Computational Social Science

Natural Language Processing: Jordan Boyd-Graber
University of Maryland

INTRODUCTION
What is Social Science?

<table>
<thead>
<tr>
<th>Study of Society and Relationships Among Individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Economics</td>
</tr>
<tr>
<td>- Sociology</td>
</tr>
<tr>
<td>- Psychology</td>
</tr>
<tr>
<td>- Political Science</td>
</tr>
</tbody>
</table>

Connection to NLP?

Many relationships happen through or are described by text!
What is Social Science?

Study of Society and Relationships Among Individuals

- Economics
- Sociology
- Psychology
- Political Science

Connection to NLP? Many relationships happen through or are described by text!
## Alternatives / Complements to Traditional Social Science

- **Polls**
- **Surveys**
- **Voting patterns**
Dataset: Social Media

- People share more readily
- Self-censorship on official channels
- Faster feedback
Dataset: Social Media

- People share more readily
- Self-censorship on official channels
- Faster feedback
Demography

Bias in social media (and associated devices)

But also a problem for traditional instruments
**Dataset: Social Networks**

- Predict links
- Predict when friendships end (and why)
More importantly . . .

- Suicide risk
- Unhealthy habits
- Bullying
More importantly . . .

- Suicide risk
- Unhealthy habits
- Bullying

Computational social science emphasizes interpretability, broader impacts, and captures what about language is unique to humans.
Tool: Topic Models

- Open-ended responses
- Questions you “forgot to ask”
- Help with interpretation
- Good for small amounts of data

---

*From the Wed., Dec. 25, 1861, issue*

—Runaway from the subscriber, on the 3d inst., my slave woman PARTHENIA. Had on a dark brown and white calico dress. She is of a ginger-bread color; medium size; the right fore-finger shortened and crooked, from a whitlow. I think she is harbored somewhere in or near Duvall’s addition. For her delivery to me I will pay $10.

G. W. H. TYLER. de 6—ts
Tool: Topic Models

- Open-ended responses
- Questions you “forgot to ask”
- Help with interpretation
- Good for small amounts of data
- Tension: prediction or interpretation
Qualitative Analysis (Grounded Theory)

Grounded Theory
Inductive coding of qualitative data to reach understanding

- low-quality models/theories and refining
- topics are “grounded” in specific word tokens
- Baumer et al. [2017]: compare themes
- model gives topic-specific “reading list”
- efficient, insulation from human biases.
Focus of today: Key Problems

- Sentiment Analysis
- Ideology Detection