

JOSHUA V. DILLON

251 Tenth Street NW, Apartment F216
Atlanta, GA 30318

jvdillon@gatech.edu
(765) 532-4643

OBJECTIVE To gain knowledge and experience in the field of statistical machine learning. It is my hope that graduate and professional endeavors will form a strong foundation and lead to a permanent research position.

EDUCATION *Ph.D., Computer Science* Jan. 2009 – Present
Georgia Institute of Technology Atlanta, GA

M.S., Electrical & Computer Engineering Aug. 2005 – Dec. 2008
Purdue University West Lafayette, IN

B.S., Electrical & Computer Engineering Aug. 2001 – Apr. 2005
Michigan Technological University Houghton, MI
Summa Cum Laude; Double Major

TEACHING EXPERIENCE *Lecturer* Jan. 2008 – Dec. 2008
Purdue University West Lafayette, IN

Two semester instructor for “Software Engineering Tools Laboratory,” an ECE undergraduate course which acquaints students with software engineering tools and scripting languages (e.g., korn shell, python). Responsible for designing curriculum and delivering weekly lectures to 60+ students. Additionally responsible for coordinating lab TAs and related material, including laboratory exercises, practical exams, and a comprehensive course project. Department-wide, one of two graduate student lecturers.

Teaching Assistant Jan. 2007 – May 2007
Purdue University West Lafayette, IN

Coordinating TA for “Microprocessor System Design and Interfacing,” an undergraduate course which introduces computer organization, assembly programming, digital/analog interfaces, and microprocessor system design. Responsible for one-on-one and laboratory guidance to help students bridge the gap between theory and application. Additionally responsible for five undergraduate TAs.

WORK EXPERIENCE *Research Intern* May 2009 – Aug. 2009
Microsoft Research Redmond, WA

This work examined robust automatic query expansion for information retrieval using convex optimization techniques. Automatic query expansion seeks to mitigate the effect of vocabulary mismatch between queries and documents, typically through unsupervised methods. Although state-of-the-art techniques perform extremely well on-average, they are practically limited by high variance across queries. We depict this shortcoming as a risk/reward tradeoff between expansion and no expansion and encode it as a convex objective which balances notions of expansion reward and risk due to term uncertainty. This work focuses on incorporating self-tuning expansion methods into a risk-aware framework and justifies the result by examining the space of such models parameterized at random.

ISCR Summer Scholar May 2006 – Aug. 2006
Lawrence Livermore National Laboratory Walnut Creek, CA

Used data mining and factor analysis techniques to analyze the Joint Genome Institute’s DNA sequencing process. By correlating resources, machines, and operators with the quality of sequenced DNA, specific negative modi operandi were isolated among the network of actualized and unactualized production paths. Both parametric and non-parametric techniques were used to isolate anomalous DNA traces. Then, through the use of statistical process control, component analysis, and data clustering, the contributing production sub-

networks were identified. Implementation made extensive use of Matlab and Perl, as well as C and SQL. Results and findings were presented to department heads and the division director in a final presentation and scholar symposium. Work funded by the DOE LLNL Institute for Scientific Computing.

Extreme Blue Internship

May 2005 – Aug. 2005

International Business Machines, Inc.

Austin, TX

Worked with a research-oriented technical and business team to implement and internally market an intelligent error-tracing tool for the Linux kernel. Designed, implemented novel techniques to identify statistically abnormal execution paths. Feature engineering included, but not limited to, kernel binary disassembly, resolving dynamic control flow, system profiling with call-stack resolution, and advanced data structure architecture to facilitate efficient access. Designed several techniques of minimizing feature complexity through graph theoretic techniques, e.g., partitioning, coloring, etc. Implemented artificial neural networks and logistic regression for classification. Researched and experimented with innovative classification techniques based on harmonically relating features. Implemented as an object-oriented modular framework using C and Perl. Work was funded jointly by the Extreme Blue program and the Linux Technology Center.

Software Engineering Internship

Feb. 2005 – May 2005

ThermoAnalytics, Inc.

Calumet, MI

Designed and implemented a Trolltech QScript (language subset) to C code translator. Efforts included lexical analysis, context free grammar specification, abstract syntax tree representation and corresponding auto-typing symbol table, and semantic routines. The majority of this project entailed designing and implementing the semantic routines as an easily maintainable and highly extensible context sensitive analyzer. Implementation utilized Flex, Bison, and object-oriented C++.

Electrical Engineering Internship

May 2004 – Aug. 2004

International Business Machines, Inc.

Rochester, MN

Designed VHDL logic for the floating-point core of the Cell processor. Designed and implemented Java-based floating-point analysis and prototyping software in to reduce design time. Significantly improved turnaround time of logic macro timing and synthesis by implementing several Perl report generating tools. The high profile nature of this project made enhanced security training and clearance necessary. Work funded by the Engineering & Technology Services group under contracts from Sony and Microsoft.

Software Engineering Internship

May 2003 – Aug. 2003

International Business Machines, Inc.

Rochester, MN

Designed, implemented, and packaged an SAP R/3 cluster management plug-in for iSeries Navigator. Design goals included extensibility, graphical ease-of-use, and an aggressive release cycle to meet clients' demands. Work was primarily done using Java. Efforts involved the coordination of domestic and German colleagues. End product was delivered in a fully packaged form, ahead of schedule.¹

PUBLICATIONS

1. J. Dillon and G. Lebanon.
Statistical and Computational Tradeoffs in Stochastic Composite Likelihood *Proc. of the 12th International Conference on Artificial Intelligence and Statistics*. 2009 (to appear).
2. G. Lebanon, Y. Mao, and J. Dillon.
The Locally Weighted Bag of Words Framework for Document Representation. *Journal of Machine Learning Research*, 8(Oct):2405-2441, 2007.
3. Y. Mao, J. Dillon, G. Lebanon.
Sequential Document Visualization. *IEEE Transactions on Visualization and*

¹For more details, see <http://www.alphaworks.ibm.com/tech/sapsysman>.

Computer Graphics, 13(6), 2007.

4. J. Dillon, Y. Mao, G. Lebanon and J. Zhang.
Statistical Translation, Heat Kernels, and Expected Distances.
Proc. of the 23rd Conference on Uncertainty in Artificial Intelligence, 2007.
5. J. Dillon, Y. Mao, G. Lebanon and J. Zhang.
Statistical Translation, Heat Kernels, and Expected Distances.
NIPS Workshop on Learning to Compare Examples, 2006.
6. J. Dillon, J. Baumohl, D. Casner, A. Copeland and D. Pletcher.
Isolation of Production Modi Operandi Correlated with Anomalous Behavior.
Lawrence Livermore National Laboratory 2006 Student Symposium, 2006.
7. D. Casner, J. Dillon and D. Pletcher.
Closed Loop Optimization of Gene Sequencing Run Times.
Lawrence Livermore National Laboratory 2006 Student Poster Symposium, 2006.

SELECTED
AWARDS
& HONORS

- US delegate, 57th Lindau Meeting of Nobel Laureates and Students, Germany, 2007. Sponsored by US DOE of Science and the NSF Directorate for Mathematical and Physical Sciences.
- Ross Graduate Fellowship, Purdue University, 2005-06.
- Summa Cum Laude, Dept. of Electrical and Computer Engineering, Michigan Tech, 2005.
- Eta Kappa Nu, Beta Gamma Chapter, Electrical and Computer Engineering Honor Society, 2005-present.
- Phi Kappa Phi, Honor Society, 2004-present.
- Tau Beta Pi, Engineering Honor Society, 2004-present.
- Award of Excellence, Dept. of Mathematics, Michigan Technological University, 2002.
- Four-year tuition scholarship, Board of Control, Michigan Technological University, 2001-05.
- Valedictorian, Cass City High School, 2001.