality.pdf.

[40] Yangqiu Song, Wen-Yen Chen, Hongjie Bai, Chih-Jen Lin, and Edward Y. Chang. Parallel spectral clustering. In *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD)*, 2008. URL http://www.csie.ntu.edu.tw/~cjlin/papers/ecml08.pdf.

[41] S. Sathiya Keerthi, Sellamanickam Sundararajan, Kai-Wei Chang, Cho-Jui Hsieh, and Chih-Jen Lin. A sequential dual method for large scale multi-class linear SVMs. In *Proceedings of the Forteenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 408–416, 2008. URL http://www.csie.ntu.edu.tw/~cjlin/papers/sdm_kdd.pdf.

[42] Cho-Jui Hsieh, Kai-Wei Chang, Chih-Jen Lin, S. Sathiya Keerthi, and Sellamanickam Sundararajan. A dual coordinate descent method for large-scale linear SVM. In *Proceedings of the Twenty Fifth International Conference on Machine Learning (ICML)*, 2008. URL http://www.csie.ntu.edu.tw/~cjlin/papers/cddual.pdf.

[43] Chih-Jen Lin, Ruby C. Weng, and S. Sathiya Keerthi. Trust region Newton method for large-scale logistic regression. In *Proceedings of the 24th International Conference on Machine Learning (ICML)*, 2007. Software available at http://www.csie.ntu.edu.tw/~cjlin/liblinear.

[44] Tzu-Kuo Huang, Chih-Jen Lin, and Ruby C. Weng. Ranking individuals by group comparisons. In *Proceedings of the Twenty Third International Conference on Machine Learning (ICML)*, 2006.

[45] Pai-Hsuen Chen, Rong-En Fan, and Chih-Jen Lin. Training support vector machines via smo-type decomposition methods. In *Proceedings of the 16th International Conference on Algorithmic Learning Theory (ALT 2005)*, pages 45–62, 2005.

[46] Tzu-Kuo Huang, Ruby C. Weng, and Chih-Jen Lin. A generalized Bradley-Terry model: From group competition to individual skill. In *Advances in Neural Information Processing Systems 17*. MIT Press, Cambridge, MA, 2005.

[47] Ting-Fan Wu, Chih-Jen Lin, and Ruby C. Weng. Probability estimates for multi-class classification by pairwise coupling. In Sebastian Thrun, Lawrence Saul, and Bernhard Schölkopf, editors, *Advances in Neural Information Processing Systems 16*. MIT Press, Cambridge, MA, 2004.

[48] Kai-Min Chung, Wei-Chun Kao, Tony Sun, and Chih-Jen Lin. Decomposition methods for linear support vector machines. In *Proceedings of ICASSP 2003*, pages 868–871, 2003.

[49] Ming-Wei Chang, Chih-Jen Lin, and Ruby C. Weng. Adaptive deterministic annealing for two applications: competing SVR of switching dynamics and travelling salesman problems. In *Proceedings of ICONIP 2002*, pages 920–924, 2002.

[50] Kai-Min Chung, Wei-Chun Kao, Tony Sun, Li-Lun Wang, and Chih-Jen Lin. Radius margin bounds for support vector machines with the RBF kernel. In *Proceedings of ICONIP 2002*, pages 893–897, 2002.

[51] Ming-Wei Chang, Chih-Jen Lin, and Ruby C. Weng. Analysis of nonstationary time series using support vector machines. In Seong-Whan Lee and Alessandro Verri, editors, *Proceedings of SVM 2002*, Lecture Notes in Computer Science 2388, pages 160–170, New York, NY, USA, 2002. Springer-Verlag Inc.

[52] Ming-Wei Chang, Chih-Jen Lin, and Ruby C. Weng. Analysis of switching dynamics with competing support vector machines. In *Proceedings of IJCNN*, pages 2387–2392, 2002.

[53] Chih-Chung Chang and Chih-Jen Lin. IJCNN 2001 challenge: Generalization ability and text decoding. In *Proceedings of IJCNN*. IEEE, 2001.

[54] Shuo-Peng Liao, Hsuan-Tien Lin, and Chih-Jen Lin. A note on the decomposition methods for support vector regression. In *Proceedings of IJCNN*, 2001.

[55] Chih-Chung Chang, Chih-Wei Hsu, and Chih-Jen Lin. The analysis of decomposition methods for support vector machines. In *Workshop on Support Vector Machines, IJCAI99*, 1999.

[56] Chih-Jen Lin, Nestor Michelena, and Romesh Saigal. Topological fixture synthesis using semidefinite programming. In *Proceedings of the Third World Congress of Structural and Multidisciplinary Optimization (WCSMO-3)*, May 17-21 1999.

[57] Chih-Jen Lin. Preconditioning dense linear systems from large-scale semidefinite programming problems. In *Proceedings of the Fifth Copper Mountain conference on iterative methods*, 1998.

- BOOK CHAPTERS

[1] Léon Bottou and Chih-Jen Lin. Support vector machine solvers. In Léon Bottou, Olivier Chapelle, Dennis DeCoste, and Jason Weston, editors, *Large Scale Kernel Machines*, pages 1–28. MIT Press, Cambridge, MA., 2007. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/bottou_lin.pdf`.

[2] Yi-Wei Chen and Chih-Jen Lin. Combining SVMs with various feature selection strategies. In Isabelle Guyon, Steve Gunn, Masoud Nikravesh, and Lofti Zadeh, editors, *Feature extraction, foundations and applications*. Springer, 2006.

[3] Soon-Yi Wu, Shu-Cherng Fang, and Chih-Jen Lin. Analytic center based cutting plane method for linear semi-infinite programming. In M. Goberna and M. Lopez, editors, *Semi-infinite programming: recent advances*. Kluwer, 2001.

[4] Chih-Jen Lin, Shu-Cherng Fang, and Soon-Yi Wu. A dual affine scaling based algorithm for solving linear semi-infinite programming problems. In D. Z. Du and J. Sun, editors, *Advances in Optimization and Application*, pages 217–234. Kluwer Academic Publishers, 1994.

- TECHNICAL REPORTS:

[1] Wei-Cheng Chang, Ching-Pei Lee, and Chih-Jen Lin. A revisit to support vector data description. Technical report, 2013. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/svdd.pdf`.

[2] Chih-Wei Hsu, Chih-Chung Chang, and Chih-Jen Lin. A practical guide to support vector classification. Technical report, Department of Computer Science, National Taiwan University, 2003. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf`.

[3] Hsuan-Tien Lin and Chih-Jen Lin. A study on sigmoid kernels for SVM and the training of non-PSD kernels by SMO-type methods. Technical report, Department of Computer Science, National Taiwan University, 2003. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/tanh.pdf`.

[4] Jen-Hao Lee and Chih-Jen Lin. Automatic model selection for support vector machines. Technical report, Department of Computer Science and Information Engineering, National Taiwan University, 2000.

[5] Chih-Jen Lin. *Study in Large-Scale optimization.* PhD thesis, University of Michigan, Ann Arbor, Michigan, 1998.

[6] Chih-Jen Lin and Romesh Saigal. A predictor corrector method for semi-definite linear programming. Technical report, Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI 48109-2117, 1995.

[7] Chih-Jen Lin and Romesh Saigal. An infeasible start predictor corrector method for semi–definite linear programming. Technical report, Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI 48109-2117, 1995.

- SOFTWARE

1. `LIBSVM`: an integrated software for support vector classification and regression, released April 2000. (with C.-C. Chang)
   (`http://www.csie.ntu.edu.tw/~cjlin/libsvm`)
   Around 53,600 Google Scholar citations (up to September 2023).
2. `LIBLINEAR`: a library for large linear classification, released April 2007. (with my research group)
   (`http://www.csie.ntu.edu.tw/~cjlin/liblinear`)
   Around 9,500 Google Scholar citations (up to September 2023).
3. `BSVM`: a decomposition method for large-scale support vector machines, released February 2000. (with C.-W. Hsu)
   (`http://www.csie.ntu.edu.tw/~cjlin/bsvm`)
4. `TRON`: a bound-constrained optimization software, released in May 1999. (with J. J. Moré)
   (`http://www.mcs.anl.gov/~more/tron`)
5. `ICFS`: an incomplete Cholesky factorization for sparse matrices, released August 1998. (with J. J. Moré)
   (`http://www.mcs.anl.gov/~more/icf`)

- INVITED TALKS AND MISCELLANEOUS PRESENTATIONS

1. "On the "rough use" of machine learning techniques." Keynote at the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval, Taipei, July 25, 2023
2. "Empirical Risk Minimization: Basic Concepts and Optimization Techniques." Invited talk at Machine Learning Research School, Thailand, August, 2023

3. "Stochastic gradient methods for deep learning." Keynote at the 34th International Meeting on Probabilistic, Combinatorial, and Asymptotic Methods for the Analysis of Algorithms, Taipei, June 26, 2023
4. "Algorithms and software for text classification." Invited talk at the 1st CASL Workshop at Mohamed bin Zayed University of Artificial Intelligence (MBZUAI), October 13, 2022
5. "Lessons from developing machine learning algorithms and systems." Keynote Talk at the ML-Labs Bootcamp, University College Dublin, Ireland, September 24, 2020 (virtual)
6. "Optimization and machine learning." Plenary talk at TWSIAM annual meeting, Taiwan, July 24, 2020
7. "Lessons from developing machine learning algorithms and systems." Keynote at Conference on Technologies and Applications of Artificial Intelligence (TAAI), Taiwan, December 1, 2017
8. "Large-scale linear classification: status and challenges." Invited talk at Criteo annual Machine Learning in the Real World Workshop, Paris, November 8, 2017.
9. "Statistical learning: status and challenges." Keynote at International Conference on Cognitive Science, Taipei, Taiwan, September 2, 2017
10. "When and When Not to Use Distributed Machine Learning." Keynote at Alibaba Data Mining Workshop, Hangzhou, China, June 29, 2017
11. "Optimization and machine learning." Invited talk at Summer School on Optimization, Big Data and Applications, Italy, July, 2017
12. "Matrix factorization and factorization machines for recommender systems." Keynote at the 4th Workshop on Large-Scale Recommender Systems, Boston, September 2016
13. "When and when not to use distributed machine learning?" Keynote at International Winter School on Big Data, Bilbao, Spain, February 2016
14. "Large-scale Linear and Kernel Classification." Invited talk at Microsoft Research India Summer School 2015 on Machine Learning, June 15, 2015
15. "Matrix factorization and factorization machines for recommender systems." Invited talk at SDM workshop on Machine Learning Methods on Recommender Systems, May 2, 2015.
16. "Big-data machine learning: status and challenges." Invited talk at China R Conference, Hangzhou, China, November 29, 2014
17. "Experiences and lessons in developing machine learning software." Invited talk at Industry Track, ACM Conference on Information and Knowledge Management (CIKM), Shanghai, November 4, 2014

18. "Large-scale linear classification: status and challenges." Invited talk at San Francisco Machine Learning Meetup, October 30, 2014
19. "Big-data machine learning." Invited speech at eBay Data Summit, Shanghai, China, October 25, 2014
20. "Big-data analytics: challenges and opportunities." Keynote speech at Taiwan Data Science Conference, Taipei, August 30, 2014.
21. "Distributed data classification." Invited talk at Workshop on New Learning Frameworks and Models for Big Data, ICML, June 25, 2014.
22. "Distributed data classification." Invited talk at Workshop on Scalable Data Analytics, PAKDD, May 13, 2014
23. "Large-scale machine learning." Invited talk at International Conference on Big Data and Cloud Computing, Xiamen, China, December 29, 2013.
24. "Distributed Newton methods for CTR (Click Through Rate) prediction." Invited talk at Mysore park workshop on distributed computing for machine learning and optimization, India, December 19, 2013.
25. "Recent advances in large-scale linear classification." Invited talk at Asian Conference on Machine Learning, November 15, 2013
26. "Experiences and lessons in developing machine learning and data mining software," Invited talk at China R Conference, Shanghai, China, November 2, 2013
27. "Optimization and machine learning." Plenary talk at 11th EUROPT Workshop on Advances in Continuous Optimization, Florence, Italy, June 26, 2013
28. "Optimization and machine learning," 25th Simon Stevin Lecture, K. U. Leuven Optimization in Engineering Center, Leuven, Belgium, January 17, 2013.
29. "Machine learning software: design and practical use," invited talk at Machine Learning Summer School (MLSS), Kyoto, August 2012.
30. "Experiences and lessons in developing industry-strength machine learning and data mining software," invited talk at Industry Practice Expo of ACM KDD 2012, Beijing, August 2012.
31. "Machine learning software: design and practical use," invited talk at Machine Learning Summer School (MLSS), Santa Cruz, July 2012.
32. "Large-scale machine learning in distributed environments," tutorial at ACM International Conference on Multimedia Retrieval, June, 2012
33. "Support vector machines and kernel methods," invited tutorial at Asian Conference on Machine Learning, Tokyo, Japan, November 8, 2010

34. "Support vector machines and kernel methods," plenary talk at International Workshop on Recent Trends in Learning, Computation, and Finance, Pohang, Korea, August 30, 2010.

35. "Training support vector machines: status and challenges," invited speaker at ICML 2008 Workshop on Large Scale Learning Challenge.

36. "Training support vector machines: status and challenges," invited speaker at Google Machine Learning Summit, May 2008.

37. "Support vector machines," invited tutorial speaker at Machine Learning Summer School (MLSS), Taipei, July 2006.

38. "Training linear and non-linear SVMs," invited talk at Workshop on Mathematics and Medical Diagnosis, Erice, Italy, July 2006.

39. "Support vector machines for data classification," invited tutorial at ICONIP 2005, Taiwan, October 30, 2005.

40. "Optimization issues in training support vector machines," the 16th international conference on Algorithmic Learning Theory, Singapore, October 9, 2005 (invited talk).

41. "Support vector machines for data classification," XXXVI Annual Conference of the Italian Operational Research Society, Camerino, Italy, September 8, 2005 (invited plenary talk).

42. "Generalized Bradley-Terry model and multi-class probability estimates," ISI (International Statistical Institute) 2005, Australia, April 6, 2005 (talk in an invited session).

43. "Report on NIPS 2003 Feature Selection Competition," NIPS workshop on feature selection competition, Canada, December 12, 2003.

44. "Optimization techniques for data mining and machine learning," invited talk in Workshop on Optimization and Control, National Cheng Kung University, Tainan, Taiwan, January 6, 2003.

45. "Support vector machines for time series segmentation," invited talk in the 2002 Taipei International Statistical Symposium and Bernoulli Society EAPR Conference, Taipei, July 7-10, 2002.

46. "Support vector machines for protein classification/prediction," invited talk at the 8th Symposium on Recent Advances in Biophysics, Taipei, May 23, 2002.

47. "Automatic model selection using the decomposition methods," NIPS workshop on kernel methods, Breckenridge, CO, December 1, 2000.

48. "Newton's method for support vector machines." Talk at the Sixth SIAM Conference on Optimization, Atlanta, May 1999.

49. "Structural optimization and semidefinite programming," Talk at INFORMS Fall meeting, Seattle, October 1998.

50. "Preconditioning dense linear systems from large-scale semidefinite programming problems," Talk at the Fifth Copper Mountain Conference on Iterative Methods, Copper Mountain, Colorado, April, 1998.

51. "Incomplete Cholesky factorizations with limited memory." Talk at the Fourth Kalamazoo Symposium on Matrix Analysis & Applications, Kalamazoo, MI, October, 1997.

52. "Newton's method for large bound-constrained optimization problems." Talk at International Symposium on Mathematical Programming, Lausanne, Switzerland , August, 1997.

53. "An unconstrained convex programming approach for solving linear semi-infinite programming problems." Talk at International Symposium on Mathematical Programming, Lausanne, Switzerland , August, 1997.

54. "An infeasible start predictor corrector method for semidefinite linear programming ." Talk at Fifth SIAM Optimization Conference, Victoria, British Columbia, Canada, May 1996.

- ACADEMIC SERVICES

1. Editorial services
   - Editorial board member, *ACM Transactions on Knowledge Discovery from DATA*, 2017–2020
   - Action editor, *Data Mining and Knowledge Discovery*, 2009–2017
   - Editorial board member, *ACM Transactions on Intelligent Systems and Technology*, 2012–2021
   - Associate editor, *IEEE Transactions on Neural Networks*, 2005–2010
   - Associate editor, *Journal of Information Science and Engineering*, 2009–2013
   - Guest editor: special issue on Support Vector Machines, *Neurocomputing*, 2003.
2. Reviewer for the following journals
   - *Journal of Machine Learning Research*
   - *Machine Learning*
   - *Neural Computation*
   - *SIAM Journal on Matrix Analysis and Applications*
   - *SIAM Journal on Optimization*
   - *IEEE Transactions on Neural Networks*

- *IEEE Transactions on Pattern Analysis and Machine Intelligence*
- *IEEE Transactions on Knowledge and Data Engineering*
- *IEEE Transactions on Big Data*
- *IEEE Transactions on Fuzzy Systems*
- *IEEE Transactions on Image Processing*
- *IEEE Transactions on Signal Processing*
- *IEEE Signal Processing Letters*
- *IEEE Transactions on Evolutionary Computation*
- *IEEE Transactions on Systems, Man, and Cybernetics*
- *IEEE Transactions on Semiconductor Manufacturing*
- *IEEE Transactions on Antennas and Propagation*
- *IEEE Transactions on Automation Science and Engineering*
- *IEEE Transactions on Audio, Speech and Language Processing*
- *Biometrika*
- *Neurocomputing*
- *Bioinformatics*
- *BMC Bioinformatics*
- *Theory of Computing Systems*
- *Neural Processing Letters*
- *Signal Processing*
- *International Journal of Pattern Recognition and Artificial Intelligence (IJPRAI)*
- *Artificial Intelligence Review*
- *Pattern Analysis & Applications*
- *Computational Intelligence and Neuroscience*
- *IIE Transactions*
- *Annals of the Institute of Statistical Mathematics*
- *Journal of Statistical Planning and Inference*
- *Statistics and Computing*
- *Communications in Statistics*
- *Pattern Recognition*
- *Pattern Recognition Letters*
- *Knowledge and Information Systems*
- *Computational Optimization and its Applications*

- *INFORMS Journal on Computing*
- *Journal of Global Optimization*
- *Optimization*
- *Numerical Algorithms*
- *Information Processing and Management*
- *Internet Electronic Journal of Molecular Design*
- *International Journal of Operations and Quantitative Management*
- *International Journal of Computer Mathematics*
- *Journal of Information Science and Engineering*
- *Journal of Computer Science and Technology (JCST)*
- *Journal of Formosan Medical Association*
- *Journal of Chinese Institute of Industrial Engineers*
- *Journal of the Chinese Institute of Engineers*
- *Journal of the Chinese Institute of Electrical Engineering*

3. Reviewer for several book chapters
4. Conference chair, Area chair, or senior program committee member:
   - Program co-chair, ACM SIGKDD international conference on Knowledge discovery and data mining (KDD), 2018
   - General co-chair, Mysore park workshop on distributed computing for machine learning and optimization, India, 2013
   - Steering committee member, Asian Conference on Machine Learning (ACML), 2015–2019
   - Senior PC, ACM SIGKDD international conference on Knowledge discovery and data mining (KDD), 2013, 2014, 2015, 2016, 2017, 2019, 2020
   - Senior PC, SIAM international conference on data mining (SDM), 2017
   - Area chair, Neural Information Processing Systems (NIPS) 2007, 2010, 2011, 2015
   - Area chair, International Conference on Machine Learning (ICML), 2016, 2017
   - General co-chair, Asian Conference on Machine Learning (ACML) 2011
   - Senior PC, Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD), 2013, 2014, 2015, 2016
   - Senior PC, AAAI 2017
   - Senior PC, IJCAI 2011 (IEAI track)
   - Senior PC, Asian Conference on Machine Learning (ACML) 2010

5. Program committee member:
   - AAAI 2016
   - International Joint Conference on Artificial Intelligence (IJCAI) , 2015
   - Workshop on Large-Scale Recommender Systems at ACM RecSys, 2014
   - The International Workshop on advances in Regularization, Optimization, Kernel methods and Support vector machines: theory and applications (ROKS-2013), Belgium.
   - ACM SIGKDD international conference on Knowledge discovery and data mining (KDD), Washington D.C. 2010, San Diego, 2011, Beijing, 2012
   - SIAM international conference on data mining (SDM), 2013
   - AI & Statistics 2010, 2014
   - NIPS Workshop on Optimization for Machine Learning (2008, 2009, 2010, 2011, 2012, 2013)
   - NIPS Workshop on AutoML (2014)
   - International Conference on Machine Learning (ICML), Helsinki 2008, Montreal 2009, Haifa 2010, Bellevue, WA 2011, Scotland, 2012, Atlanta, 2013, Beijing, 2014
   - European Conference on Machine Learning (ECML) and European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD), 2008, 2010, 2011
   - International Joint Conference on Neural Networks (IJCNN), Hong Kong 2008, San Jose, CA 2011
   - Pacific-Rim Conference on Multimedia (PCM), Hong Kong 2007, Bangkok, Thailand 2009, Shanghai, China 2010.
   - IEEE International Conference on Multimedia & Expo (ICME), Beijing 2007, Hannover 2008.
   - Asian Conference on Machine Learning (ACML), 2009
   - NIPS 2006 Workshop on Machine Learning Open Source Software.
   - ACM Multimedia Conference (ACM MM), Santa Barbara 2006
   - International Colloquium on Grammatical Inference (ICGI), Japan 2006
   - International workshops on Statistical Techniques in Pattern Recognition (SPR), Hong Kong 2006, Orlando, Florida, 2008, Turkey, 2010.
   - Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD), Singapore 2006, China 2007, Osaka, Japan 2008, Thailand 2009
   - International Conference on Neural Information Processing (ICONIP), India 2004, Hong Kong 2006

- International Workshop on Pattern Recognition with Support Vector Machines (SVM2002), Canada
- Fourth Asia-Pacific Conference on Industrial Engineering and Management Systems, 2002, Taiwan

6. Reviewer for the following conferences
   - Neural Information Processing Systems (NIPS), 2003, 2004, 2005, 2006, 2014, 2016
   - Conference on Learning Theory (COLT), 2003, 2009
   - International Joint Conference on Neural Networks (IJCNN), 2003, 2004, 2005
   - IEEE International Conference on Multimedia & Expo (ICME), 2009
   - First Asia-Pacific Bioinformatics Conference, Australia, 2003
   - The Seventh Pacific Rim International Conference on Artificial Intelligence, (PRICAI-02)
7. Other conference planning and administration
   - Special session organizer, ICONIP 2002, Singapore
8. Thesis External Reviewers:
   - The University of British Columbia: Huang Fang (Ph.D. 2021)
   - University of Technology Sydney: Yuangang Pan (Ph.D. 2019)
   - Nanyang Technological University: Mingkui Tan (Ph.D. 2014)
   - University of Trento, Italy: Nicola Segata (Ph.D. 2009)
   - Australian National University: Jin Yu (Ph.D. 2009)
   - Ruhr-Universität Bochum: Tobias Glasmachers (Ph.D. 2008)
   - Hongkong University of Science and Techonology: Ivor Tsang (Ph.D. 2007)
   - National University of Singapore: Chu Wei (Ph.D. 2003), Kaibo Duan (Ph.D. 2003), Jianbo Yang (Ph.D., 2011)
   - Chinese University of Hongkong: Wan Zhang (M. Phil. 2003)
9. Proposal Reviewers:
   - Research Grants Council, Hong Kong, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2016, 2022, 2023
   - American University of Beirut, 2009
   - Czech Science Foundation, 2010
10. Other Services:
    - Member of the strategic planning committee, Taiwan AI Center of Excellence, October 2022 – September 2024

- IEEE CS society fellow evaluation committee member (2011, 2012)

- TALKS IN ACADEMIC INSTITUTES AND INDUSTRY

  - International:

    1. A*STAR Centre for Frontier AI Research, Singapore, April 1, 2024
    2. Fudan University, November 15, 2023
    3. Shanghai Jiao Tong University, November 14, 2023
    4. Chulalongkorn University, Bangkok, Thailand, August 3, 2023
    5. Mohamed bin Zayed University of Artificial Intelligence, Abu Dhabi, UAE, April 4, 2023
    6. Bloomberg, New York, November 14, 2022 (virtual)
    7. Mohamed bin Zayed University of Artificial Intelligence, Abu Dhabi, UAE, October 13, 2022
    8. Cedars-Sinai Medical Center, CA, January 30, 2020
    9. Statistics Department, University of California, Los Angeles, CA, November 12, 2019
    10. Huawei research, Shenzhen, April 17, 2019
    11. University of California, Los Angeles, CA, November 6, 2018
    12. University of Florence, Florence, Italy, October 26, 2018
    13. Alibaba, Sunnyvale, California, May 1, 2018
    14. Lightinthebox, Beijing, China, September 27, 2017
    15. Criteo, California, April 26, 2017
    16. Google, California, April 25, 2017
    17. LinkedIn, California, April 24, 2017
    18. Huawei, Shenzhen, China, April 19, 2017
    19. University of Technology Sydney, Australia, April 6, 2017
    20. University of Waikato, New Zealand, April 5, 2017
    21. Tencent, Shenzhen, China, March 15, 2017
    22. Alibaba, Hangzhou, China, February 28, 2017
    23. Nanyang Technological University, Singapore, January 16, 2017
    24. Microsoft, Redmond, Washington, October 6, 2016
    25. Guangdong University of Technology, Guangzhou, China, June 20, 2016
    26. Samsung Research America, California, June 10, 2016
    27. UC Davis, California, May 4, 2016

28. Netflix, California, May 3, 2016
29. Huawei Research Labs, Shenzhen, China, April 19, 2016
30. Chinese University of Hong Kong, Shenzhen, China, April 18, 2016
31. Facebook, California, November 13, 2015
32. Quora, California, November 12, 2015
33. Nanjing University, December 25, 2014
34. University of Electronic Science and Technology of China, Chengdu, China, November 30 and December 1, 2014 (two talks)
35. Twitter, California, October 31, 2014
36. eBay China, October 24, 2014
37. Microsoft Research, New York City, August 22, 2014 (open machine learning software workshop)
38. Microsoft, Redmond, Washington, August 18, 2014
39. Criteo, California, August 1, 2014
40. Databricks, California, July 31, 2014
41. Research Center on Fictitious Economy and Data Science, Chinese Academy of Sciences, Beijing, China, June 27, 2014
42. Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, Beijing, China, June 26, 2014
43. Samsung Research America, California, May 23, 2014
44. Walmart Labs, California, April 23, 2014
45. Pandora, California, April 22, 2014
46. Alibaba, Hangzhou, China, December 27, 2013
47. eBay China, November 1, 2013
48. Shanghai Jiao Tong University, October 31, 2013
49. Microsoft Research, Redmond, August 15, 2013
50. University of Rome "La Sapienza," June 25, 2013
51. K. U. Leuven, Belgium, January 14-16, 2013
52. Baidu, China, October 24, 2012
53. Luminescent technology, California, August 24, 2012
54. eBay Machine Learning Forum, San Jose, California, February 17, 2012
55. City University of Hong Kong, December 30, 2011
56. NEC Labs, Cupertino, California, August 26, 2011
57. Adobe, California, August 25, 2011

58. eBay research, San Jose, California, December 7, 2010
59. Facebook, Palo Alto, California, December 6, 2010
60. Baidu, China, September 3, 2010
61. Google Research New York, July 29, 2010
62. Yahoo! Research, Santa Clara, California, July 23, 2010
63. China Agriculture University, October 16, 2009
64. Microsoft Research Asia, October 13, 2009
65. Department of Computer Science and Engineering, Hong Kong University of Science and Technology, February 5, 2009
66. Department of Computer Science and Technology, Tsinghua University, China, September 5, 2008
67. HP Labs China, June 26, 2008
68. IBM T. J. Watson Research Center, May 16, 2008
69. Department of Industrial and Operations Engineering, University of Michigan, August 15, 2007
70. Yahoo! Research, Santa Clara, California, February 20, 2007
71. NEC Labs, Princeton, New Jersey, February 15, 2007
72. Siemens Corporate Research, Princeton, New Jersey, February 14, 2007
73. AT&T Research, February 13, 2007
74. California Institute of Technology, November 14, 2006
75. School of Information and Computer Science, University of California, Irvine, November 6, 2006
76. Yahoo! Research, Burbank, California, August 30, 2006
77. Mathematics and computer science division, Argonne National Lab., June 23, 2006
78. Chinese University of Hong Kong, Hone Kong, December 12, 2005
79. Nanyang Technological University, Singapore, October 10, 2005
80. Università di Roma "La Sapienza" and Isituto di Analisi dei Sistemi ed Informatica del CNR, Italy, September 1-2, 2005 (a short course).
81. CWI (Dutch National Research Institute for Mathematics and Computer Science), February 9, 2004
82. Department of Electronics and Computer Science, University of Southampton, February 2-6, 2004 (two talks)
83. Department of Computer Science, University of Essex, January 22, 2004

84. Department of Statistics and Probability Theory, Vienna University of Technology, September 4, 2003
85. Fraunhofer Institute for Computer Architecture and Software Technology, Germany, August 18, 2003
86. Department of Computer Science, University of Essex, August 13, 2003
87. University of Freiburg, Germany, July 15, 2003
88. Max Planck Institute of Biological Cybernetics, Germany, July 9, 2003
89. University of Tuebingen, Germany, July 8, 2003
90. KXEN Corporation, Suresnes, France, February 17, 2003
91. Max Planck Institute of Informatics (Computer Science), Germany, February 10-16, 2003 (two talks)
92. Max Planck Institute of Biological Cybernetics, Germany, January 12-February 10, 2003 (three talks)
93. Department of Electrical and Computer Engineering, University of Michigan-Dearborn, August 27, 2002
94. Siemens Corporate Research, Princeton, New Jersey, August 21, 2002
95. Department of Computer Science, Binghamton University, August 19, 2002
96. Merck research Lab., New Jersey, August 16, 2002
97. Agilent Inc., Colorado, July 31, 2001
98. Ford Research Lab., Michigan, July 24, 2001
99. Department of Electrical Engineering, Ohio State University, August 29, 2000.

– Domestic:

1. Department of Computer Science, University of Taipei, March 27, 2024
2. College of science, National Chung Hsing University, April 14, 2022
3. Institute of Statistical Science, Academia Sinica, March 21, 2022
4. Asus Inc., January 18, 2019
5. Department of Computer Science, Providence University, May 15, 2018
6. National Center for Theoretical Sciences, Mathematics Division, March 9, 2018
7. Institute of Statistics, National Tsing Hua University, April 29, 2016
8. Industrial Technology Research Institute, October 7, 8, 21, and 22, 2015 (a short course on data mining)
9. Institute of Communication Engineering, National Tsing Hua University, October 17, 2014
10. Department of Computer Science, University of Taipei, October 15, 2014

11. Industrial Technology Research Institute, July 18, 22, 24, and August 12, 2014 (a short course on data mining)
12. Interdisciplinary Science Program, National Chiao Tung University, March 28, 2014
13. Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, September 24, 2012
14. Department of Mathematics, National Taiwan University, October 17, 2011
15. Department of Financial and Computational Mathematics, Providence University, September 22, 2011
16. Department of Mathematics, National Taiwan Normal University, April 20, 2011
17. Department of Applied Informatics, Fo Guang University, April 14, 2011
18. Department of Information Management, National Taiwan University, February 25, 2011
19. Institute of Information Science, Academia Sinica, February 16, 2011
20. Department Computer Science and Information Engineering, Chaoyang University of Technology, October 29, 2010
21. Graduate Institute of Communication Engineering, National Taiwan University, September 27, 2010
22. Department Computer Science and Information Engineering, National Central University, November 12, 2008
23. Department of Information Management, Chaoyang University of Technology, October 30, 2007
24. Department Computer Science and Information Engineering, National Cheng-Kung University, October 26, 2007
25. Department of Computer Science, National Chengchi University, November 10, 2005
26. Department of Computer Science, National Chi-Nan University, September 24, 2004
27. Institute of Information Science, Academia Sinica, April 15, 2004
28. Department of Statistics, National Chiao Tung University, April 9, 2004
29. Computer and Communications Research Laboratories, Industrial Technology Research Institute, February 27 and March 3, 2004 (8 hours)
30. Computer and Communications Research Laboratories, Industrial Technology Research Institute, November 18, 2003
31. Department of Information Management, Chaoyang University of Technology, November 4, 2003
32. Graduate Institute of Industrial Engineering, National Taiwan University, April 23, 2003

33. Department Mathematics, National Taiwan University, March 10, 2003
34. Department of Information and Computer Engineering, Chung Yuan Christian University, December 16, 2002
35. Department of Statistics, Feng Chia University, November 1, 2002
36. Department of Statistics, National Chengche University, October 14, 2002
37. Asian BioInnovations Corporation, Taipei, June 14, 2002
38. Graduate Program in Bioinformatics, National Yang Ming University, March 29, 2002
39. Department of Information Science and Management, Providence University, March 22, 2002
40. Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, March 11, 2002
41. Institute of Statistical Science, Academia Sinica, January 16, 2002
42. Department Mathematics, National Taiwan University, January 5, 2002
43. Institute of Computer Science and Information Engineering, Chang Gung University, December 4, 2001.
44. Graduate Institute of Medical Informatics, Taipei Medical University, November 22, 2001.
45. Department of Information Management, National Taichung Institute of Technology, October 23, 2001.
46. Graduate Institute of Industrial Engineering, National Taiwan University, October 3, 2001
47. Department of Information Management, National Taiwan University of Science and Technology, September 27, 2001
48. Department of Biological Science and Technology, National Chiao Tung University, September 26, 2001
49. Institute of Information Science, Academia Sinica, August 28-29, 2001
50. Department Computer Science and Information Engineering, National Cheng-Kung University, May 25, 2001
51. Department of Information and Computer Education, National Taiwan Normal University, April 9, 2001
52. Institute of Statistical Science, Academia Sinica, February 19, 2001
53. Department Computer Science and Information Engineering, National Central University, January 17, 2001

54. Division of Biostatistics and Bioinformatics, National Health Research Institutes, December 6, 2000
55. Institute of Biochemistry, National Yang-Ming University, June 5, 2000
56. Department Computer Science and Information Engineering, National Chung Cheng University, May 22, 2000
57. Institute of Information Science, Academia Sinica, November 19, 1999
58. Department of Computer Science, National Tsing Hua University, June 2, 1999
59. Department Computer Science and Information Engineering, National Taiwan University, March 5, 1999
60. Department of Industrial Engineering, National Tsing Hua University, December 24, 1998
61. Department Computer Science and Information Engineering, National Taiwan University, December 26, 1997
62. Department Mathematics, National Cheng-Kung University, December 19, 1997
63. Institute of Information Management, National Chi-Nan University, December 18, 1997
64. Department of Industrial Engineering, National Tsing Hua University, December 17, 1997
65. Department Mathematics, National Cheng-Kung University, May, 1997

- TEACHING EXPERIENCE

1. Optimization methods for deep learning (Spring 2019 UCLA, Winter 2020, Winter 2021, Winter 2023)
2. Optimization and machine learning (Fall 2010, Fall 2011, Winter 2014, Fall 2015, Winter 2017, Winter 2018)
3. Introduction to the theory of computation (Fall 2003, Fall 2004, Fall 2005, Fall 2007, Fall 2008, Fall 2009, Fall 2010, Fall 2011, Fall 2012, Fall 2013, Fall 2014, Fall 2015, Fall 2017, Fall 2018, Fall 2020, Fall 2021, Fall 2022, Fall 2023)
4. Numerical methods (Winter 2001, Winter 2002, Winter 2003, Winter 2009, Winter 2010, Winter 2011, Winter 2013, Winter 2014, Winter 2016, Winter 2017, Winter 2018, Winter 2021, Winter 2022, Winter 2023)
5. Machine learning: theory and practice (Winter 2007, Winter 2010, Winter 2013)
6. Data mining and machine learning (Fall 2001, Fall 2002, Winter 2004, Winter 2005, Winter 2006, Winter 2007)

7. Statistical learning theory (Fall 1999, Fall 2000, Fall 2001, Fall 2002, Fall 2003, Fall 2004, Fall 2005)
8. Scientific computing (Winter 1999, Winter 2000)
9. Operations Research (Fall 1998, Fall 1999, Fall 2000)

- CONSULTING SERVICES

1. Taboola, 2022 – present
2. Asus, November 2019 – August 2023
3. HTC, 2012–2014

- MEMBERSHIPS: IEEE (fellow), ACM (fellow), AAAI (fellow)