

## RESEARCH INTERESTS

Machine Learning (especially Active Learning, Experimental Design), Data Mining and Analysis, Information Retrieval and Theory.

## EDUCATION

### **Tokyo Institute of Technology**

*PhD Computer Science*

Degree Expected: Sep.2008

Member of Machine Learning Lab (Sugiyama)

Supervisor: Prof. Sugiyama

### **University of Massachusetts, Amherst**

*MSc Computer Science 2005*

Thesis: Fuzzy Ranking for Information Retrieval Systems

GPA 3.9/4.0

Supervisors: Prof. Alan, Prof. Kilmer

### **Brigham Young University**

*BSc Computer Science 1999*

GPA 3.7/4.0

## PUBLICATIONS

N. Rubens and M. Sugiyama. Influence information criterion for active learning in collaborative settings. *IEICE Transactions on Information and Systems*, 2008 [under review].

N. Rubens, V. Sheinman, T. Tokunaga, and M. Sugiyama. Order retrieval. In *Proceedings of the 3rd International Conference on Large-scale Knowledge Resources (LKR 2008)*, Lecture Notes in Artificial Intelligence (LNAI). Springer-Verlag, 2008.

N. Rubens and M. Sugiyama. Influence-based collaborative active learning. In *Proceedings of the 2007 ACM conference on Recommender systems (RecSys 2007)*. ACM, 2007. [pdf](#)

N. Rubens and M. Sugiyama. Explorative active learning for collaborative filtering. In *Proceedings of the Japanese Society for Artificial Intelligence 67th Meeting of Special Interest Group on Fundamental Problem in Artificial Intelligence*, 2007. [pdf](#)

N. Rubens and M. Sugiyama. Coping with active learning with model selection dilemma: Minimizing expected generalization error. In *Proceedings of 2006 Workshop on Information-Based Induction Sciences (IBIS 2006)*, 2006. [pdf](#) [bib](#)

N. Rubens. The application of fuzzy logic to the construction of the ranking function of information retrieval systems. *Computer Modelling and New Technologies*, 10(1):20-27, 2006. [pdf](#) [bib](#) [code](#)

M. Sugiyama, N. Rubens, and K.-R. Müller. Dataset Shift in Machine Learning, chapter A conditional

expectation approach to model selection and active learning under covariate shift. MIT Press, Cambridge, 2008.

M. Sugiyama and N. Rubens. A batch ensemble approach to active learning with model selection. Neural Networks, 2008 [conditionally accepted].

M. Sugiyama and N. Rubens. Active learning with model selection in linear regression. In SIAM International Conference on Data Mining (SDM 2008), 2008.

V. Sheinman, N. Rubens, and T. Tokunaga. Commonly perceived order within a category. In Proceeding of OntoLex Workshop at 6th International Semantic Web Conference (ISWC 07 ), 2007. [pdf](#)

V. Sheinman, N. Rubens, and T. Tokunaga. Word sequences for second language acquisition. MAPLL Workshop. Technical report of IEICE. Thought and language, 107(138), 2007.

N. Rubens. DNA sequence alignment with the use of reinforcement learning. Technical report, 2004.

N. Rubens. Application of fuzzy logic to the refinement of an answer set of information retrieval systems. Technical report, 2004.

N. Rubens. Detecting network intrusion with the use of belief networks. Technical report, 2003.

## FUNDING AWARDS

**Research Fellowship** Apr.2007 - Aug.2008

*Japanese Government* (MEXT (Monbukagasho) / Tokyo Institute of Technology) \$50,000

**Research Fellowship** Oct.2005 - Apr.2007

*Japanese Government* (MEXT (Monbukagasho) / Tokyo Institute of Technology) \$50,000

**Research Fellowship** Oct.2005 - Aug.2008

*Japanese Government* (MEXT (Monbukagasho) / JAIST) \$100,000

Declined due to conflicting offer

**Travel / Conference Grant** Oct.2006

Information-Based Induction Sciences (IBIS)

## PROFESSIONAL EXPERIENCE

**Tokyo Institute of Technology / Machine Learning Lab (Sugiyama)** Oct.2005 - present

*Research Fellow*

Artificial Intelligence/Machine Learning: development of fundamental theories and practical algorithms

**Fincross** May.2003 - Aug.2005

*Senior Software Engineer (Consultant)*

Application of Artificial Intelligence and Information Retrieval to Financial Analysis

**CTS** Dec.2001-May.2003

*Software Engineer*

Analysis and processing of telecommunication data

**NIC** Jan.2000-Sep.2001

*Software Engineer*

Development of server side applications with high reliability/scalability requirements

**ETS** May.1999-Jan.2000

*Software Engineer*

Analysis and processing of standardized assessment data

**Brigham Young University / Nissan Sep.1998-May.1999**

*Research Assistant*

R&D of neuro-fuzzy controller for Accumulative Cruise Control

**Corel Apr.1998-Sep.1998**

*Software Engineer*

Primarily responsible for maintaining and improving Word Perfect Expert module - combines the features from QuickTasks, templates, Help, and coaches user to help accomplish work more efficiently

**Frank Phillips College Aug.1996-May.1997**

*Lab Supervisor*

Developing and maintaining IT infrastructure

Training/Supervising teaching assistants

## PROJECTS

### **E-Business**

*Software Engineer primarily responsible for (NIC):*

- DMV Renewal (revenue: \$10 Million / year)
- Payment Processing Framework (revenue: \$20 Million / year)
- Shopping Cart (used by several US state agencies)

### **Data Analysis / Processing**

- financial data (Fincross)
- standardized assessment data (ETS)
- telecommunication data (CTS)

## SKILLS

- Numerical Analysis: Matlab, R, S, SAS, SQL
- Programming Languages: Java, C#, C++

## CITIZENSHIP

- Canada, Japan (permanent resident)
- Working visa is not required for: Canada, United States, Japan and certain countries in European Union

## REFERENCES

Available upon request.

---