

Linguistic Models of Deceptive Opinion Spam

Myle Ott

Facebook

myleott@gmail.com

Abstract of the talk

Consumers increasingly inform their purchase decisions with opinions and other information found on the Web. Unfortunately, the ease of posting content online, potentially anonymously, combined with the public's trust and growing reliance on this content, creates opportunities and incentives for abuse. This is especially worrisome in the case of online reviews of products and services, where businesses may feel pressure to post deceptive opinion spam---fictitious reviews disguised to look like authentic customer reviews.

In recent years, several approaches have been proposed to identify deceptive opinion spam based on linguistic cues in a review's text. In this talk I will summarize a few of these approaches. I will additionally discuss some of the challenges researchers face when studying this problem, including the difficulty of obtaining labeled data, uncertainties surrounding the prevalence of deception, and how linguistic cues to deceptive opinion spam vary with the text's sentiment (e.g., 5-star vs 1- and 2-star reviews), domain (e.g., hotel vs. restaurant reviews) and the domain expertise of the author (e.g., crowdsourced vs. employee-written deceptive opinion spam).