RICH CARUANA

172 Pleasant Grove Road
Ithaca, NY 14850

caruana@cs.cornell.edu
cell: 607-351-2546

EDUCATION

1997 CARNEGIE MELLON UNIVERSITY, PITTSBURGH, PENNSYLVANIA
Ph.D., Computer Science. Thesis: *Multitask Learning*

1984 VILLANOVA UNIVERSITY, VILLANOVA, PENNSYLVANIA
M.Sc., Computer Science

1982 VILLANOVA UNIVERSITY, VILLANOVA, PENNSYLVANIA
B.Sc., Mathematics. Minors in Physics and Chemistry

RESEARCH POSITIONS

July 2001–Present
CORNELL UNIVERSITY, ITHACA, NEW YORK
Assistant Professor, Department of Computer Science

July 2000–July 2001
CARNEGIE MELLON UNIVERSITY, PITTSBURGH, PENNSYLVANIA
Research Faculty, Center for Automated Learning and Discovery

September 1998–July 2000
UNIVERSITY OF CALIFORNIA AT LOS ANGELES, LOS ANGELES, CALIFORNIA
Assistant Professor, Departments of Radiology and Computer Science
CARNEGIE MELLON UNIVERSITY, PITTSBURGH, PENNSYLVANIA
Visiting Professor, Center for Automated Learning and Discovery (CALD)

July 1996–March 2000 (part-time starting September 1998)
JUSTSYSTEM PITTSBURGH RESEARCH CENTER (JPRC), PITTSBURGH, PENNSYLVANIA
Research Scientist

October 1986–August 1989
PHILIPS LABS, BRIARCLIFF MANOR, NEW YORK
Research Scientist, Department of Machine Learning

September 1984–October 1986
GTE GOVERNMENT SYSTEMS, MOUNTAIN VIEW, CALIFORNIA
Knowledge Engineer, Adaptive and Expert Systems Group

August 1980–August 1983
GEOMETRIC DATA COMPANY, VALLEY FORGE, PENNSYLVANIA
Research Scientist, Biomedical Image Recognition Research Group

PROFESSIONAL ACTIVITIES

- Program Co-Chair (with X. Wu) for the 2007 ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2007.
- Funding Co-Chair (with L. Getoor) for the International Machine Learning Society (IMLS) and International Conferences on Machine Learning (ICML), 2008.

- Funding Chair for the International Machine Learning Society (IMLS) and International Conferences on Machine Learning (ICML), 2005–2007.
- Area Chair for the International Conference on Machine Learning (ICML), 2003, 2005, 2006.
- Senior Program Committee for the National Conference on Artificial Intelligence (AAAI), 2006.
- Action Editor for the *Journal of Machine Learning Research* (JMLR), 2006–present.
- Action Editor for the *Machine Learning Journal* (MLJ), 2006–present.
- Co-Chair (with K. Bennet and D. Silver) of Workshop on "Inductive Transfer: Ten Years Later," Neural and Information Processing Systems (NIPS), 2005. An MLJ Special Issue and a book of the workshop proceedings are in preparation.
- Co-Chair (with T. Joachims) of KDD Cup for the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2004.
- Co-Chair (with T. Joachims) of Workshop on "Predicting Probabilities," Neural and Information Processing Systems (NIPS), 2004.
- Area Chair for Neural and Information Processing Systems (NIPS), 2003.
- Senior Program Committee for International Conference on Data Mining (ICDM), 2003.
- Senior Program Committee for ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2003.
- Senior Program Committee for American Medical Informatics Association Conference (AMIA), 2003.
- Co-Chair (with T. Joachims) of Workshop on "Beyond Classification and Regression: Learning Rankings, Preferences, Equality Predicates, and Other Structures," Neural and Information Processing Systems (NIPS), 2002.
- Senior Program Committee for the International Conference on Artificial Neural Networks in Medicine and Biology (ANNIMAB), 2000.
- General Co-Chair (with V. de Sa) of Workshops, Neural and Information Processing Systems (NIPS), 2000.
- General Co-Chair (with S. Becker) of Workshops, Neural and Information Processing Systems (NIPS), 1999.
- Co-Chair (with V. de Sa) of Workshop on "Integrating Supervised and Unsupervised Learning," Neural and Information Processing Systems (NIPS), 1998.
- Co-Chair (with J. Orr) of Workshop on "Tricks of the Trade," Neural and Information Processing Systems (NIPS), 1996.
- Co-Chair (with J. Baxter, T. Mitchell, L. Pratt, D. Silver, S. Thrun) of Workshop on "Learning to Learn," Neural and Information Processing Systems (NIPS), 1995.

HONORS AND AWARDS

- Best Student Paper Award at the European Conference on Machine Learning (ECML), 2007 for D. Sorokina, R. Caruana, and M. Riedewald, "Additive Groves of Regression Trees."
- Solomonov Lecture, Josef Stefan Institute (JSI), Slovenia, May 2006.
- Merrill Faculty Award, Cornell University 2005.
- Best Student Paper Award at the International Conference on Machine Learning (ICML), 2005 for A. Niculescu-Mizil and R. Caruana, "Predicting Good Probabilities with Supervised Learning."
- NSF CAREER Award, "Meta Clustering," May 2004-2009.
- Sonny Yau Teaching Award, Cornell University, 2003-2004.
- Nomination for the Best Paper Award at the American Medical Informatics Conference (AMIA), 2003 for R. Caruana, S. Niculescu, B. Rao, and C. Simms, "Evaluating the C-section Rate of Different Physician Practices: Using Machine Learning to Model Standard Practice."
- Award for Outstanding Research Contribution, Philips Labs, 1988.

TALKS

- Invited Speaker, NIPS 2007 Workshop on Overcoming Computational Bottlenecks in Machine Learning, December 2007.
- Invited Speaker, AAAI Workshop on Evaluation Methods for Machine Learning, July 2007.
- Invited Speaker, Department of Biomedical Informatics, University of Pittsburgh, February 2007.

- Invited Speaker, University of Ottawa, March 2007.
- Invited Speaker, University of Texas at Austin, February 2007.
- Invited Speaker, Brigham Young University, February 2007.
- Invited Speaker, NIPS 2006 Workshop on the Testing and Evaluation of Deployable Learning Systems, December 2006.
- Invited Speaker (three invited talks), Josef Stefan Institute (JSI), Slovenia, May 2006.
- Invited Speaker, Wharton School Department of Statistics, University of Pennsylvania, March 2006.
- Invited Tutorial, American Meteorology Conference (AMS), January 2006.
- Invited Speaker, Toyota Technical Institute at Chicago (TTI-C), November 2005.
- Invited Speaker, University of Illinois at Urbana-Champaign, November 2005.
- Invited Speaker, Microsoft Research, Redmond WA, October 2005.
- Invited Instructor, Toyota Technical Institute at Chicago (TTI-C) Machine Learning Summer School, May 2005.
- Invited Speaker, Rensselaer Polytechnic Institute (RPI), April 2005.
- Invited Speaker, University of California at San Diego, CA, January 2005.
- Invited Speaker, Yahoo! Research Labs, Pasadena, CA, January 2005.
- Invited Speaker, Fair Isaac Research, San Diego, CA, January 2005.
- Invited Speaker, American Meteorology Conference (AMS), January 2005.
- Invited Speaker, "Methods for Learning From Imbalanced Data." American Association for Artificial Intelligence (AAAI) Workshop on Learning From Imbalanced Data, July 2000.
- Invited Speaker (three invited talks), "Unabridged and Multitask Learning," "Machine Learning, Medical Decision Making, and Explanation," "Semi-Supervised Clustering." University of Edinburgh Research Jamboree, Edinburgh, Scotland, May 2000.
- Invited Lecturer, "Multitask Learning." University of Skövde, Skövde, Sweden, December 1994.

JOURNAL PUBLICATIONS

A. Niculescu-Mizil and R. Caruana, "Predicting Good Probabilities With Supervised Learning." *Machine Learning Journal* (MLJ) invited paper, accepted subject to minor revisions, 2008.

E. Ipek, S. McKee, K. Singh, R. Caruana, B. Supinski, and M. Schulz, "Efficient Architectural Design Space Exploration via Predictive Modeling." *ACM Transactions on Architecture and Code Optimization* (TACO), Vol. 4, No. 4 January 2008, pp.1-33.

K. Singh, E. Ipek, S.A. McKee, B.R. de Supinski, M. Schulz, R. Caruana, "Predicting Parallel Application Performance via Machine Learning." Concurrency and Computation: Practice and Experience (CCPE), Vol. 19, No. 13, 10 December 2007, pp. 2219-2235.

W.M. Hochachka, R. Caruana, D. Fink, S. Kelling, A. Munson, M. Riedewald, and D. Sorokina, "Data Mining for Discovery of Pattern and Process in Ecological Systems." *Journal of Wildlife Management,* Vol. 71, No. 7, pp. 2427–2437, 2007.

G.F. Cooper, V. Abraham, C.F. Aliferis, J.M. Aronis, B.G. Buchanan, R. Caruana, M.J. Fine, J.E. Janosky, G. Livingston, T. Mitchell, S. Montik, and P. Spirtes, "Predicting Dire Outcomes of Patients with Community Acquired Pneumonia." *Journal of Biomedical Informatics*, Vol. 38, No. 5, 2005, pp. 347-366.

R. Caruana and V.R. de Sa, "Benefitting from the Variables that Variable Selection Discards." *Journal of Machine Learning Research*, Vol. 3, 2003, pp.1245-1264.

A. Goldenberg, G. Shmueli, R. Caruana, and S. Fienberg, "Early Statistical Detection of Anthrax Outbreaks by Tracking Over-the-Counter Medication Sales." *Proceedings of the National Academy of Sciences*, 99, 2002, pp. 5237-5240.

C.J. Simms, L. Meyn, R. Caruana, R.B. Rao, T. Mitchell, and M. Krohn, "Predicting Cesarean Delivery with Decision Tree Models." *The American Journal of Obstetrics and Gynecology*, Vol. 183, No. 5, 2000, pp. 1198-1206.

G.F. Cooper, C.F. Aliferis, R. Ambrosino, J. Aronis, B.G. Buchanan, R. Caruana, M. J. Fine, C. Glymour, G. Gordon, B.H. Hanusa, J.E. Janosky, C. Meek, T. Mitchell, T. Richardson, and P. Spirtes, "An Evaluation of Machine Learning Methods for Predicting Pneumonia Mortality." *Artificial Intelligence in Medicine*, Vol. 9, 1997, pp. 107-138.

R. Caruana, "Multitask Learning." *Machine Learning*, Vol. 28, 1997, pp. 41-75.

T. Mitchell, R. Caruana, D. Freitag, J. McDermott, and D. Zabowski, "Experience with a Learning Personal Assistant." *Communications of the ACM*, Vol. 37, No. 7, 1994, pp. 80-91.

R. Caruana, R.B. Searle, and S.I. Shupack, "Additional Capabilities for a Fast Algorithm for the Resolution of Spectra." *Journal of Analytical Chemistry*, 1988.

R. Caruana, R.B. Searle, T. Heller, and S.I. Shupack, "Fast Algorithm for the Resolution of Spectra." *Journal of Analytical Chemistry*, 1986.

REFEREED CONFERENCES

N. Nguyen and R. Caruana, "Improving Classification with Pairwise Constraints: A Margin-based Approach." To appear in *The Proceedings of The European Conference on Machine Learning and Practice of Knowledge Discovery in Databases* (ECML PKDD), September 2008.

N. Nguyen and R. Caruana, "Classification with Partial Labels." To appear in *The Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (KDD), August 2008.

R. Caruana, A. N. Karaampatziakis, and Y. Yessenalina, "An Empirical Evaluation of Supervised Learning in High Dimensions." To appear in *The Proceedings of the International Conference on Machine Learning* (ICML), July 2008.

D. Sorokina, R. Caruana, M. Riedewald, and D. Fink, "Detecting Statistical Interactions with Additive Groves of Trees." To appear in *The Proceedings of the International Conference on Machine Learning* (ICML), July 2008.

E. Ipek, O. Mutlu, J. Martinez, and R. Caruana, "Self-Optimizing Memory Controllers: A Reinforcement Learning Approach." To appear in *The Proceedings of The 35th International Symposium on Computer Architecture* (ISCA), June 2008.

N. Nguyen and R. Caruana, "Consensus Clusterings." The *Proceedings of the Seventh International Conference on Data Mining* (ICDM), October 2007.

D. Sorokina, R. Caruana, and M. Riedewald, "Additive Groves of Regression Trees." The *Proceedings of the European Conference on Machine Learning* (ECML), September 2007 (Best Student Paper Award).

D. Skalak, A. Niculescu-Mizil, and R. Caruana, "Classifier Loss Under Metric Uncertainty." The *Proceedings of the European Conference on Machine Learning* (ECML), September 2007.

A. Niculescu-Mizil and R. Caruana, "Inductive Transfer for Bayesian Network Structure Learning." The *Proceedings of the International Conference on Artificial Intelligence and Statistics* (AISTATS), March 2007.

R. Caruana, A. Munson, and A. Niculescu-Mizil, "Getting the Most Out of Ensemble Selection." The *Proceedings of the Sixth International Conference on Data Mining* (ICDM), December 2006.

R. Caruana, M. Elhawary, N. Nguyen, and C. Smith, "Meta Clustering." The *Proceedings of the Sixth International Conference on Data Mining* (ICDM), December 2006.

E. Ipek, S. McKee, B. de Supinski, M. Schulz, and R. Caruana, "Efficiently Exploring Architectural Design Spaces via Predictive Modeling." The *Proceedings of the 12th International Conference on Architectural Support for Programming Languages and Operating Systems* (ASPLOS), October 2006.

C. Bucila, R. Caruana, and A. Niculescu-Mizil, "Model Compression." *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (KDD), August 2006.

R. Caruana, M. Elhawary, D. Fink, W. Hochachka, S. Kelling, A. Munson, M. Riedewald, and D. Sorokina, "Mining Citizen Science Data to Predict Prevalence of Wild Bird Species." *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (KDD), August 2006, pp. 909-915.

L. Backstrom and R. Caruana, "C2FS: An Algorithm for Feature Selection in Cascade Neural Nets." *Proceedings of the IEEE World Congress on Computational Intelligence and the International Joint Conference on Neural Nets* (IJCNN), July 2006, pp. 400-408.

R. Caruana and A. Niculescu-Mizil, "An Empirical Comparison of Supervised Learning Algorithms." *Proceedings of the 23rd International Conference on Machine Learning* (ICML), June 2006, pp. 161-168.

A. Niculescu-Mizil and R. Caruana, "Predicting Good Probabilities with Supervised Learning." International Conference on Machine Learning (ICML), 2005. (Distinguished Student Paper award.)

A. Niculescu-Mizil and R. Caruana, "Obtaining Calibrated Probabilities from Boosting." International Conference on Uncertainty in Artificial Intelligence, 2005, pp. 413-420.

R. Caruana, A. Niculescu-Mizil, G. Crew, and A. Ksikes, "Ensemble Selection from Libraries of Models." International Conference on Machine Learning (ICML), 2004.

R. Caruana and A. Niculescu-Mizil, "Data Mining in Metric Space." International Conference on Knowledge Discovery and Data Mining (KDD), 2004.

R. Caruana, S. Niculescu, R.B. Rao, C.J. and Simms, "Evaluating the C-Section Rate of Different Physician Practices: Using Machine Learning to Model Standard Practice." The American Medical Informatics Conference (AMIA), November 2003 (nominated for best paper award).

J. Langford and R. Caruana, "(Not)Bounding the True Error." *Neural and Information Processing Systems*, Vol. 14 (Proceedings of NIPS*2001). MIT Press, 2002.

R. Caruana, S. Niculescu, R.B. Rao, and C.J. Simms, "Machine Learning for Sub-Population Assessment: Evaluating the C-Section Rate of Different Physician Practices." The American Medical Informatics Conference (AMIA), November 2002, pp. 126-130.

R. Caruana, "A Non-Parametric EM-Style Algorithm for Imputing Missing Values." *Proceedings of the International Conference on Artificial Intelligence and Statistics* (AISTATS), January 2001.

R. Caruana, S. Lawrence, and L. Giles, "Overfitting in Artificial Neural Nets Trained with Backpropagation, Conjugate Gradient, and Early Stopping." *Neural and Information Processing Systems*, Vol. 13 (Proceedings of NIPS*2000). MIT Press, 2001.

A. Berger, R. Caruana, D. Cohn, D. Freitag, and V. Mittal, "Bridging the Lexical Chasm: Automatic FAQ Answer Finding." Special Interest Group on Information Retrieval (SIGIR),-July 2000.

J. O'Sullivan, J. Langford, R. Caruana, A. Blum, "Unabridged Learning." International Conference on Machine Learning (ICML), July 2000.

C.J. Simms, R. Caruana, M.J. Krohn, L. Meyn, T. Mitchell, R.B. Rao, and I. Schmeuking, "Predicting Caesarean Section with Decision Trees." Annual Meeting of the Society of Fetal and Maternal Medicine, February 2000.

R. Caruana, "Case-Based Explanation of Artificial Neural Nets." Artificial Neural Nets in Medicine and Biology (ANNIMAB), May 2000.

R. Caruana, H. Kangarloo, J.D.N. Dionisio, U. Sinha, D.B. Johnson, "Case-Based Explanation of Non–Case-Based Learning Methods." Proceedings of the 1999 American Medical Informatics Association (AMIA) Symposium, 1999, pp. 212-215.

R. Caruana, J.D.N. Dionisio, D.B. Johnson, R.K. Taira, H. Kangarloo, "Automatic Imaging Protocol Selection." American Radiology Conference (IRAS), 1999.

R. Caruana, "Applying Case-Based Explanation to Non–Case-Based Methods such as Artificial Neural Nets or Decision Trees." American Radiology Conference (IRAS), 1999.

R. Caruana and J. O'Sullivan, "Multitask Pattern Recognition for Autonomous Robots." IEEE International Conference on Intelligent Robotic Systems (IROS), October 1998.

R. Caruana and V.R. de Sa, "Using Feature Selection to Find Inputs that Work Better as Extra Outputs." The International Conference on Neural Nets (ICANN), September 1998.

R. Caruana and J. O'Sullivan, "Multitask Pattern Recognition for Vision-Based Autonomous Robots." The International Conference on Neural Nets (ICANN), September 1998.

R. Caruana and V.R. de Sa, "Promoting Poor Features to Supervisor: Some Inputs Works Better as Outputs." *Neural and Information Processing Systems*, Vol. 9 (Proceedings of NIPS*96). MIT Press, 1997, pp. 389-395.

R. Caruana, "Algorithms and Applications for Multitask Learning." *Machine Learning, Proceedings of the 13th International Conference on Machine Learning* (ICML). Morgan Kauffmann, 1996, pp. 87-95.

R. Caruana, S. Baluja and T. Mitchell, "Using the Future to 'Sort Out' the Present: Rankprop and Multitask Learning for Medical Risk Evaluation." *Advances in Neural Information Processing Systems,* Vol. 8 (Proceedings of NIPS*95). MIT Press, 1996, pp. 959-965.

S. Baluja and R. Caruana, "Removing the Genetics from the Standard Genetic Algorithm." *Proceedings of the 12th Annual Conference on Machine Learning*, 1995, pp. 38-46.

R. Caruana, "Learning Many Related Tasks at the Same Time with Backpropagation." *Advances in Neural Information Processing Systems 7* (Proceedings of NIPS*94). MIT Press, 1995 pp. 657-664.

R. Caruana and D. Freitag, "Greedy Attribute Selection." *Machine Learning, Proceedings of the Eleventh International Conference on Machine Learning* (ICML 1994). Morgan Kauffmann, 1994, pp. 28-36.

R. Caruana, "Multitask Connectionist Learning." *Proceedings of the 1993 Connectionist Models Summer School*, 1993, pp. 372-379.

R. Caruana, "Multitask Learning: A Knowledge-Based Source of Inductive Bias." *Proceedings of the 10th International Conference on Machine Learning*, 1993, pp. 41-48.

R. Caruana and J.D. Schaffer, "Using Multiple Representations to Control Inductive Bias: Gray and Binary Codes for Genetic Algorithms." *Proceedings of the Sixth International Workshop on Machine Learning* (ML 1989). Morgan Kaufmann, 1989, pp. 375-378.

J.D. Schaffer, R. Caruana, and L.J. Eshelman, "Designing Neural Nets that Generalize Optimally with Genetic Algorithms." Los Alamos Conference on Emergent Computation, May 1989.

L.J. Eshelman, R. Caruana, and J.D. Schaffer, "Biases in the Crossover Landscape." International Conference on Genetic Algorithms, June 1989.

J.D. Schaffer, R. Caruana, L.J. Eshelman, and R. Das, "A Study of Control Parameters Affecting Online Performance of Genetic Algorithms for Function Optimization." International Conference on Genetic Algorithms, June 1989.

R. Caruana, L.J. Eshelman, and J.D. Schaffer, "Representation and Hidden Bias II: Eliminating Defining Length Bias in Genetic Search with Shuffle Crossover." International Joint Conference on Artificial Intelligence (IJCAI), August 1989.

R. Caruana, J.D. Schaffer, "Representation and Hidden Bias: Gray vs. Binary Coding for Genetic Algorithms." International Conference on Machine Learning, June 1988.

BOOK CHAPTERS AND MAGAZINE ARTICLES

E. Ipek, S. McKee, B. de Supinski, M. Schulz, and R. Caruana, Invited article to appear in IEEE Micro discussing our 2006 ASPLOS paper "Efficiently Exploring Architectural Design Spaces via Predictive Modeling."

R. Caruana and T. Joachims, "KDD Cup 2004: Results and Analysis." SIGKDD Explorations 6(2):95-108, 2004.

R. Caruana, "15 Useful Tricks with Extra Outputs." *Neural Networks: Tricks of the Trade*. G.B. Orr and K.-R. Muller (Eds.), Springer-Verlag, 1998.

R. Caruana, "Multitask Learning." *Learning to Learn*. S. Thrun and L. Pratt (Eds.), Kluwer Academic Publishers, 1997.

R. Caruana, D. Freitag, "How Useful Is Relevance?" *Intelligent Relevance: Papers from the 1994 Fall Symposium*. American Association for Artificial Intelligence (AAAI) Report FS-94-02, ISBN 0-929280-76-8, 1994.

R. Caruana, "The Automatic Training of Rule Bases that Use Numerical Uncertainty Representations." *Uncertainty in Artificial Intelligence*, Vol. 3. North-Holland, 1988.

WORKSHOPS, SYMPOSIA, AND NON-REFEREED CONFERENCES

M. Hochachka, D. Fink, D.N. Bonter, R. Caruana, S.T. Kelling, A. Munson, M. Riedewald, and D. Sorokina, "Exploring the Ecological Consistency of Bird Conservation Regions Across a Gradient of Human Density." North American Ornithological Congress, Veracruz, Mexico. October 2006.

D. Fink, W.M. Hochachka, R. Caruana, S.T. Kelling, A. Munson, M. Riedewald, and D. Sorokina. "Data Mining to Explore Spatial and Temporal Variation in Bird Distribution: Irruptive Winter Migrants." North American Ornithological Congress, Veracruz, Mexico. October 2006.

M. Riedewald, R. Caruana, D. Fink, W.M. Hochachka, A. Munson, S.T. Kelling, D. Sorokina, "Advances in Data Mining to Support Bird Conservation." Avian Knowledge Network Workshop, Cornell Lab of Ornithology, June 2006.

A. Munson, R. Caruana, D. Fink, W.M. Hochachka, M. Riedewald, S.T. Kelling, D. Sorokina, "Decision Trees and Sweep Analysis of Project Feeder Watch Data." Avian Knowledge Network Workshop, Cornell Lab of Ornithology, June 2006.

D. Sorokina, R. Caruana, D. Fink, W.M. Hochachka, A. Munson, M. Riedewald, S.T. Kelling, "Variable Interaction Detection." Avian Knowledge Network Workshop, Cornell Lab of Ornithology, June 2006.

D. Fink, R. Caruana, W.M. Hochachka, A. Munson, S.T. Kelling, M. Riedewald, D. Sorokina, "Statistics and Data Mining." Avian Knowledge Network Workshop, Cornell Lab of Ornithology, June 2006.

W.M. Hochachka, R. Caruana, D. Fink, A. Munson, S.T. Kelling, M. Riedewald, D. Sorokina, "Analysis at Scale: New Strategies for Analysis of Biodiversity Data." Avian Knowledge Network Workshop, Cornell Lab of Ornithology, June 2006.

A. Niculescu-Mizil and R. Caruana, "Learning the Structure of Related Tasks." To appear in the *Proceedings of Workshop on Inductive Transfer: 10 Years Later*, held at the Neural Information Processing Conference (NIPS), December 2005.

R. Caruana and A. Niculescu-Mizil, "Predicting Good Probabilities with Supervised Learning." Oral presentation at the Machines that Learn Workshop, Snowbird, Utah, April 2005.

B. Shaparenko, R. Caruana, J. Gehrke, and T. Joachims, "Identifying Temporal Patterns and Key Players in Document Collections." In *Proceedings of the IEEE ICDM Workshop on Temporal Data Mining: Algorithms, Theory and Applications (TDM-05)*, pp. 165–174, Houston, TX, USA, 2005.

R. Caruana and A. Niculescu-Mizil, "An Empirical Evaluation of Supervised Learning for ROC Area." *The Proceedings of the First Workshop of ROC Analysis in Artificial Intelligence* (ROCAI), August 2004.

R. Caruana, D. Cohn, and A. McCallum, "Semi-Supervised Clustering with User Feedback." Oral presentation at the Machines that Learn Workshop, Snowbird, Utah, April 2000.

R. Caruana and P. Hodor, "A High-Precision Workbench for Extracting Information from the Protein Data Bank (PDB)." Knowledge and Data Discovery (KDD) Workshop on Text and Information Extraction, 2000.

R. Caruana and M. Mullin, "Estimating the Number of Local Minima in Complex Search Spaces." International Joint Conference on Artificial Intelligence Workshop on Optimization, (IJCAI), 1999.

D.B. Johnson, J.D.N. Dionisio, R.K. Taira, R. Caruana, W.W. Chu, and H. Kangarloo, "Tailoring Medical Literature Searches to Support Evidence-Based Medical Practice." Scientific Assembly and Annual Meeting of the Radiological Society of North America (RSNA), infoRAD Exhibit, 1999.

R. Caruana, "15 Useful Tricks with Extra Outputs." Neural and Information Processing Systems (NIPS) Workshop on Tricks of the Trade, 1996.

R. Caruana and D. Freitag, "How Useful Is Relevance?" American Association for Artificial Intelligence (AAAI) Fall Symposium on Relevance, 1994.

R. Caruana, "Generalization vs. Network Size." Neural and Information Processing Systems (NIPS) Workshop on Generalization, 1993.

L. Chrisman, R. Caruana, and W. Carriker, "Intelligent Agent Design Issues: Internal Agent State and Incomplete Perception." American Association for Artificial Intelligence (AAAI) Fall Symposium on Sensory Aspects of Robotic Intelligence, 1991.

R. Caruana, "The Automatic Training of Rule Bases that Use Numerical Uncertainty Representations." American Association for Artificial Intelligence (AAAI) Workshop on Uncertainty in Artificial Intelligence, 1987.

TECHNICAL REPORTS

A. Munson and R. Caruana, "Cluster Ensembles for Network Anomaly Detection." Cornell University Technical Report, TR2006-2047, 2006.

R. Caruana and A. Niculescu-Mizil, "An Empirical Comparison of Supervised Learning Algorithms Using Different Performance Criteria." Cornell University Technical Report, TR2005-1973, 2005.

D. Cohn, R. Caruana, and A. McCallum, "Semi-Supervised Clustering with User Feedback." Cornell University Technical Report, TR2003-1892, 2003.

R. Caruana, P. Artigas, A.Goldenberg, and A. Likhodedov, "Meta Clustering." Cornell University Technical Report, TR2002-1884, 2002.

R. Caruana, "Multitask Learning." Ph.D. Dissertation, School of Computer Science, Carnegie Mellon University, CMU-CS-97-203, 1997.

W. Buntine and R. Caruana, "Introduction to IND and Recursive Partitioning." NASA Ames Research Center, TR# FIA-91-28, 1991.

R. Caruana, "BANDIT: A Fast Algorithm for the Resolution of Spectra." Master's Thesis, Departments of Computer Science and Chemistry, Villanova University, 1988.

R. Caruana, "Estimating the Number of Minima in Complex Search Spaces." Philips Labs, TR-88-159, 1988.

R. Caruana and J.D. Schaffer, "Optimizing Digital Filters with Simulated Annealing and Genetic Algorithms." Philips Labs, TR-88-123, 1988.

R. Caruana and J.D. Schaffer, "An Investigation of Parameter Sets for Genetic Algorithms." Philips Labs, TR-88-045, 1988.

R. Caruana and B.J. Coffey, "Searching for Optimal FIR Multiplierless Digital Filters with Simulated Annealing." Philips Labs, TR-88-031, 1988.

R.N. Pelavin, B.J. Coffey, and R. Caruana, "Research in Diagnosis Using Design Knowledge." Philips Labs, TR-88-001, 1988.

R. Caruana and J.D. Schaffer, "Gray vs. Binary Coding for Genetic Algorithm Function Optimizers." Philips Labs, TR-87-080, 1987.

D.P. Benjamin and R. Caruana, "Partial-Matching as Search." Philips Labs, TR-87-019, 1987.

R. Caruana, "Experiments in Rule-Based Learning in Systems Using Numerical Uncertainty Representations." GTE Western Division Technical Report, 1986.

PATENTS

R. Sukthankar, R. Caruana, K. Hasegawa, and M. Mullin, "Using Active Monitor Illumination for 3-D Active Imaging." United States Patent 6,704,447; filed April 2000, granted March 9, 2004. Assignee: JustSystem Pittsburgh Research Center, Pittsburgh, Pennsylvania.

R. Caruana, "Iterated K-Nearest Neighbor Method and Article of Manufacture for Filling in Missing Values." United States Patent 6,047,287; filed May 5, 1998, granted April 4, 2000. Assignee: JustSystem Pittsburgh Research Center, Pittsburgh, Pennsylvania.

RELEASED SOFTWARE

PERF
Software for computing more than 20 different supervised learning performance measures. **PERF** was released in April 2004 and has been downloaded more than 2000 times.

IND (INDUCE)
Software for growing decision trees, widely used both in published research and for teaching. Wray Buntine is the principal author of this software. With Buntine, improved **IND** and prepared it for public release while working as an intern at NASA. **IND** was released by NASA in 1991 and has been downloaded more than 5,000 times.

REVIEWING

- Action Editor for *Journal of Machine Learning Research* (JMLR)
- Action Editor for *Machine Learning Journal* (MLJ)
- American Association for Artificial Intelligence (AAAI)
- IEEE Transactions on Knowledge and Data Engineering
- IEEE Transactions on Neural Nets
- IEEE Transactions on Systems, Men and Cybernetics
- International Conference on Artificial Neural Nets (ICANN)
- International Conference on Artificial Neural Networks in Medicine and Biology (ANNIMAB)
- International Conference on Data Mining (ICDM)
- International Conference on Machine Learning (ICML)
- International Joint Conference on Artificial Intelligence (IJCAI)
- International Joint Conference on Artificial Neural Nets (IJCNN)
- *Journal of Analytical Chemistry*
- *Journal of Artificial Intelligence Research* (JAIR)
- *Journal of Machine Learning Research* (JMLR)
- International Conference on Knowledge Discovery and Data Mining (KDD)
- *Machine Learning Journal* (MLJ)
- Neural and Information Processing Systems (NIPS)
- Portuguese Conference on Artificial Intelligence

RESEARCH GRANTS, GIFTS, AND AWARDS

NSF Grant, 2007-2010
"Computer Architecture Optimization: A Machine Learning Approach," Co-Principal Investigator with J. Martinez. Amount: 35% of $500,000.

Google Research Gift, 2006-2007
"Model Compression." Principal Investigator. Amount: 100% of $130,000.

NSF Grant, 2006-2009

"Ecological Discovery & Inference: Tools for Data-Driven Exploration and Testing of Observational Data." Co–Principal Investigator with S. Kelling, D. Fink, M. Riedewald, and M. Wells. Amount: 33% of $987,334.

KD-D Grant, 2005
"KD-D Evaluation Proposal." Co–Principal Investigator with J. Gehrke and T. Joachims. Amount: 33% of $51,386.

KD-D Grant, 2005
"KD-D Open Task Proposal." Co–Principal Investigator with J. Gehrke and T. Joachims. Amount: 33% of $100,000.

NSF CAREER Award, 2004-2009
"Meta Clustering." Principal Investigator. Amount: 100% of $507,000.

NSF Grant, 2004-2008
"Tracking Environmental Change through the Data Resources of the Bird-Monitoring Community," Principal Investigator (with S. Kelling, D. Fink, and M. Riedewald). Amount: 20% of $2,691,684.

NSF Grant, 2004-2007
"Optimizing Classification Models to Application-Specific Performance Metrics," Principal Investigator with T. Joachims. Amount: 90% of $270,000.

NSF Grant, 2004-2007
"Discriminative Methods for Learning with Dependent Outputs," Co–Principal Investigator with T. Joachims. Amount 10% of $270,000.

NSF Grant, 2002-2006
"Multi-Algorithm Parallel Optimization of Costly Functions," Co–Principal Investigator with C. Shoemaker. Amount: 10% of $379,999.

AFOSR/AFRL Grant, 2002-2006
"AFRL/Cornell Information Assurance Institute: Intrusion Detection," Senior Personnel (Principal Investigator F. Schneider). Amount: 100% of $165,000.