

Department of Computer Science
University of Houston

DISTINGUISHED LECTURER SEMINAR 2012

WHEN: MONDAY, SEPTEMBER 10, 2012
WHERE: PGH 232
TIME: 11:00 AM

SPEAKER: Dr. Pushpak Bhattacharyya, IIT Bombay

Host: Dr. Ioannis Kakadiaris

TITLE: Sentiment Analysis in New Terrain: New Domain, New Features, New Language and New Tasks

Abstract: Sentiment analysis is an exciting new field of research in Artificial Intelligence combining Natural Language Processing, Machine Learning and Psychology. Since 2000, due to the proliferation of huge amounts of opinions in electronic form on the web, on social networks and on blogs, automatic means of polarity (positive, negative and neutral) detection in texts flourished in leaps and bounds. Individual and organizations with public interface can no longer afford to be oblivious of sentiments expressed about them in electronic form.

In the present talk we describe our research on sentiment analysis that attempts to advance the state of the art by tackling new text form (Tweets), a new task (Thwarting), new language (Indian languages) and new features (word senses instead of words). Tweets are noisy texts; they need text cleaning and normalization. Exploitation of discourse relations like 'although', 'still', 'but' and so on improve the accuracy considerably. Languages differ in terms of annotated resources. However, projecting the parameters learnt from one language to another can ameliorate the problem of resource scarcity. We show the use of these ideas for sentiment analysis of Indian languages. We demonstrate that working with senses instead of just words is not only advantageous from the point of view of multilinguality, but also for accuracy. Finally, problems like thwarting (reversal of polarity just by a single statement at a critical place in the text) and sarcasm are hard problems. We present some of our findings in these tasks too.

The presentation is based on our recent publications in machine learning and NLP fora like EMNLP, ACL, ECML, CIKM and so on. The students involved are Balmurali (PhD), Subhbrata (Masters- graduated), Aditya (masters- graduated) and Akshat (masters- graduated).

BIO: Dr. Pushpak Bhattacharyya is a Professor of Computer Science and Engineering at IIT Bombay. He received his B.Tech from IIT Kharagpur, M.Tech from IIT Kanpur and PhD from IIT Bombay. He has held visiting positions at MIT, Cambridge, USA, Stanford University, USA and University Joseph Fourier, Grenoble, France. Dr. Bhattacharyya's research interests include Natural Language Processing, Machine Learning and Cross Lingual Search. He has had more than 130 publications in top conferences and journals and has served as program chair, area chair, workshop chair and PC member of top fora like ACL, COLING, LREC, SIGIR, CIKM, NAACL, GWC and others. He has guided 12 PhDs and over 150 masters and undergraduate students in their thesis work. Dr. Bhattacharyya plays a leading role in India's large scale projects on Machine Translation, Cross Lingual Search, and Wordnet and Dictionary Development. Dr. Bhattacharyya received a number of prestigious awards including IBM Innovation Award, Yahoo Faculty Award, United Nations Research Grant, Microsoft Research Grant, IIT Bombay's Patwardhan Award for Technology Development and Ministry of IT and Digital India Foundation's Manthan Award. Recently he has been appointed Associate Editor of the prestigious journal, ACM Transactions on Asian Language Information Processing.