

COMPUTER SCIENCE
SEMINAR

*Opinion Mining for the Internet: Models, Algorithms and
Predictive Analytics*

Arjun Mukherjee
University of Illinois at Chicago

Abstract: The massive amounts of user generated content in social media offers new forms of actionable intelligence. Public sentiments in debates, blogs, and news comments are crucial to governmental agencies for passing new bills/policies, gauging social unrest, predicting elections, and socio-economic indices. The goal of my research is to build robust statistical models for opinion mining with applications to marketing, social, and behavioral sciences. To achieve this goal, a number of research challenges need to be addressed. The first challenge is fine-grained information extraction which can capture diverse types of opinions (e.g., agreement/disagreement; contention/controversy, etc.) and various other latent sentiments expressed in social conversations and discussions. The state-of-the-art machinery (e.g., topic modeling) falls short for such a task. I develop several novel knowledge induced sentiment topic models which respect notions of human semantics. The second challenge is that social sentiments are inherently dynamic and change over time. To leverage the sentiments over time for predictive analytics (e.g., predicting financial markets), I develop Bayesian nonparametric topic based sentiment time-series and vector autoregression models. The third challenge is to filter deceptive opinion spam/fraud. It is estimated that 15-20

In this talk, I will present novel statistical models for sentiment analysis and talk about two key frameworks: (1) Semi-supervised graphical models for mining fine-grained opinions in social conversations, and (2) Bayesian nonparametrics, sentiment time-series, and vector autoregression models for stock market prediction.

In the later part of the talk, I will discuss the problem of opinion spam and throw light on some techniques for filtering opinion spam. The focus will be on modeling collusion and combating group spam in e-Commerce reviews. The talk will conclude with a discussion about my ongoing research and future research vision in opinion contagions, forecasting socio-economic indices, and healthcare.

Thursday, March 20, 2014, 4:00 pm
Mathematics and Science Center: W302

MATHEMATICS AND COMPUTER SCIENCE
EMORY UNIVERSITY