

# Large Scale Learning at Twitter

Aleksander Kolcz

Twitter, USA  
alek@ir.iit.edu

Twitter represents a large complex network of users with diverse and continuously evolving interests. Discussions and interactions range from very small to very large groups of people and most of them occur in the public. Interests are both long and short term and are expressed by the content generated by the users as well as via the Twitter follow graph, i.e. who is following whose content.

Understanding user interests is crucial to providing good Twitter experience by helping users to connect to others, find relevant information and interesting information sources.

The manner in which information is spread over the network and communication attempts are made can also help in identifying spammers and other service abuses.

Understanding users and their preferences is also a very challenging problem due to the very large volume information, the fast rate of change and the short nature of the tweets. Large scale machine learning as well as graph and text mining have been helping us to tackle these problems and create new opportunities to better understand our users. In the talk I will describe a number of challenging modeling problems addressed by the Twitter team as well as our approach to creating frameworks and infrastructure to make learning at scale possible.