

Amit Dhurandhar

Present Address

310, Main St.,
Cold Spring, NY 10516

Contact

Homepage: <http://researcher.ibm.com/person/us-adhuran>
Email: adhuran@us.ibm.com
Phone: (352) 235-4789

Research Interests

Machine Learning, Data Mining, Pattern Recognition, Fuzzy Systems

Education

P.h.d. in Computer Science, University of Florida, USA 2006-2009
M.S. in Computer Science, University of Florida, USA 2004-2005
B.E. in Computer Science, Pune University, India 2000-2004

Employment

Research Staff Member, 2010-present
Business Analytics and Math. Sci. Dept.
IBM T.J. Watson, USA

Postdoctoral Researcher, 2009-2010
Business Analytics and Math. Sci. Dept.
IBM T.J. Watson, USA

Research Assistant, 2005-2009
Computer and Inf. Sci. and Engg.
University of Florida, USA

Highlights

- One of the lead authors of a **Science** paper that was highlighted in AAAS meeting as a **breakthrough** in Olfaction research over the past 3 decades, 2017.
- Third Plateau Invention Achievement award for creative contributions to IBM, 2017.
- Appeared in IBM Journal of Eminence, 2016.
- IEEE ICDM Best Paper Candidate invited to KAIS Journal, 2015.
- Winner of (IBM) Shark Tank for proposing a fresh new direction with significant research and product potential, 2015.
- Second Plateau Invention Achievement award for creative contributions to IBM, 2015.
- Deployed Application (Paper) Award, in Innovative Applications of Artificial Intelligence (IAAI) track at Assoc. for the Adv. of Artificial Intelligence (AAAI) conference, 2015.
- Paper nominated for IBM Pat Goldberg Best Paper Award from the data mining area, 2014.
- Manager's Choice Award winner for conceptualizing, designing and developing analytics for IBM Risk Analysis Tool deployed across multiple organizations within IBM that daily analyzes millions of invoices and tens of thousands of vendors spread over 150 countries, 2014.
- Knowledge and Information Systems (KAIS) reviewer certificate, 2014.
- Appointed Knowledge Discovery and Data Mining Professional Interest Community (KDD PIC) *chair* for IBM TJ Watson Research, 2013.

Responsibilities include but are not limited to deciding on conference sponsorships, managing funds for visiting speakers, nominating candidates (with machine learning/data mining background) for IBM Phd fellowships, organizing workshops/speaker series and creating budget proposals for the following year.

- Co-lead (with Vikas Sindhwani) for Business Analytics and Mathematical Sci. (BAMS) Dept. in IBM research for creating, refining and selecting visionary research proposals that will be the big waves of the future. Our strategy had the *highest* number of submissions across IBM, 2013.
- First Plateau Invention Achievement award for creative contributions to IBM, 2012.
- Top IBMER award, for winning the 5th Annual IBM Chess Tournament, (researchers from the Deep Blue team also participated) 2011.
- First patent application invention achievement award presented by IBM, 2011.
- 2nd prize in all India project competition, Concepts 2004 sponsored by Microsoft, Cybage, Calsoft and other top companies.
- 2nd prize in all India Robotics competition, Techfest 2003 at IIT Bombay.
- Among the top 10 students out of a total of 12000+ B.E. students at Pune University, 2001.

Research and Teaching

IBM T.J. Watson

2009-present

- Interpretable modeling.
- Data driven understanding of human olfaction (smell).
- Building predictive models for condition based maintenance in complex instrumented domains such as the chip manufacturing industry, petroleum industry, automobile industry.
- Performing research in new model selection frameworks, online learning with multiple feature additions, active instance completion, actionable clustering, relational learning.
- Developing (Analytics lead) a comprehensive proprietary risk/fraud analysis solution for global procurement services, which encapsulates online learning, text analytics, social network analytics and unsupervised learning.
- Performing research in stochastic modeling for applications such as business compliance analytics, spend forecasting.
- Building automated learning models for electricity companies for theft detection.

Research Assistant at UF

2006-2009

- *Machine Learning and Data Mining:* a) Worked with Dr. Alin Dobra on semi-analytical methods for analyzing classification models and model selection measures in the non-asymptotic regime. This included theory and efficient computation using Monte Carlo and non-linear optimization techniques. b) Worked on issues related to collective classification in Statistical Relational Learning. These issues included comparing collective and independent classification and finding distribution free bounds in the relational setting.
- *Fuzzy Systems:* Worked with Dr. Paul Gader on analyzing a generic aggregation function namely, the Choquet Integral.

Teaching Assistant at UF

Fall 2007

- Course: Discrete Mathematics
Responsibilities included lecturing, grading and holding office hours. The instructor reviewed my performance as *excellent*.

Centre for Development of Advanced Computing (C-DAC)

2003-2004

- Developed a proprietary character recognition algorithm for printed Devanagari script documents. Devanagari script has many more characters than the Roman script. Moreover, combination of these characters is also allowed to produce composite characters which makes the problem of efficient and accurate recognition challenging.

Patents

Granted

- **Amit Dhurandhar** and Jun Wang. A System and Method for Relational Transductive Learning. *U.S. Patent Serial No. US 9355367*.
- **Amit Dhurandhar** and Jayant Kalagnanam. Multistep Time Series Prediction in Complex Instrumented Domains. *US Patent Serial No. US20120150489 A1*.
- Pawan Chowdhary, **Amit Dhurandhar**, Markus Ettl, Soumyadip Ghosh, Bruce Graves, Bill Schaefer and Yu Tang. Method and system for optimizing procurement spend compliance. *US Patent Serial No. US8732090 B2*.
- **Amit Dhurandhar**. Improving Predictions using Aggregate Information. *US Patent Serial No. US8762314 B2*.
- Robert Baseman, **Amit Dhurandhar**, Sholom Weiss and Brian White. A System and Method for Continuous Prediction of Expected Chip Performance Throughout the Production Lifecycle. *US Patent Serial No. US8793106 B2*.

Pending

- Guillermo Cecchi, **Amit Dhurandhar**, Etkin Gutterez and Pablo Meyer. Prediction of Olfactory and Taste Perception through Semantic Encoding. *P20170003US01*.
- Ioana Baldini, **Amit Dhurandhar**, Abhishek Kumar, Aleksandra Mojsilovic, Kein T Pham, Kush R Varshney and Maja Vukovic. Humanitarian Crisis analysis using secondary information gathered by focused web crawler. *YOR920161631*.
- **Amit Dhurandhar**, Sechan Oh and Marek Petrik. Interpretable Rule Generation using loss-preserving transformation. *U.S. Provisional Patent Application Serial No. 15/489,418*.
- **Amit Dhurandhar**, Bruce Graves, Rajesh Ravi and Markus Ettl. A System and Method for Identifying Procurement Fraud/Risk. *US Provisional Patent Serial No. 14/186,071*.
- **Amit Dhurandhar**, Stuart Seigal, Yada Zhu and Jayant Kalagnanam. A System and Method for Detecting Electricity Theft via Meter Tampering Using Statistical Methods of Anomaly Detection. *US Provisional Patent Serial No. 13/909,239*.
- **Amit Dhurandhar**, Robert Baseman and Fateh Tipu. A System and Method for Identifying Significant Consumable Insensitive Trace Features. *YOR8020140228*.
- **Amit Dhurandhar**, Bruce Graves, Rajesh Ravi, Gopikrishnan Maniachari, Markus Ettl, Anthony Mazzatti. A Robust System for Ranking and Tracking Suspicious Procurement Entities. *YOR8020140315*.
- Guillermo Cecchi, **Amit Dhurandhar**, Stacey M Gifford, Raquel Norel, Pablo Meyer, Kahn Rhrissorakrai and Bo Zhang. Predicting User Preferences based on Olfactory Characteristics. *YOR8020161333*.
- Guillermo Cecchi, **Amit Dhurandhar** and Pablo Meyer. Correlating Olfactory Perception with Molecular Structure. *YOR920161332*.

Publications

([†] implies equal contribution)

Journals (Peer Reviewed)

- Andreas Keller, Richard Gerkin, Yuanfang Guan, **Amit Dhurandhar**, Gabor Turu, Bence Szalai, Joel Mainland, Russ Wolfinger, Celine Vens, Leander Schietgat, Kurt De Grave, Raquel Norel, Gustavo Stolovitzky, Guillermo Cecchi, Leslie Vosshall, Pablo Meyer. Predicting Human Olfactory Perception from Chemical Features of Odor Molecules. *Science*, 2017. (Highlighted in AAAS meeting as a breakthrough in the last 3 decades in olfactory research.)
- Real-Time Understanding of Humanitarian Crises via Targeted Information Retrieval. Kien T. Pham, Prasanna Sattigeri, **Amit Dhurandhar**, Arpith C. Ja-

- cob, Maja Vukovic, Patrice Chataigner, Juliana Freire, Aleksandra Mojsilovi, and Kush R. Varshney. *IBM Journal of Research and Development*, 2017.
- Tsuyoshi Ide and **Amit Dhurandhar**. Supervised Item Response Models for Informative Prediction. *Knowledge and Information Systems (KAIS)*, 2016. (invited)
 - **Amit Dhurandhar** and Karthik Sankarnarayanan. Improving Classification Performance through Selective Instance Completion. *Machine Learning Journal (MLJ)*, 2015. (with presentation slot at ECML 2015.)
 - **Amit Dhurandhar** and Marek Petrik. Efficient and Accurate Methods for Updating Generalized Linear Models with Multiple Feature Additions. *Journal of Machine Learning Research (JMLR)*, 2014.
 - **Amit Dhurandhar**. Bounds on the Moments for an Ensemble of Random Decision Trees. *Knowledge and Information Systems (KAIS)*, 2014.
 - Sholom Weiss, **Amit Dhurandhar**, Robert Baseman, Brian White, Ronald Logan, Jonathan Winslow and Daniel Poindexter. Continuous Prediction of Manufacturing Outcomes Throughout the Production Lifecycle. *Journal of Intelligent Manufacturing (JIMS)*, 2014.
 - **Amit Dhurandhar** and Jun Wang. Single Network Relational Transductive Learning. *Journal of Artificial Intelligence Research (JAIR)*, 2013.
 - **Amit Dhurandhar**. Using Coarse Information for Real Valued Prediction. *Data Mining and Knowledge Discovery (DMKD)*, 2013. (nominated for IBM Pat Goldberg Award)
 - **Amit Dhurandhar** and Alin Dobra. Probabilistic Characterization of Nearest Neighbor Classifier. *International Journal of Machine Learning and Cybernetics (IJMLC)*, 2012. (invited)
 - **Amit Dhurandhar** and Alin Dobra. Distribution free bounds for Relational Classification. *Knowledge and Information Systems (KAIS)*, 2012.
 - **Amit Dhurandhar** and Alin Dobra. Semi-analytical Method for Analyzing Models and Model Selection Measures based on Moment Analysis. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 2009.
 - **Amit Dhurandhar** and Alin Dobra. Probabilistic Characterization of Random Decision Trees. *Journal of Machine Learning Research (JMLR)*, 2008.
 - **Amit Dhurandhar** and Alin Dobra. Test Set Bounds for Relational data that vary with Strength of Dependence. submitted

Conferences/Workshops (Peer Reviewed)

(Conference acceptance rates are typically < 30%)

- **Amit Dhurandhar**, Vijay Iyengar, Ronny Luss and Karthikeyan Shanmugam. TIP: Typifying the Interpretability of Procedures. submitted
- Karthik Gurumoorthy, **Amit Dhurandhar** and Guillermo Cecchi. ProtoDash: Fast Interpretable Prototype Selection. submitted
- **Amit Dhurandhar**, Vijay Iyengar, Ronny Luss and Karthikeyan Shanmugam. A Formal Framework to Characterize Interpretability of Procedures. *Human Interpretability in Machine Learning workshop in Intl. Conference on Machine Learning (ICML)*, 2017.
- **Amit Dhurandhar**, Margareta Ackerman and Xiang Wang. Uncovering Population Trends with Accordant Clustering. *SIAM Intl. Conference on Data Mining (SDM)*, 2017. [Oral]
- **Amit Dhurandhar**, Sechan Oh and Marek Petrik. Building Interpretable Recommender via Loss-Preserving Transformation. *Human Interpretability in Machine Learning workshop in Intl. Conference on Machine Learning (ICML)*, 2016.
- **Amit Dhurandhar**, Bruce Graves, Rajesh Ravi, Gopikrishnan Maniachari and Markus Ettl. Big Data System for Analyzing Risky Entities. *ACM SIGKDD conference on data mining (KDD)*, 2015. [Oral]

- Tsuyoshi Ide and **Amit Dhurandhar**. Informative and Accurate Prediction based on Ordinal Questionnaire Data. *IEEE Intl. conference on data mining (ICDM), 2015. (Best Paper Candidate)*
- **Amit Dhurandhar**, Rajesh Ravi, Bruce Graves, Gopikrishnan Maniachari and Markus Ettl. Robust System for Identifying Procurement Fraud. *Innovative Applications of Artificial Intelligence track in Assoc. for the Adv. of Artificial Intelligence (AAAI), 2015. (Deployed Application Award)*
- **Amit Dhurandhar** and Karthik Gurumoorthy. Symmetric Submodular Clustering with Actionable Constraint. *Discrete Optimization workshop in Adv. of Neural Inf. Proc. Systems (NIPS), 2014.*
- Rajesh Ravi, **Amit Dhurandhar**, Markus Ettl, Bruce Graves. Procurement Fraud Risk Analytics Tool. *Information on Demand Conference (IOD), 2014.*
- Sholom Weiss, **Amit Dhurandhar** and Robert Baseman. Improving Quality Control by Early Prediction of Manufacturing Outcomes. *ACM SIGKDD conference on data mining (KDD), 2013. [Oral]*
- Karthik Sankarnarayanan[†] and **Amit Dhurandhar**[†]. Intelligently Querying Incomplete Instances for Improving Classification Performance. *ACM International Conference on Information and Knowledge Management (CIKM), 2013 . [Full paper, Oral]*
- **Amit Dhurandhar**. Auto-correlation Dependent Bounds for Relational Data. *Mining and Learning over Graphs workshop in ACM SIGKDD conference on data mining (KDD), 2013.*
- **Amit Dhurandhar**. Improving Predictions using Aggregate Information. *ACM SIGKDD conference on data mining (KDD), 2011.*
- Pawan Chowdhary, Markus Ettl, **Amit Dhurandhar**, Soumyadip Ghosh, Gopikrishna Maniachari, Bruce Graves, Bill Schaefer and Yu Tang. Identify and Manage Procurement Savings using Advanced Compliance Analytics. *IEEE International Conference on e-Business Engineering (ICEBE), 2011 .*
- **Amit Dhurandhar**. Multistep Time Series Prediction in Complex Instrumented Domains. *Large scale analytics in complex instrumented domains workshop in IEEE International Conference on Data Mining (ICDM), 2010. This paper was also invited to Chance Discovery workshop in (IJCAI), 2011.*
- **Amit Dhurandhar**. Learning Maximum Lag for Grouped Graphical Granger Models. *Knowledge Discovery from Climate Data Prediction, Extremes, and Impacts workshop in IEEE International Conference on Data Mining (ICDM), 2010.*
- Robert Baseman, **Amit Dhurandhar**, Michal Ozery and Naama Perush. Statistical Assessment of dissimilarities in trace data of unusual and nominal wafers. *ISMI Manufacturing Week, 2010.*
- Dan Connors, **Amit Dhurandhar**, Markus Ettl, Mary Helander, Jayant Kalagnanam, Shubir Kapoor, Ramesh Natarajan, Stuart Seigal, Zhackary Xue. Demand forecasting and supply chain optimization using freshness. *Information on Demand Conference (IOD), 2010.*
- Robert Baseman, Frances Clougherty, **Amit Dhurandhar**, Lyndon Logan, Daniel Poindexter, Brian White, Sholom Weiss, Jonathan Winslow, Denis Zhereschin. Early Predictions of Device Performance for Enhanced Process Control and Operations Optimization. *ISMI Symposium on Manufacturing Excellence, 2010.*
- John Andrews, Robert Baseman, Michael Biagetti, **Amit Dhurandhar**, Hong Lin, Michal Ozery-Flato, Stuart A Siegel, Naama Parush-Shear-Yashuv, Adam Ticknor. Utilization of Equipment Trace Data in a 300mm Semiconductor Fab. *ISMI Symposium on Manufacturing Excellence, 2010.*
- **Amit Dhurandhar** and Alin Dobra. Evaluating Evaluation Measures. *Evaluation Methods in Machine Learning workshop in International Conference on Machine Learning (ICML), 2009.*

- **Amit Dhurandhar** and Alin Dobra. Study of Classification Algorithms using Moment Analysis. *New Challenges in Theoretical Machine Learning workshop in Neural Information Processing Systems (NIPS), 2008.*
- **Amit Dhurandhar**, Kartik Shankar and Rakesh Jawale. Robust Pattern Recognition Scheme for Devanagari Script. *IEEE International Conference on Computational Intelligence and Security (CIS) 2005.*

Technical Reports

- **Amit Dhurandhar**, Steve Hanneke and Liu Yang. Learning with Changing Features.
- **Amit Dhurandhar** and Paul Gader. Output Distribution of Choquet Integral.
- **Amit Dhurandhar** and Alin Dobra. Insights into Cross-validation.
- **Amit Dhurandhar** and Alin Dobra. Independent vs Collective Classification in Statistical Relational Learning.

Book Chapters

Helped add/edit chapter on Generalized Linear Models (Aug. 2014) in book *Data Classification*, by Charu Aggarwal.

Invited Talks/Presentations

- Typifying Interpretability of Procedures, *CMU, 2017.*
- Real Time Personalized Recommendations, *Lowes, 2016.*
- Lessons Learned from DREAM Olfaction Challenge, *RECOMB RSG, 2015.*
- Robust System for Identifying Procurement Fraud. *AAAI, 2015.*
- Symmetric Submodular Clustering with Actionable Constraint, *NIPS, 2014.*
- IBM Risk Analytics Tool, *Volkswagen, 2014.*
- Intelligently Querying Incomplete Instances for Improving Classification Performance, *ACM CIKM 2013.*
- Intelligently Querying Incomplete Instances for Improving Classification Performance, *UC Davis 2013.*
- Intelligently Querying Incomplete Instances for Improving Classification Performance, *Apple Inc. 2013.*
- Procurement Fraud/Risk Detection, *Walmart (CIO and top execs.), 2013.*
- Tracer: Large Scale Monitoring of Process Trace Data for IBM Semiconductor Manufacturing, *IBM Lab Around, 2013.*
- Improving Quality Control by Early Prediction of Manufacturing Outcomes, *KDD, 2013.*
- Auto-correlation Dependent Bounds for Relational Data, *KDD, 2013.*
- Condition Monitoring and Predictive Maintenance, *Workshop with Samsung, 2011.*
- Improving Predictions using Aggregate Information, *KDD, 2011.*
- Multistep Time Series Prediction in Complex Domains, *ICDM, 2010.*
- Learning Maximum Lag for Grouped Graphical Granger Models, *ICDM, 2010.*
- Condition based monitoring, *Workshop with Petrobras, 2010.*
- Semi-analytical Method for Analyzing Models and Model Selection Measures, *IBM T.J. Watson Research, 2009.*
- Evaluating Evaluation Measures, *ICML, 2009.*
- Semi-analytical Method for Analyzing Models and Model Selection Measures, *Yahoo Research, 2008.*
- Study of Classification Algorithms using Moment Analysis, *NIPS, 2008.*

**Professional
Activities**

Grant Panelist

National Science Foundation - Small Business Innovative Research (**NSF-SBIR**) grant panelist, 2016-2017.

Organizing

- Helped organize 2nd DREAM ALS Stratification Challenge. This was an international competition with tens of thousands of dollars cash prize with the hope making strides in terms predicting ALS disease progression, 2015.
- **Main organizer of KDD Speaker Day** at IBM TJ Watson Research on varied topics in machine learning and data mining. Speakers included Jieping Ye, Geoff Webb, Dafna Shahaf, Jennifer Neville and Cynthia Rudin, 2014.

Reviewing

- **Area Chair/Session Chair Invitations:** ICLR 2018, CIKM 2013, IJCAI 2015
- **Program Committee Member:** KDD 2017, NIPS 2017, ICML 2017, KDD 2016 (both tracks), IJCAI 2016, NIPS 2016, IJCAI 2015, ICML 2015, NIPS 2015, Advanced Manufacturing Workshop in IEEE Conference on Big Data 2015.
- **Journal Reviewing:** Pattern Recognition (PR), Journal of Machine Learning Research (JMLR), IEEE Transactions on Information Theory (IEEE-IT), ACM Transactions on Knowledge Discovery in Data (TKDD), Applied Intelligence (APIN), Knowledge and Inf. Systems (KAIS), Statistical Analysis and Data Mining (SAM), Intl. Journal on Machine Learning and Cybernetics (IJMLC), Intelligent Systems in Accounting, Finance and Management.
- **Conference Reviewing:** ACM conference on Knowledge Discovery and Data Mining (KDD), Siam Conference on Data Mining (SDM), Intl. Conference on Machine Learning (ICML).

Skills

- Proficient in major programming languages and platforms.
- Strong mathematical skills (in particular statistical analysis).
- Fluent in English, Hindi and Marathi.

**Other Activities
and Leadership**

- Volunteered at Pets Alive no-kill animal shelter.
- Led sessions in the Algorithm Theory Group Meetings in Fall 2008.
- Co-ordinator for the Database Center Weekly Seminars in Spring 2008.
- College team captain of the chess team in undergrad.
- College team member of the badminton team in undergrad.

References

Will be made available on request.