

Empirical Methods for CS: Reading List

- [1] Peter J. Denning. Is computer science science? *Communications of the ACM*, 48(4), 2005.
- [2] Craig Loehle. A guide to increased creativity in research—inspiration or perspiration? *Bioscience*, 40(2), 1990.
- [3] Walter F. Tichy. Should computer scientists experiment more? *Computer*, 31(5), 2002.
- [4] Duncan J. Watts and Steven H. Strogatz. Collective dynamics of ‘small-world’ networks. *Nature*, 393:440–442, June 1998.
- [5] Albert-Laszlo Barabasi and Reka Albert. Emergence of scaling in random networks. *Science*, 286:509–512, October 1999.
- [6] George H. John. Robust decision trees: Removing outliers from databases. In *Proc. KDD*, 1995.
- [7] Tim Oates and David Jensen. The effects of training set size on decision tree complexity. In *Proc. ICML*, 1997.
- [8] Tim Oates and David Jensen. Toward a theoretical understanding of why and when decision tree pruning algorithms fail. In *Proc. AAAI*, 1999.
- [9] Sally Floyd and Van Jacobson. The synchronization of periodic routing messages. *IEEE/ACM Transactions on Networking*, 2, 1994.
- [10] John P. A. Ioannidis. Why most published research findings are false. *PLoS Medicine*, 2(8), 2005.
- [11] Amos Tversky and Daniel Kahneman. Judgment under uncertainty: heuristics and biases. *Science*, 185(4157):1124–1131, 1974.
- [12] Lars Backstrom, Dan Huttenlocker, Jon Kleinberg, and Xiangyang Lan. Group formation in large social networks: Membership, growth, and evolution. In *Proc. KDD*, 2006.
- [13] Pedro Domingos and Michael Pazzani. On the optimality of the simple Bayesian classifier under zero-one loss. *Machine Learning*, 29:103–130, 1997.
- [14] Susan Dumais, Edward Cutrell, JJ Cadiz, Gavin Jancke, Raman Sarin, and Daniel C. Robbins. Stuff i’ve seen: a system for personal information retrieval and re-use. In *Proc. SIGIR*, 2003.
- [15] Susan Dumais, Edward Cutrell, and Hao Chen. Optimizing search by showing results in context. In *Proc. SIGCHI*, 2001.
- [16] Matthew Hertz and Emery D. Berger. Quantifying the performance of garbage collection vs explicit memory management. In *Proc. OOPSLA*, 2005.
- [17] David D. Jensen and Paul R. Cohen. Multiple comparisons in induction algorithms. *Machine Learning*, 38:309–338, 2000.
- [18] Stephen M. Blackburn et al. Wake up and smell the coffee: evaluation methodology for the 21st century. *Communications of the ACM*, 51(8), 2008.
- [19] Marc Liberatore, Brian Neil Levine, and Chadi Barakat. Maximizing transfer opportunities in bluetooth dtns. In *Proc. CoNEXT*, 2006.
- [20] Tony Hey, Stewart Tansley, and Kristin Tolle, editors. *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Microsoft Research, 2009.
- [21] Scott Kirkpatrick and Bart Selman. Critical behavior in the satisfiability of random boolean expressions. *Science*, 264:1297–1301, 1994.

- [22] Peter Medawar. Is the scientific paper a fraud? BBC Third Programme; Experiment: A Series of Scientific Case Histories, 1964.
- [23] Richard W. Hamming. You and your research. Presented at the Bell Communications Research Colloquium Seminar, 1986.