David C. Kale

Data Scientist and Ph.D. Student
USC Viterbi School of Engineering
Los Angeles, CA 90089
http://www-scf.usc.edu/~dkale/

dkale@usc.edu
davekale@cs.stanford.edu
479-769-0570
Linkedin: http://www.linkedin.com/in/davekale
Github: http://github.com/turambar

Research Interests

My research aims to use machine learning to extract insight from digital data in high impact domains, including but not limited to health care. My primary interest is in developing robust methods for mining and modeling multivariate time series, which are increasingly common, especially in medicine and mobile health. Currently, I am working on techniques to incorporate structured domain knowledge and causal inference into deep learning architectures, as well as on making neural networks more interpretable (essential for domains like medicine). I have also worked on kernelized locality-sensitive hashing to enable fast similarity search over multivariate time series data, and on active learning and transfer learning, two complementary paradigms for performing supervised learning in the absence of carefully prepared and labeled data sets.

Education

2012- ★ **Ph.D., Computer Science – University of Southern California**
Advisor: Yan Liu, Topic: Machine Learning. *Expected graduation: 2017*
Viterbi Dean's Doctoral Fellow and *Alfred E. Mann Innovation in Engineering Doctoral Fellow*

2011 ★ **M.S., Computer Science – Stanford University**
Specialization: Artificial Intelligence

2005 ★ **B.S. with distinction, Symbolic Systems – Stanford University**
Inducted into *Phi Beta Kappa* Academic Honor Society.

Select Professional and Research Experience

2014 ★ **Research Intern – Microsoft Research, Redmond**
Working with Scott Sapponas to develop cutting edge wearable health sensors for continuous monitoring of cardiovascular health.

2014- ★ **Judge – Qualcomm Tricorder XPRIZE**
I am a judge in the $10 million global competition to build a first-generation *tricorder*, a self-contained health diagnostic device easy enough to be used by non-expert consumers.

2011- ★ **Co-founder and Advisor – Podimetrics**
Boston-area medical device start-up building a sensor and data analytics platform for detection and prevention of foot ulcers in patients suffering from diabetic neuropathy.

2009-2013 ★ **Lead Data Scientist – Whittier VPICU, Children's Hospital Los Angeles**
Lead data science team in application of statistical learning to complex clinical data sets.

2006–2009 ★ **Research Assistant – Stanford Artificial Intelligence Lab**
Worked on projects in areas such as computer vision, robotics, automated driving, and natural language processing.

2005 ★ **Engineering Intern – Google, Inc.**
Extended "Mustang" (Google's web search framework at the time) to handle structured data as a part of Froogle and Google Base teams.

Research and Publications

Refereed computer science

- Z. Che,* D. Kale,* W. Li, M.T. Bahadori, and Y. Liu. *Deep Computational Phenotyping*. Proceedings of the *21st ACM International Conference on Knowledge Discovery and Data Mining* (SIGKDD), 2015.
- M.T. Bahadori, D. Kale, Y. Fan, and Y. Liu. *Functional Subspace Clustering with Application to Time Series.* Proceedings of the *31st International Conference on Machine Learning* (ICML), 2015.
- D. Kale, M. Ghazvininejad, A. Ramakrishna, J. He, and Y. Liu. *Hierarchical Active Transfer Learning*. Proceedings of the 2015 *SIAM International Conference on Data Mining* (SDM), 2015.
- D. Kale,* D. Gong,* Z. Che,* G. Medioni, R. Wetzel, P. Ross, and Y. Liu. *An Examination of Multivariate Time Series Hashing with Applications to Health Care.*. Proceedings of the *IEEE 14th International Conference on Data Mining (ICDM)*, 2014.
- D. Kale and Y. Liu. *Accelerating Active Learning with Transfer Learning*. Proceedings of the *IEEE 13th International Conference on Data Mining (ICDM)*, 2013.
- D. Kale, S. Di, Y. Liu, and Y. Gil. *Capturing Data Analytics Expertise with Visualization in Workflows*. *AAAI Fall Symposium Series Discovery Informatics Workshop (DIS)*, 2013.
- D. Kale and D. Stork. *Estimating the Position of Illuminants in Paintings Under Weak Model Assumptions: An Application to the Works of Two Baroque Masters*. In B. E. Rogowitz and T. N. Pappas (eds.), *Electronic Imaging: Human Vision and Electronic Imaging* XIV, vol. 7240, pp. 72401M112. SPIE/IS&T, Bellingham, 2009.

Refereed health and medical

- D. Kale, Z. Che, Y. Liu, and R. Wetzel. *Computational discovery of physiomes in critically ill children using deep learning*. 1st Workshop on Data Mining for Medical Informatics: Electronic Phenotyping, *American Medical Informatics Assocation*, 2014.
- D. Epstein, M. Reibel, J. B. Unger, M. Cockburn, L. A. Escobedo, D. Kale, J. C. Chang, and J. I. Gold. *The Effect of Neighborhood and Individual Characteristics on Pediatric Critical Illness*. To appear in *Journal of Community Health*: Feb. 2014.
- E. B. Celikkaya, C. Shelton, D. Kale, R. Wetzel, and R. Khemani. *Non-invasive Blood Gas Estimation for Pediatric Mechanical Ventilation*. Proceedings of the *NIPS 2013 Workshop on Machine Learning for Clinical Data Analysis and Healthcare*.
- R. G. Khemani, E. B. Celikkaya, C. R. Shelton, D. Kale, P. A. Ross, R. C. Wetzel, and C. J. L. Newth. *Algorithms to estimate $PaCO_2$ and pH using non invasive parameters for children with Hypoxemic Respiratory Failure*. In *Respiratory Care*, Dec. 2013.

- E. Ingram, D. Kale, and R. Balfour. *Hemilaminectomy for thoracolumbar Hansen Type I intervertebral disk disease in ambulatory dogs with or without neurologic deficits*. *Journal of Veterinary Surgery*: v. 42, #8, pg. 924-931, Nov. 2013.

- B. Stubbs and D. Kale. *Sim*TwentyFive: An Interactive Visualization System for Data-Driven Decision Support*. Proceedings of the *American Medical Informatics Association Annual Symposium (AMIA)*, 2012.

- B. Marlin, D. Kale, R. Khemani, and R. Wetzel. *Unsupervised Pattern Discovery in Electronic Health Care Data Using Probabilistic Clustering Models*. Proceedings of the *2nd ACM SIGHIT International Health Informatics Symposium (IHI)*: pg. 389-98, 2012.

- D. Crichton, C. Mattmann, A. Hart, D. Kale, R. Khemani, P. Ross, S. Rubin, P. Veeravatanayothin, A. Braverman, C. Goodale, and R. Wetzel. *An Informatics Architecture for the Virtual Pediatric Intensive Care Unit*. In proceedings of the *24th IEEE International Symposium on Computer-Based Medical Systems (CBMS)*, pages 1-6, 2011.

- D. Kale, A. Hart, C. Mattmann, R. Khemani, P. Ross, P. Vee, J. Terry, R. Wetzel, and D. Crichton. *An Open Source, Grid-based Software Framework for Management and Sharing of Pediatric ICU Data*. In proceedings of the *9th International Conference on Complexity in Acute Illness (ICCAI)*, 2010.

Other

- B. Sankaran, D. Kale, X. He, L. Cohen, and M. Ghazvininejad. *Learning and Optimization with Submodular Functions*. arXiV.

- D. Kale. *Unsupervised Pattern Discovery in Sparsely Sampled Clinical Time Series*. At *From Data to Knowledge Workshop*, University of California Berkeley, 2012.

- D. Kale. *Probabilistic Modeling of Electronic Health Records Data*. AI/ML Weekly Seminar at University of California, Irvine, Center for Machine Learning and Intelligent Systems. October 2011.

- D. Kale, B. Marlin, R. Khemani, and R. Wetzel. *A Novel Application of Unsupervised Learning to Electronic Health Care Records Data*. At *1st Southern California Workshop on Machine Learning (SoCaML 2011)*, 2011.

- D. Kale, B. Marlin, R. Khemani, and R. Wetzel. *Using Probabilistic Clustering to Find Patterns in Digital Medical Data*. At *1st Meaningful Use of Complex Medical Data Symposium (MUCMD 2011)*, 2011.

- A. Hart, D. Kale, R. Khemani, and H. Kincaid. *Distributed, Modular Grid Software for Data Management and Exploration of Patient-Centric Healthcare IT Information*. In proceedings of the OReilly Open Source Convention: Special Session on Healthcare Technology (OSCON 2010), 2010.

Funded Grants

2009 ★ **Advanced Computational Framework for Decision Support in Critically Ill Children**
Co-PI, $1,000,000 to Children's Hospital Los Angeles (NASA JPL sub-contractor)

Teaching Experience

Teaching Assistant	USC	CSCI 109: Introduction to Computing
	Stanford	CS 221: Introduction to Artificial Intelligence CS 121: Introduction to Artificial Intelligence CS 103A: Discrete Mathematics for Computer Science CS 103B: Discrete Structures
Section Leader	Stanford	CS 107: Computer Organization and Systems

Professional Activities and Awards

2015 ★ **NSF Travel Award – 21st ACM Int'l Conf. on Knowl. Disc. and Data Mining (SIGKDD)**

2014 ★ **Student Travel Award – IEEE 14th Int'l Conference on Data Mining (ICDM)**

2014 ★ **Alfred E. Mann *Innovation in Engineering* Fellowship**
One of three fellows conducting innovative interdisciplinary research.

2014 ★ **Organizing Committee – AAAI 1st Modern AI for Health Analytics Workshop**

2014 ★ **Program Committee – AAAI Discovery Informatics Workshop**

2013 ★ **NSF Travel Award – IEEE 13th Int'l Conference on Data Mining (ICDM)**

2013 ★ **Session Moderator – Health 2.0 7th Annual Fall Conference**

2013 ★ **Organizing Committee – Meaningful Use of Complex Medical Data Symposium**

2013 ★ **Session Moderator – Health 2.0 Health:Refactored Conference**

2012 ★ **USC Viterbi School of Engineering Dean's Doctoral Fellowship**
Fellowship given to promising incoming Ph.D. students in engineering at USC

2012 ★ **Organizing Committee – Meaningful Use of Complex Medical Data Symposium**

2011 ★ **Organizing Committee – Meaningful Use of Complex Medical Data Symposium**

2011– ★ **Organizer – LA Machine Learning Meetup**

2011– ★ **Co-founder, *Podimetrics***
★ **Winner of Life Sciences Track – MIT100K Business Plan Competition**
★ **Second Runner-Up – MIT100K Elevator Pitch Competition**
★ **Winner – MIT Hacking Medicine Pitching Competition**
Co-founded medical device company *Podimetrics* building a remote monitoring platform for early detection and prevention of diabetic foot ulcers; raised A round funding in early October of 2012. Competed in 2011-2012 MIT $100 Entrepreneurship Competition (Finalist and Winner of Life Sciences Track in the Business Plan Competition; Second Runner-Up and Audience Choice Award in Elevator Pitch Competition. One of six winning teams in inaugural MIT Hacking Medicine Pitch Competition in 2011.

2010– ★ **Project Management Committee – Object Oriented Data Technology Project, Apache Software**

2003 ★ **Inducted into Phi Beta Kappa Academic Honor Society**
Stanford University, Beta Chapter of California; inducted as junior (top 5% of class).

U.S. Patents

2012-2013 ★ **Patents pending – Podimetrics**
Podimetrics has filed several provisional patents related to detection and prediction of diabetic foot ulcers.

2011 ★ **Patent #7,933,900 (granted April 26, 2011; filed October 23, 2005) – Google, Inc.**
Abstract: A user can refine a search over structured data by specifying that a label or an attribute value be used to further filter the results of a query.
Inventors: Bindu Reddy, Jonathan Brunsman, Ning Mosberger, Gaurav Bhaya, Sarah Sirajuddin, David Kale, Jennifer Kozenski, Arvind Sundararajan, and Puneet Agarwal.