



# Siavash Kazemian

1208-8 McKee Ave.  
Toronto, ON. M2N 7E5  
T 416-888-4481  
kazemian@cs.toronto.edu  
[www.cs.toronto.edu/~kazemian](http://www.cs.toronto.edu/~kazemian)

## Education

University of Toronto, Department of Computer Science - M. Sc. 2008

University of Toronto, Department of Electrical and Computer Engineering - B. A. Sc. 2005

## Skills

- Programming: Python, Fortran, CSH scripting, ActionScript and Adobe Flex, Glade, Tcl/Tk, ASP.NET, C/C++, C#
- Sound understanding of numerical precision issues in software
- Statistical/Linguistic Models: Hidden Markov Models (HMM), Neural Networks, Deep Belief Nets, Restricted Boltzman Machines, Singular Value Decomposition (SVD), Decision Trees, AdaBoost, TextTiling
- Leadership, Written and oral communication skills, Business skills

## Relevant Projects

### Speech Retrieval System - January 2008 to August 2008

- Designed and implemented a system that allows a user to search and browse over 100 hours of speech data. All of the coding was done in Python and Csh scripts.
- Compared the performance of the most cutting edge indexing schemes for speech data: Word Lattice, Position Specific Posterior Lattice (PSPL), and Confusion Networks.
- Employed a simple yet powerful and efficient technique for this task called Set-of-Words. This method's retrieval performance produced an F-measure of 82.7% compared with state of the art: PSPL's performance of 84.5% while performing the search task 5 times faster.
- Designed a web-based Flash interface for this system using Adobe Flex, allowing the user to search and browse the corpus.

### Text Segmentation and Summarization - Feb 2006

- Wrote a program to segment a text document according to topics covered using the TextTiling algorithm. Then wrote a summarization module employing Singular Value Decomposition (SVD) that used paragraph extraction to create an informative summary.
- Achieved great performance both in topic segmentation and summarization. For topic segmentation, I achieved an F-measure of 72% compared with the baseline system's performance: F-measure of 30%. Summarization was evaluated manually and achieved an accuracy of 70%.



### **Genre Classification and End-of-Sentence Detection - March 2006**

- Implemented a heuristic algorithm to find the end of sentence boundaries. This implementation improved the performance of baseline method that picks all "." as end-of-sentence boundaries by 12%.
- Wrote a program to identify genre of a documents. Processed the text documents to collect relevant statistics. Then used the collected statistics to train a decision tree for the classification task.

## **Employed Positions**

### **Teaching Assistant for Natural Language Processing, University of Toronto - 2008**

- Lectured the class about Text Segmentation and Text Summarization
- Was in charge of an assignment that employed Text Segmentation and Text Summarization.

### **Quality Verification Specialist, Redknee Inc. Toronto, ON - 2006 to 2007**

- Performed major functionality and performance verification of our company's most complex product: Customer Rating Management (CRM).
- Designed Test cases for new features added to our product.
- Took initiative to make training material for our product. The material is now used by new hires and sales and management staff.

### **Engineering Trainee, AN GROUP A/S, Lyngø, Denmark - 2006**

- Worked with a senior engineer on the specifications of a reporting system for ECOLAB: an industrial cleaning company. The product would be installed in 10-15 locations across Denmark.
- Developed the reporting system in ASP.NET using C#. The system would allow the end user to visualize and export the data in various formats. The data visualization functionality is a 10,000 USD asset that AN GROUP A/S now owns and has used in other projects.

## **Awards and Leadership Appointments**

- Isabel A. King Scholarship for Leadership in community services. Sept 2001
- IEEE Canadian Foundation Scholarship (2003)
- Professional Development Director, IEEE Student Branch, University of Toronto; 2003 – 05

## **Publications**

- S. Kazemian, F. Rudzicz, G. Penn, C. Munteanu. "A Critical Assessment of Spoken Utterance Retrieval through Approximate Lattice Representations" Proceeding of ACM Multimedia Information Retrieval 2008
- X. Xhu, S. Kazemian, G. Penn. "Identifying Salient Utterances of Online Spoken Documents Using Descriptive Hypertext". Submitted to IEEE Spoken Language Understanding Workshop 2008.

## **Reference**

Available upon request