

CEM AKKAYA

210 S. Bouquet St.,
Rm. 5422 Sennott Square Building
Pittsburgh, PA 15260

<http://www.cs.pitt.edu/~cem>
412-736-7724
cem@cs.pitt.edu

RESEARCH INTEREST

My research interests consist of **Natural Language Processing** and **Machine Learning**. Primarily, I work on the application of Word Sense Disambiguation to Subjectivity/Sentiment Analysis. Machine Learning is an integral part of my research. I am particularly interested in active learning and semi-supervised learning methods and also crowdsourcing.

EDUCATION

University of Pittsburgh, Pittsburgh, PA 2006-present
Ph.D. in Intelligent Systems Program (Expected Graduation: December 2012)
Advisor: Dr. Janyce Wiebe

Vienna University of Technology (TU Wien), Vienna, Austria 2000–2005
B.S. and M.S. in Informatics
Advisor: Dr. Silvia Miksch, Dr. Katharina Kaiser

RESEARCH EXPERIENCES

Natural Language Processing Group, University of Pittsburgh 2007-present
Working with Janyce Wiebe on Subjectivity and Sentiment Analysis:
Application of Word Sense Disambiguation to Subjectivity and Sentiment Analysis
Design of an expressive lexicon representation

Information Engineering Group, TU Wien 2004-2005
Worked with Katharina Kaiser and Silvia Miksch on Information Extraction from
clinical practice guidelines

PUBLICATIONS

Cem Akkaya, Janyce Wiebe, and Rada Mihalcea (2012). Utilizing Semantic Composition in Distributional Semantic Models for Word Sense Discrimination and Word Sense Disambiguation. Sixth IEEE International Conference on Semantic Computing (IEEE ICSC 2012)

Cem Akkaya, Janyce Wiebe, Alexander Conrad and Rada Mihalcea (2011). Improving the Impact of Subjectivity Word Sense Disambiguation on Contextual Opinion Analysis. *Conference on Computational Natural Language Learning (CoNLL 2011)*.

Cem Akkaya, Alexander Conrad, Janyce Wiebe, and Rada Mihalcea (2010). *Amazon Mechanical Turk for Subjectivity Word Sense Disambiguation*. NAACL-HLT 2010 Workshop on Creating Speech and Language Data With Amazon's Mechanical Turk

Cem Akkaya, Janyce Wiebe and Rada Mihalcea (2009). *Subjectivity Word Sense Disambiguation*. Conference on Empirical Methods in Natural Language Processing (EMNLP 2009).

Yaw Gyamfi, Janyce Wiebe, Rada Mihalcea, and Cem Akkaya (2009). *Integrating Knowledge for Subjectivity Sense Labeling*. Joint Conference of the North American Chapter of the Association for Computational Linguistics and the Human Language Technologies Conference (HLT-NAACL 2009).

Katharina Kaiser, Cem Akkaya, Silvia Miksch (2007). *How can Information Extraction ease formalizing treatment processes in clinical practice guidelines? A method and its evaluation*. *Artificial Intelligence in Medicine*, 39(2):151-163, 2007.

Katharina Kaiser, Cem Akkaya, Silvia Miksch (2005). *Gaining Process Information from Clinical Practice Guidelines Using Information Extraction*. 10th Conference on Artificial Intelligence in Medicine (AIME 2005).

INDUSTRIAL EMPLOYMENTS

Software Developer, Denizbank Digital Kusak Bankaciligi 2005-2006

Developer for Intertech Financial Advanced Computing Environment (inter-F.A.C.E.), a banking solution for DenizBank. I worked on all three tiers of the project doing Javascript and HTML programming on the presentation tier, C# programming on the application tier and SQL programming on the database tier.

Summer Intern, Yahoo! Inc. 2010

I worked under the supervision of Dr. Yookyung Kim in the Infinite Browse Team. We applied Machine Learning techniques to improve parts of the existing logic.

ACADEMIC EMPLOYMENTS

Graduate Student Researcher 2007-present

Intelligent Systems Program, University of Pittsburgh

Teaching Assistant for CS7: Introduction to Programming 2008

Department of Computer Science, University of Pittsburgh

ATTENDED PROJECTS

Design and Implementation of the treasury component for inter-F.A.C.E., a banking solution for DenizBank. Inter-F.A.C.E is based on a three-tier architecture model. I worked on all three tiers of the project.

Used Technologies and Languages: Javascript, Ajax, HTML, C#, SQL, Microsoft SQL Server

Application of machine learning methods to entity recommendation from news articles for Yahoo's "Infinite Browse". The machine learning component I designed went to production.

Used Technologies and Languages : Java, Weka machine learning software

Design and Implementation of a new version of OpinionFinder, a widely used opinion analysis system.

Used Technologies and Languages: Java, Weka machine learning software

Design and Implementation of a GUI based information extraction system. The system was used to extract treatment options from clinical practice guidelines.

Used Technologies and Languages: Java and Swing Java GUI toolkit

Design of an expressive lexicon representation to store clue patterns for opinion analysis. The patterns are very flexible and can be defined in terms of n-grams, gappy n-grams and syntactic relations.

Used Technologies and Languages: XML, XML Schema

Implementation of a tool to match clue patterns from the designed lexicon in text documents.

Used Technologies and Languages: Python, Java

Implementation of a tool for Constrained Clustering

Used Technologies and Languages: MATLAB

Maintaining Intelligent Systems Program webpage

Used Technologies and Languages: Drupal Content Management System, HTML

Implementation of a web page for class registration as part of a class project

Used Technologies and Languages : Java, JavaServer Pages, SQL, MySQL Server, Apache Tomcat Server

Implementation of a reasoning engine based on first-order predicate logic.

Used Technologies and Languages : LISP

HONORS & AWARDS

| | |
|---|-----------|
| Arts and Sciences Graduate Fellowship, University of Pittsburgh | 2006-2007 |
| Andrew Mellon Pre-Doctoral Fellowship | 2010-2011 |

SERVICES AND ORGANIZATIONS

| | |
|--|--------------|
| Reviewer for EMNLP 2012 | 2012 |
| Reviewer for ACL 2012 | 2012 |
| Program Committee for NAACL 2012 | 2012 |
| Reviewer for NAACL 2010 | 2010 |
| Member of the Natural Language Processing Group, University of Pittsburgh | 2007-present |
| Webmaster, Intelligent Systems Program, University of Pittsburgh | 2008-2009 |
| Member of the Turkish & American Student Association at University of Pittsburgh | 2006-present |

LANGUAGE PROFICIENCY

English: Fluent

German: Fluent

Turkish: Native

REFERENCES

Janyce Wiebe (University of Pittsburgh – Professor)

Email : wiebe@cs.pitt.edu

Phone : 412-624-9590

Yookyung Kim (Yahoo! Inc. – Linguistics Lead)

Email : ykim@yahoo-inc.com

Phone : 650-269-3818

Milos Hauskrecht (University of Pittsburgh – Professor)

Email : milos@cs.pitt.edu

Phone: 412-624-8845