



Curriculum Vita

Daniel T. Larose, Ph.D.

Professor of Statistics and Data Mining
Founder, Data Mining @CCSU
Department of Mathematical Sciences
Central Connecticut State University
New Britain, Connecticut, 06050

Marcus White 105 Phone: 860-204-1661

EEmail: larosed@ccsu.edu Web: www.ccsu.edu/faculty/larose

Education

Ph.D. in Statistics, August 1996, University of Connecticut, Storrs.

Dissertation: *Bayesian Approaches to Meta-Analysis*. Advisor: Dr. Dipak K. Dey.

Master of Science in Statistics, May 1993, University of Connecticut, Storrs, CT.

Bachelor of Arts in Mathematics, May 1990, Westfield State College, Westfield, MA.

Bachelor of Science in Computer Science and Linguistics, May 1979.

Honors: *Natural Language Processing*, University of Massachusetts, Amherst, MA.

Employment

August 2006 to Present: Professor of Statistics and Data Mining

August 2001 to August 2006: Associate Professor of Mathematical Sciences

August 1996 to August 2001: Assistant Professor of Mathematical Sciences

Department of Mathematical Sciences, CCSU

July 2006 to Present: Book Series Editor: *Wiley Series on Methods and Applications in Data Mining*, John Wiley and Sons, Inc., Hoboken, NJ

May 2006 to Present: Owner, *Data Mining Consultant.com, LLC*

January 2001 to August, 2016: Coordinator, Data Mining @CCSU

Master of Science and Graduate Certificate Programs

Credit Load Activity

Courses Taught

Because I developed the curricula for the Data Mining programs here at CCSU (see below, Service), I teach the following graduate courses:

- **Stat 521 *Introduction to Data Mining***
- **Stat 522 *Clustering and Affinity Analysis***
- **Stat 523 *Predictive Analytics***
- **Stat 525 *Web Mining***
- **Stat 529 *Special Topics: Enhancing Classification Model Performance***

I also teach undergraduate statistics, including

- **Stat 104 *Elementary Statistics***
- **Stat 200 and 201 *Business Statistics 1 and 2***
- **Stat 215 *Statistics for the Behavioral Sciences*.**

Master's Thesis Advisees

Since 2001, I have served as Master's Thesis Advisor for about 70% of all Data Mining graduates.

Fall 2017:

- Rajini Prabhu, *Predictive Modeling of Credit Card Default*.

Fall 2016:

- Brian Ironside, *Improving the Performance of Ensemble Classifier Models through the Local Specialization of Base Classifiers*.

Spring 2016:

- James Cunningham, *Employing Data Mining Methods to Assess the Efficacy of Classical Credit Risk Models*.
- Cathy Farrell, *Trinary Predictive Classification of Diabetic Episode Recurrence*.

Fall 2015:

- Tarek Abdel-Azim, *Performance of Various Base and Ensemble Clustering Approaches*.
- Curtis Lentz, *Data Analytic Approaches to Predicting Success in Bank Telemarketing*.

Spring 2015:

- Xavier Antony, *Classification of Fetal Heart Rate Patterns*.
- Marc Glettenberg, *Predictive Classification of Skill Level Using Telemetric Data*.

Fall 2014:

- Eric Flores-Acosta, *Implementation of Stability-Based Cluster Validation Using Prediction Strength*.
- Benjamin Dickman, *Using Data Mining to Predict Success in a Nursing Program*.
- Andrew Hendrickson, *Classification Modeling of Freddie Mac Home Loan Delinquency*.
- Jeffrey Richardson, *Applying Association Rules to Optimize Enterprise Software Development*.

Master's Thesis Advisees (continued)

Spring 2014:

- John Almeda, *Using Predictive Analysis to Enhance the Efficiency of Commuter Transportation Networks.*
- Richard Aceves, *Predicting Change in County-Level Presidential Election Voter Turnout Using Data Mining Methods.*
- Jill Willie, *Assessment of Similarity-Based Cluster Validation Methods.*
- Sairam Tadigadapa, *The Application of Decision Trees for Diagnosing Liver Disease.*
- Paolo Carbone, *Distributor Price Optimization Using Market Segmentation.*

Fall 2013:

- Thomas Wilk, *Improving Workplace Accident Fatality Classification Models with Text Mining and Ensemble Methods.*
- George DeVarenes, *Applying Cost-Benefit Analysis to a Trinary Classification Model.*
- Daniel Aloï, *Using Crime Prevention Models to Aid Law Enforcement in Resource Allocation and Decision Making.*
- William Rowe, *Classifying Web Pages by Image Attributes.*

Spring 2013

- Martin Couture, *Applying Data Mining Techniques in Classifying Personal Automobile Insurance Risk.*
- Steven Cultrera, *Analysis of the Impact of Weather on Runs Scored in Baseball Games at Fenway Park.*
- Kay Batta, *Applying Misclassification Costs to Ameliorate the False Positive Rate in Bioassay Screening.*

Fall 2012

- Senthil Murugan, *Mining for Profitable Low-Risk Delta-Neutral Long Straddle Options Strategies.*
- Rajiv Sambisavan, *Modeling of Flight Delays.*
- Edwin Rivera, *Anti-Money Laundering Behavior: Reducing the Number of Non-Productive Alerts in Structuring through Effective Data Mining.*

Spring 2012

- Judith Gu, *Using Data Mining to Model Market Reaction to Corporate Earnings Announcements.*
- Sampson Adu-Poku, *Comparing Classification Algorithms in Data Mining.*

Prior to 2012:

- Thierry Vallaud, *Estimating Customer Potential Value Using Classification of Customer Data.*
- Kathleen Alber, *Identifying Patterns of Potentially Preventable Emergency Department Utilization by American Children.*
- Steven Barbee, *The Discovery by Data Mining of Rogue Equipment in the Manufacture of Semiconductor Devices.*
- Eric Taylor, *Comparing Unsupervised Multivariate Normal Cluster Results between Datasets and Consolidating Similar Clusters.*
- James Steck, *NETPIX, A Method of Feature Selection Leading to Accurate Sentiment-Based Classification Models.*
- Rafiqul Islam, *Knowledge Discovery in Microarray Data.*

Creative Activity

In the eleven year period, 2005 – 2016, I published ten textbooks.

Creative Activity - Statistics Textbooks – Part 1

Discovering Statistics

Third Edition,

By Daniel Larose,

W. H. Freeman Publishers, 2016.

Discovering Statistics covers the two-semester introductory statistics sequence, while implementing the GAISE Guidelines. These Guidelines include:

1. Emphasize statistical literacy and develop statistical thinking.
2. Use real data.
3. Stress conceptual understanding instead of mere knowledge of procedures.
4. Foster active learning in the classroom.
5. Use technology for developing conceptual understanding and analyzing data.

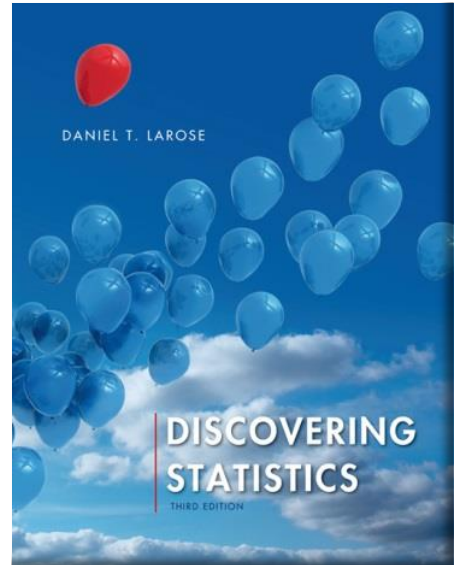
The focus of *Discovering Statistics* is to **encourage student motivation**, for example, through using data of interest to students, such as video games, cell-phone apps, and music videos. Helpful features, such as *Check it Out* and *Now You Can Do* clarify for the student which examples go with which exercises, and vice-versa, thus helping students learn.

Other features include *The Big Picture*, *Your Turn*, *Matched Objectives*, *Bringing it All Together*, *Developing Your Statistical Sense*, *What Does this Number Mean*, *What-If Scenarios*, *What Results Might We Expect*, new Chapter Case Studies, and **Step-by-Step Technology Guides** covering the TI-83/84, Excel, Minitab, SPSS, JMP, and CrunchIt technology. The Step-by-Step Technology Guides were written by my daughter, Chantal Larose, PhD.

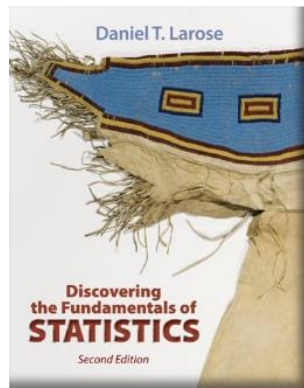
60% of the examples and 50% of the exercises are new or revised for the Third edition. Over 4,100 exercises are now included in the book.

The book chapters are as follows:

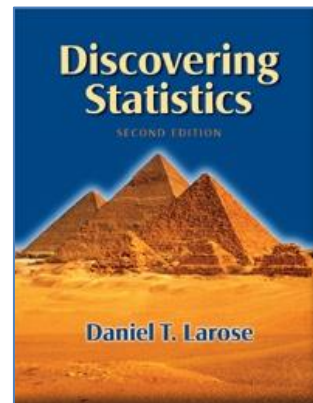
1. The Nature of Statistics
2. Describing Data Using Graphs and Tables
3. Describing Data Numerically
4. Correlation and Regression
5. Probability
6. Probability Distributions
7. Sampling Distributions
8. Confidence Intervals
9. Hypothesis Testing
10. Two-Sample Inference
11. Categorical Data Analysis
12. Analysis of Variance
13. Inference in Regression
14. Nonparametric Statistics



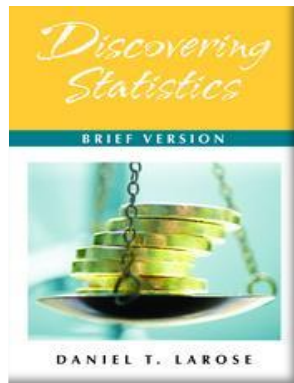
Creative Activity - Statistics Textbooks – Part 2



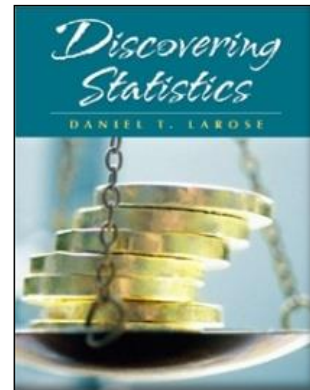
*Discovering the Fundamentals
of Statistics*
Second Edition
W.H. Freeman, 2014



Discovering Statistics
Second Edition
W.H. Freeman, 2013



*Discovering Statistics,
Brief Edition*
W.H. Freeman, 2011

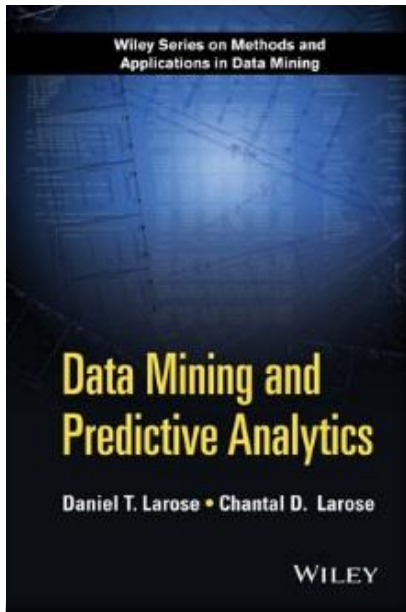


Discovering Statistics
First Edition
W.H. Freeman, 2009

For each of these editions, about 50% of the examples and 50% of the exercises are revised from the previous edition.

A custom edition of *Discovering Statistics, Second Edition* has been adopted by the Department of Mathematical Sciences for the course Stat 104 *Elementary Statistics*. Prior to 2016, the CCSU royalties for all sales of Dr. Larose's statistics and data mining books, totaling \$8,500, had been donated to the Larose Scholarship for Excellence and Service in Data Mining.

Books on Data Mining and Analytics – Part 1



Data Mining and Predictive Analytics Second Edition

Daniel Larose and Chantal Larose
John Wiley and Sons, Hoboken, NJ, 2015

Data Mining and Predictive Analytics is a comprehensive tome detailing data mining and predictive analytics from start to finish.

This textbook has been adopted by the Department of Mathematical Sciences for these courses:

- Stat 521 Intro to Data Mining
- Stat 522 Clustering and Affinity Analysis
- Stat 523 Predictive Analytics

Table of Contents

Part 1: Data Preparation

1. An Introduction to Data Mining and Predictive Analytics
2. Data Preprocessing
3. Exploratory Data Analysis
4. Dimension Reduction Methods

Part 2: Statistical Analysis

5. Univariate Statistical Analysis
6. Multivariate Statistics
7. Preparing to Model the Data
8. Simple Linear Regression
9. Multiple Regression and Model Building

Part 3: Classification

10. K-Nearest Neighbor Algorithm
11. Decision Trees
12. Neural Networks
13. Logistic Regression
14. Naive Bayes and Bayesian Networks
15. Model Evaluation Techniques
16. Cost-Benefit Analysis Using Data-Driven Costs
17. Cost-Benefit Analysis for Trinary and k-Nary Classification Models
18. Graphical Evaluation of Classification Models

Part 4: Clustering

19. Hierarchical and K-Means Clustering
20. Kohonen Networks
21. BIRCH Clustering
22. Measuring Cluster Goodness

Part 5: Association Rules

23. Association Rules: A Priori Algorithm and Generalized Rule Induction

Part 6: Enhancing Model Performance

24. Segmentation Models
25. Ensemble Methods: Bagging and Boosting
26. Model Voting and Propensity Averaging

Part 7: Further Topics

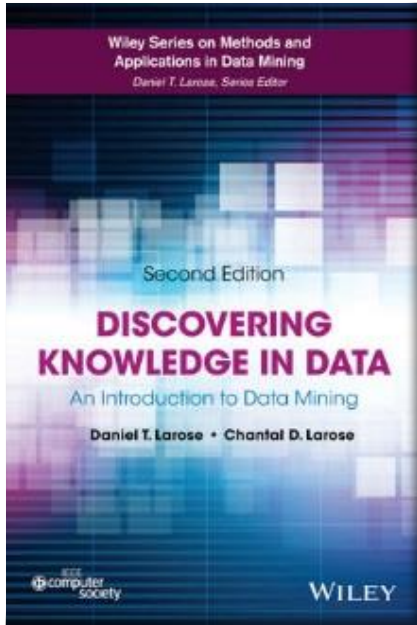
27. Genetic Algorithms
28. Imputation of Missing Data

Part 8: Case Study: Predicting Response to Direct-Mail Marketing

29. Case Study, Part 1: Business Understanding, Data Preparation, and EDA
30. Case Study, Part 2: Clustering and Principal Components Analysis
31. Case Study, Part 3: Modeling and Evaluation for Performance and Interpretability
32. Case Study, Part 4: Modeling and Evaluation for High Performance Only

Appendix A: Data Summarization and Visualization

Books on Data Mining and Analytics – Part 2



Discovering Knowledge in Data: An Introduction to Data Mining Second Edition

Daniel Larose and Chantal Larose
John Wiley and Sons, Hoboken, NJ, 2014
IEEE Computer Society

Discovering Knowledge in Data: An Introduction to Data Mining provides readers with:

- The models and techniques to uncover hidden nuggets of information,
- The insight into how the data mining algorithms really work, and
- The experience of actually performing data mining on large data sets.

The R Zone, provided in every chapter, equips readers with the R code required to do the analysis shown in the book.

Table of Contents

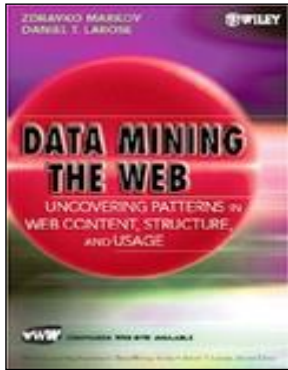
- 1. An Introduction to Data Mining**
- 2. Data Preprocessing**
- 3. Exploratory Data Analysis**
- 4. Univariate Statistical Analysis**
- 5. Multivariate Statistics**
- 6. Preparing to Model the Data**
- 7. K-Nearest Neighbor**
- 8. Decision Trees**
- 9. Neural Networks**
- 10. Hierarchical and K-Means Clustering**
- 11. Kohonen Networks**
- 12. Association Rules**
- 13. Imputation of Missing Data**
- 14. Model Evaluation Techniques**

Larose's books on data mining and predictive analytics were originally written to support the new programs in data mining that he developed for CCSU in the early 2000's.

Now, the books are used at campuses around the country and around the world.

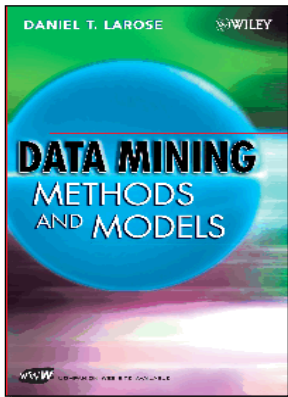
For more information about these books, please feel free to visit DataMiningConsultant.com.

Books on Data Mining and Analytics – Part 3



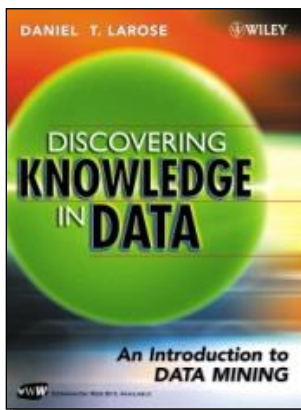
Data Mining the Web:
Uncovering Patterns in Web Content, Structure, and Usage
(with Dr. Zdravko Markov of CCSU's Computer Science Department)
John Wiley and Sons, Hoboken, NJ, 2007

This book has been adopted for Stat 525 *Web Mining*.



Data Mining Methods and Models
John Wiley and Sons, Hoboken, NJ, 2006

Formerly, this book had been adopted for Stat 521, 522, and 523, but has now been superseded by *Data Mining and Predictive Analytics*.



Discovering Knowledge in Data:
An Introduction to Data Mining
John Wiley and Sons, Hoboken, NJ, 2005

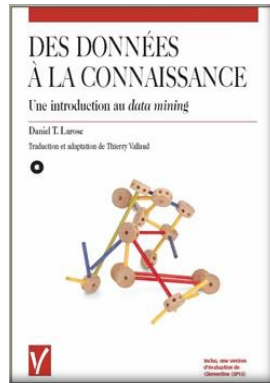
Formerly, this book had been adopted for Stat 521, 522, and 523, but has now been superseded by *Data Mining and Predictive Analytics*.

Translations into French and Polish



Exploration de Donnees: Methodes et Modeles de Data Mining

French translation of
Data Mining Methods and Models
Les Editions Vuibert, France, 2012



Des Donnees a la Connaissance

French translation of
Discovering Knowledge in Data
Les Editions Vuibert, France,
2005

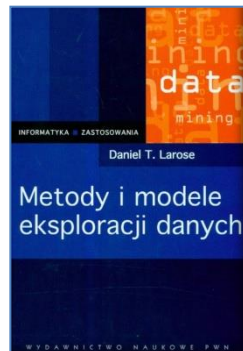
The French translations were done by Thierry Vallaud, Paris resident, and graduate of CCSU's online Master of Science in Data Mining.

Thierry is now Director of Data Science, BVA, Paris.



Eksploracja Zasobow Internetowych

Polish translation of
Data Mining the Web
Wydawnictwo Naukowe Pwn,
2010



Metody i Modele Eksploracji Danych

Polish translation of
*Data Mining Methods and
Models*
Wydawnictwo Naukowe Pwn,
2008



Odkrywanie Wiedzy z Danych

Polish translation of
*Discovering Knowledge in
Data*
Wydawnictwo Naukowe Pwn,
2006

Other Book

Ready, Set, Run! A Student Guide to SAS Software for Windows

(with Chun Jin, PhD from the Department of Mathematical Sciences), McGraw-Hill, 1998.

Research Publications

Sensitivity and clinical utility of the anti-cytosolic 5'-nucleotidase 1A (cN1A) antibody test in sporadic inclusion body myositis: report of 40 patients from a single neuromuscular center, by Dr Allan Metzger, Dr Charles Whitaker, Dr Guo Shen, Dr Randall Barton, Dr Daniel Larose, and Dr Qian Wu. **Neuromuscular Disorders**. August, 2018.

The Role of Patient History and Body Site Surveillance Cultures as Predictors of Colonization in a Long-Term Acute Care Hospital Setting, by Brenda Nurse, MD, Randall Barton, PhD, and Daniel Larose. **Connecticut Medicine**. February, 2017.

Developing Artificial Neural Networks to Predict Function One Year After Traumatic Spinal Cord Injury, by Timothy Belliveau, PhD, Alan Jette, PT, PhD, Subramani Seetharama, MD, Jeffrey Axt, MBA, David Rosenblum, MD, Daniel Larose PhD, Bethlyn Houlihan, MSW, Mary Slavin, PT, PhD, and Chantal Larose, PhD, **Archives of Physical Medicine and Rehabilitation**, October, 2016.

Low Rate of Bacteremia with a Subcutaneously Implanted Central Venous Access Device, with Brenda A. Nurse, MD, Rita Bonczek, APRN, and Randy Barton, PhD, **The Journal of Vascular Access**, Volume 15, Number 1, 51 – 55, 2013.

Status: A Tool for Understanding Intra-Family Conflict
by Daniel Wiener (CCSU Professor of Counseling and Family Therapy) and Daniel Larose. **International Journal of Humanities and Social Science**, Volume 1, Number 10, 36 – 46, 2011.

Constrained Statistical Decisions in Evolving Environments
by Elijah Gaioni, Dipak Dey, and Daniel Larose. **Journal of Intelligent Systems**, Volume 18, Issue 3, 171 – 192, 2011.

A Bayesian Meta-Analysis of the Relationship Between Duration of Estrogen Exposure and Endometrial Cancer
in *Meta-Analysis in Medicine and Health Policy*, D. Berry and D. Stangl, Eds., Marcel-Dekker, 2000.

Hands-On Statistics: Fun Activities for Learning Inference
Proceedings of the Section on Statistics Education, American Statistical Association, 2000.

Bayesian Duration-Response Meta-Analysis of Estrogen and Endometrial Cancer
Proceedings of the Section on Bayesian Statistical Science, American Statistical Association, 1998.

Modeling Publication Bias Using Weighted Distributions in a Bayesian Framework
by Daniel Larose and Dipak Dey. **Computational Statistics and Data Analysis**, Vol 26, 279-302, 1998.

Grouped Random Effects Models for Bayesian Meta-Analysis
by Daniel Larose and Dipak Dey. **Statistics in Medicine**, Vol. 16, 1817-1829, 1997.

Weighted Distributions Viewed in the Context of Model Selection: a Bayesian Perspective by Daniel Larose and Dipak Dey. **Test, The Journal of the Spanish Statistical Society**, Vol. 5, 1, 1996.

Publication Bias Models Using Weighted Distributions in a Bayesian Setting
Proceedings of the Section on Bayesian Statistical Science, American Statistical Association, 1995.

Modeling Heterogeneity and Extraneous Variation using Weighted Distributions
by Fengchun Peng, Dipak Dey, and Daniel Larose. *Model-Oriented Data Analysis 4*, Kitsos and Muller, eds, Physica Verlag, Heidelberg, Germany, 1995.

Selected Presentations

Gender, Age, Ethnicity and Student Performance in Online Statistics Courses

AMATYC Annual Conference, Minneapolis, MN, November 2007.

Invited Author Presentation: Data Mining Methods and Models

62nd Annual Deming Conference on Applied Statistics, Atlantic City, NJ, December 2006.

Invited Speaker, Session Chair and Organizer: Web Usage Mining

International Chinese Statistical Association Annual Conference, UConn, Storrs, June 2006.

Invited Speaker: Getting from Data to Knowledge

Center for Internet Data Research and Intelligent Systems, UConn School of Business, February 2006.

Demographic Factors and Student Performance in Online Statistics Courses

June Baker Higgins Gender Studies Conference, CCSU, April, 2005.

Invited Speaker: An Introduction to the CART and C4.5 Algorithms in Data Mining

New England Statistics Symposium, University of Connecticut, Storrs, CT, April, 2005.

Invited Speaker: Modeling Response to Direct Mail Marketing Using Data Mining Methods

Mini-Conference on Data Mining, Conn of American Statistical Association, Bristol-Myers / Squib, Wallingford, CT, March, 2004.

Using Item Response Theory as a Tool to Assess the Science Prowess of Connecticut's Children

Connecticut Academy for Education in Math, Science and Technology, Middletown, Conn., September 2000.

Bayesian Duration-Response Meta-Analysis of Estrogen and Endometrial Cancer

Joint Statistical Meetings, Dallas, Texas, August, 1998.

Grouped Random Effects Models

Joint Statistical Meetings, Anaheim, California, August, 1997.

Is Estrogen Linked to Cancers of the Breast and Endometrium?

New England Statistics Symposium, Storrs, Connecticut, April, 1997.

Extramural Grants

November 2015

Optimal identification of ALS-related cognitive and behavioral decline using home-based cognitive monitoring and artificial neural network predictive models,

Kevin Felice, DO, and Timothy Belliveau, PhD, Principal Investigators, Daniel Larose, PhD, Biostatistician. Amyotrophic Lateral Sclerosis Association Clinical Management Grant, 2016. \$96,693. Unsuccessful.

January 2007 – December 2012, \$27,000

W.H. Freeman Publisher's Grant to assist in the writing of *Discovering Statistics* and to develop the statistics online courseware *StatsPortal*.

September 2004 - December 2006, \$750,000

Phase II: ***Storage Efficient Data Mining for High-Speed Data Streams***

United States Air Force, Office of Scientific Research, STTR AF03-T011
Sonalysts, Inc., Waterford, CT, and the University of Connecticut

June 2003 - June 2004, \$100,000

Phase I: ***Storage Efficient Data Mining for High-Speed Data Streams***

United States Air Force, Office of Scientific Research, STTR AF03-T011
Sonalysts, Inc., Waterford, CT, and the University of Connecticut

January 2001 – December 2004, \$50,000 + 50,000 match from CSU System Office

Connecticut Distance Learning Consortium.

Program Director: *Developing an Online Master of Science in Data Mining*

January 2001. \$25,000 + \$25,000 match from CSU System Office

Connecticut Distance Learning Consortium.

Program Director: *Developing an Online Undergraduate Certificate in Data Mining*

Selected Intramural Grants

April 2006. AAUP University Research Grant. \$4000

Predicting Interesting Weather Patterns by Modeling Data Streams

April 2005. AAUP University Research Grant. \$3500

Meta-Modeling and Model Averaging in Classification Data Mining

April 2004. AAUP University Research Grant. \$4500

Mining Data Streams Using Non-Linear Probability-Based Decision Models

April 2002. AAUP University Research Grant. \$2985.

Estrogen Therapy and Cancer of the Endometrium: Which Lifestyle Factors Affect a Woman's Risk?

April 2001. AAUP University Research Grant. \$3970.

Which Factors Influence a Woman's Risk of Developing Breast Cancer When Taking Estrogen?

April 1998. AAUP University Research Grant, \$3199.

Does Longer Duration of Estrogen Exposure Increase a Woman's Risk of Developing Endometrial Cancer?

Service to the Department and the University

Coordinator of Data Mining @CCSU

Master of Science and Graduate Certificate programs in data mining.



From January 2001 to August 2016, I served as Coordinator of the data mining programs here at CCSU since I developed the programs in 2001. Here is an overview of some of the duties of this position.

- **Point of Contact.** I fielded email and phone inquiries from prospective students at the rate of more than one per day, seven days a week, 365 days a year. These represent responses to questions and queries coming from students and interested parties from around the world.
- **Program Advisement.** *I was the advisor for nearly all students in the programs,* thereby giving me about 70 active advisees and about 40 inactive advisees. Since most of these students are not within driving distance of CCSU, advisement is especially important if the students are not to feel isolated.
- **Acceptance Committee.** I served as the sole member of the acceptance committee (by agreement with the Dept of Mathematical Sciences) for both the Master of Science and the Graduate Certificate programs.
- **Data Mining Portal.** The coordinator is responsible for maintaining the CCSU data mining portal. This portal, located at www.ccsu.edu/datamining, is at or near the top of sites listed by *Google* for the search *data mining master's* (and several other searches) meaning that our site has the highest traffic and most important pointers to it, for people making such a search. This portal offers students an online nexus for accessing data mining resources from around the world, as well as information about the MS program, the Certificate program, course descriptions, faculty bios and photos, student comments, job listings, FAQ's, a contact form, and a data mining resources page, including links to data mining groups, data sets, research papers, commercial software. People search for data mining, and they find CCSU.
- **Program evaluation and assessment.** I wrote all NEASC reports and all program assessment reports.
- **Campus Representation.** The coordinator represents the program at scheduled campus functions, such as the Graduate School Open House.

- **Scholarship Committee.** I served as the chair of the selection committee for the Larose Scholarship for Excellence and Service in Data Mining. Each semester, students are nominated and selected to receive a total of \$2000 in scholarships, funded by my royalties donations and matched by the Fuller Endowment.
- **Curriculum.** In 2001, I designed and developed the original Master of Science program in data mining. In 2006, I developed the Graduate Certificate in Data Mining. In 2006 – 2007, I shepherded through the new Graduate Certificate in Data Mining, along with a substantial revision of the Master of Science in Data Mining. This entailed several meetings with CCSU curricular bodies, and four meetings beyond CCSU, including with the CSU Board of Trustees, and the Connecticut Department of Higher Education. In 2009 – 2010, at the suggestion of Dean Paulette Lemma, designed and shepherded through a substantial revision to the Master of Science and Graduate Certificate programs in Data Mining. These revisions have resulted in a substantial increase in the number of graduates.
- **Organizer of the Bi-Annual *Data Mining Thesis Presentation Day*.** Every semester, we come together to celebrate the latest graduates of the Master of Science in Data Mining, as we gather for the presentation of their theses.

Other Service

Extending Graduate Scholarships Across the Mathematical Sciences.

As a member of the Endowment Committee, I wrote the proposal to use funds from the Fuller Endowment to provide scholarships to new and current graduate students across all of the majors in the Department of Mathematical Sciences.

Helping Colleagues Publish Books

Since 2006 I have been the Series Editor for the *Wiley Series on Methods and Applications in Data Mining*. As Series Editor, I have helped two of our faculty members publish books with John Wiley and Sons, Inc. Dr. Roger Bilisoly published *Practical Text Mining* and Dr. Darius Dziuda published *Data Mining for Genomics and Proteomics*, both as part of the Series. Both texts are used for the relevant course curricula here at CCSU.

Graduate Studies Committee

Serving as Data Mining representative to the Graduate Studies Committee, 2016 – 2018.

Course Evaluation Questionnaire for Online Courses

I developed the course evaluation questionnaire for online courses for the Department of Mathematical Science. It was edited and approved by the Department in 2013.

Professional Activity

Book Series Editor: Wiley Series on Methods and Applications in Data Mining

John Wiley and Sons, Inc., Hoboken, NJ.

I canvass new book proposals and guide the authors toward completion of their projects.

Founder, *DataMiningConsultant.com*, LLC, a consulting small business company.

Founded 2006. Clients include the following:

- Microsoft, Inc.
- Oracle, Inc.
- *Forbes* Magazine
- *The Economist* Magazine
- Deloitte, Inc.
- The CIT Group
- KPMG International
- Booz-Allen-Hamilton, Inc.
- Computer Associates
- Sonalysts, Inc.
- The Hospital for Special Care, New Britain.

Analyst for published paper

I performed the data analysis required for the publication of the research paper, *An Analysis of Clinical Outcomes and Costs of a Long Term Acute Care Hospital*, by John Votto, Paul Scalise, Randall Barton, and Christine Vogel, *Journal of Medical Economics*, Volume 14, pages 141 – 146, January 2011.

Analyst for NIH Grant Application

I served as statistical analyst for the National Institutes of Health grant application, "PGRMC1 regulates bovine oocyte quality and is a potential biomarker of fertility", at the invitation of John J. Peluso, PhD, University of Connecticut Health Center, Principal Investigator.

Professional Membership

- American Statistical Association

Of Note

2002. CSU Certificate of Appreciation, Pioneer in Online Education.

2001. Project Kaleidoscope Faculty for the 21st Century.

2000. Teaching Excellence Award Honor Roll.

1998. Teaching Excellence Award Honor Roll.

1998. Founding Faculty Member, *OnlineCSU*.

1980. Phi Beta Kappa.