Dong Ryeol Lee — Curriculum vitae

CONTACT — E-mail: drselee@gmail.com WWW: http://sites.google.com/view/dongryeollee

BIO

I am a researcher with expertise in large scale machine learning and with experience working on projects in aviation, finance, internet, and automotive sectors. I hold a Ph.D. in CS and a M.S. in Math, and am currently an Adjunct Associate Professor at Columbia University.

Previously I was a senior data scientist at Faraday Future and consulted on machine learning requirements from automotive manufacturing, validation, and sales departments. Prior to this I was a scientist at Yahoo Research working on web search relevance problems. Prior to this I worked as a researcher at GE Global Research with internal stakeholders on building software infrastructure for fleet level asset prognostics and capital finance.

RESEARCH INTERESTS

Computational geometry, computational science, computational statistics, data structures, high-performance computing, machine learning, numerical linear algebra

EDUCATION

Georgia Institute of Technology, Atlanta, GA. **August 2005 – May 2012**

Ph.D., Computer Science, May 2012 GPA: 3.85 (4.0 scale)

Minor: Optimization and Statistics.
Thesis: *A Distributed Kernel Summation Framework for Machine Learning and Scientific Simulations.*. Advised by Alexander G. Gray.
Maintainer of **MLPACK** http://mlpack.org
DHS Graduate Fellowship, August 2006 – August 2009.
Upsilon Pi Epsilon CS International Honor Society, Fall 2007.

M.S., Mathematics, May 2011

Carnegie Mellon University, Pittsburgh, PA. **August 2001 – May 2005**

B.S., Computer Science, May 2005 GPA: 3.87 (4.0 scale)

Graduation with university and college honors.
Thesis: *New algorithmic techniques for generalized n-body problems*.
Dean's List for 6 out of 8 semesters, Fall 2001 – Spring 2005.
National Society of Collegiate Scholars inductee, Fall 2002.
Phi Beta Kappa inductee, Spring 2005.
Phi Kappa Phi inductee, Spring 2005.
Senior Leadership Award, Spring 2005.
University Scholarship, Fall 2001 – Spring 2005.

B.S., Mathematical Sciences, May 2005

PROFESSIONAL EXPERIENCE

Adjunct Associate Professor at Columbia, New York, NY. **January 2020 – Present**

Taught COMS W4721 Machine Learning for Data Science to students and industry professionals.

Data Scientist at Faraday Future, Gardena, CA. **July 2016 – August 2019**

Prognostics and diagnostics for electric vehicle.

Scientist at Yahoo Research, Sunnyvale, CA. **January 2015 – July 2016**

Relevance ranking for Yahoo web search.

Scientist at GE Global Research, Niskayuna, NY. **July 2012 – January 2015**

Prognostics and diagnostics for various GE assets; capital finance.

Co-founder at Analytics 1305, Atlanta, GA. **March 2009 – August 2010**

Startup consulting on industrial and business problems.

Teaching Experience

Carnegie Mellon University, Pittsburgh, PA.

Grader for Department of Mathematical Sciences **August 2004 – December 2004**

Grader for 21-355 Principles of Real Analysis I.

Teaching Assistant for School of Computer Science **January 2004 – May 2004**

Held office hours and graded assignments for *15-113 System Skills in C*.

Carnegie Mellon University Academic Development **August 2002 – May 2005**

Tutored introductory/advanced courses in mathematics and computer science. College Reading & Learning Association Level 3 Master certification.

Journal Publications

[JMLR 2015] R. R. Curtin, D. Lee, W. B. March, and P. Ram. Plug-and-play dual-tree algorithm runtime analysis. *The Journal of Machine Learning Research*, 16(1):3269–3297, 2015

[SAM 2013] D. Lee, P. Sao, R. Vuduc, and A. G. Gray. A distributed kernel summation framework for general-dimension machine learning. *Statistical Analysis and Data Mining*, 7(1):1–13, 2014

[JCP 2012] D. Lee, A. Ozakin, and A. G. Gray. Multibody multipole methods. *Journal of Computational Physics*, 231(20):6827–6845, 2012

Book Chapters

[AMLDMA 2012] W. B. March, A. Ozakin, D. Lee, R. Riegel, and A. G. Gray. Multitree algorithms for large-scale astrostatistics. In *Advances in Machine Learning and Data Mining for Astronomy*, pages 463–483. CRC Press, 2012

Conference Publications

[NeurIPS 2012] N. Mehta, D. Lee, and A. G. Gray. Minimax multi-task learning and a generalized loss-compositional paradigm for mtl. In *Advances in Neural Information Processing Systems*, pages 2150–2158, 2012

[SC 2012] W. B. March, K. Czechowski, M. Dukhan, T. Benson, D. Lee, A. J. Connolly, R. Vuduc, E. Chow, and A. G. Gray. Optimizing the computation of n-point correlations on large-scale astronomical data. In *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis*, page 74. IEEE Computer Society Press, 2012

[CVPR 2012] K. Kim, D. Lee, and I. Essa. Detecting regions of interest in dynamic scenes with camera motions. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1258–1265. IEEE, 2012

[SDM 2012B] P. Ram, D. Lee, and A. G. Gray. Nearest-neighbor search on a time budget via max-margin trees. In *Proceedings of the 2012 SIAM International Conference on Data Mining*, pages 1011–1022. SIAM, 2012

[SDM 2012A (Best paper)] D. Lee, R. Vuduc, and A. G. Gray. A distributed kernel summation framework for general-dimension machine learning. In *Proceedings of the 2012 SIAM International Conference on Data Mining*, pages 391–402. SIAM, 2012

[ICCV 2011] K. Kim, D. Lee, and I. Essa. Gaussian process regression flow for analysis of motion trajectories. In *2011 International Conference on Computer Vision*, pages 1164–1171. IEEE, 2011

[NeurIPS 2009B (Poster spotlight)] P. Ram, D. Lee, W. March, and A. G. Gray. Linear-time algorithms for pairwise statistical problems. In *Advances in Neural Information Processing Systems*, pages 1527–1535, 2009

[NeurIPS 2009A] P. Ram, D. Lee, H. Ouyang, and A. G. Gray. Rank-approximate nearest neighbor search: Retaining meaning and speed in high dimensions. In *Advances in Neural Information Processing Systems*, pages 1536–1544, 2009

[NeurIPS 2008] D. Lee and A. G. Gray. Fast high-dimensional kernel summations using the monte carlo multipole method. In *Advances in Neural Information Processing Systems*, pages 929–936, 2009

[AISTATS 2007] P. Wang, D. Lee, A. Gray, and J. M. Rehg. Fast mean shift with accurate and stable convergence. In *Artificial Intelligence and Statistics*, pages 604–611, 2007

[UAI 2006] D. Lee and A. Gray. Faster gaussian summation: theory and experiment. In *Proceedings of the Twenty-Second Conference on Uncertainty in Artificial Intelligence*, pages 281–288. AUAI Press, 2006

[NeurIPS 2005] D. Lee, A. W. Moore, and A. G. Gray. Dual-tree fast gauss transforms. In *Advances in Neural Information Processing Systems*, pages 747–754, 2006

PROGRAMMING LANGUAGES — C, C++, MapReduce framework (Spark, Flink), OpenMPI, Python

LANGUAGES — Fluent in English and Korean. Conversational level in Japanese and Mandarin.