Gender, styles, and social networks in Twitter
Tyler Schnoebelen, David Bamman and Jacob Eisenstein
Stanford University, Carnegie Mellon University, Georgia Tech

Session 12: Computer-Mediated Communication

The use of social media for informal communication enables large-scale statistical modeling of language style and author properties like gender, age, race, and geographical location. Strong correlations between language and such categories enable predictive models that are disarmingly accurate. But this leads to an oversimplified and misleading picture of how language conveys personal identity. We present evidence from large-scale social media data that supports an alternative view. Rather than two styles—male and female—we find multiple styles, many of which are gendered but differently so. We use "style" to refer to combinations of linguistic resources; we call a group of authors that use a particular style a "cluster". Each author has their own social network, which may or may not be part of the same cluster—we define an author's social network as the people they repeatedly send and receive direct messages from.

We begin with a predictive model that achieves state-of-the-art gender prediction accuracy of 88.9% on a corpus of 14,464 Twitter users. The predictive power comes from words with the strongest gender associations: for women, many of these words are non-standard spellings or emoticons; for men, many of the words are named entities in the sports and technology domains. A clustering analysis tells a more nuanced story. We cluster authors by their text words alone and find that many clusters have strong gender associations—some of these clusters bear out the high-level gender difference of non-standard spellings versus named entities, but others directly contradict it. For example, in the author-based analysis, a cluster of authors that is 72% male heavily favors alternative spellings, a pattern our gender-only classifier learned to be a hallmark of female language. The statistics make it clear that the clusters enact gender, but they do so in a way which is at odds with the population-level statistics. By dividing authors based only on gender, we would not be aware that these gendered styles were even possible.

Next, we consider the social network between the authors in our corpus. The network displays strong homophily: 63% of the connections are between same-gender individuals. If social network connections and text were each independently driven by the basic category of gender, making gender predictions with both should be better than using just one. But in fact, individuals who use linguistic resources from the other gender (or who do not use linguistic resources from their own gender) have consistently denser social network connections to the other gender. A static binary model of gender is inadequate in describing linguistic resources and social network connections.

Finally, we combine the social network and clustering analysis. In general, the gender composition of the social networks of the members of each cluster tracks the gender composition of the cluster itself. For example, women in the sports-related clusters have far more male friends than average—though they still have fewer male friends than the male members of the cluster. Rather than revealing essential categories, these styles reflect an interplay of authors, audiences, and topics.