Social Media and Political Participation

Lab 4

pablo.barbera@nyu.edu

January 10, 2014

Today

- Facebook: what is it? Main features.
- Introduction to the Facebook API
- Collecting Facebook data using the API
- Quantitative analysis of Facebook data
- In-class exercise: capture and analyze your own Facebook data

Facebook

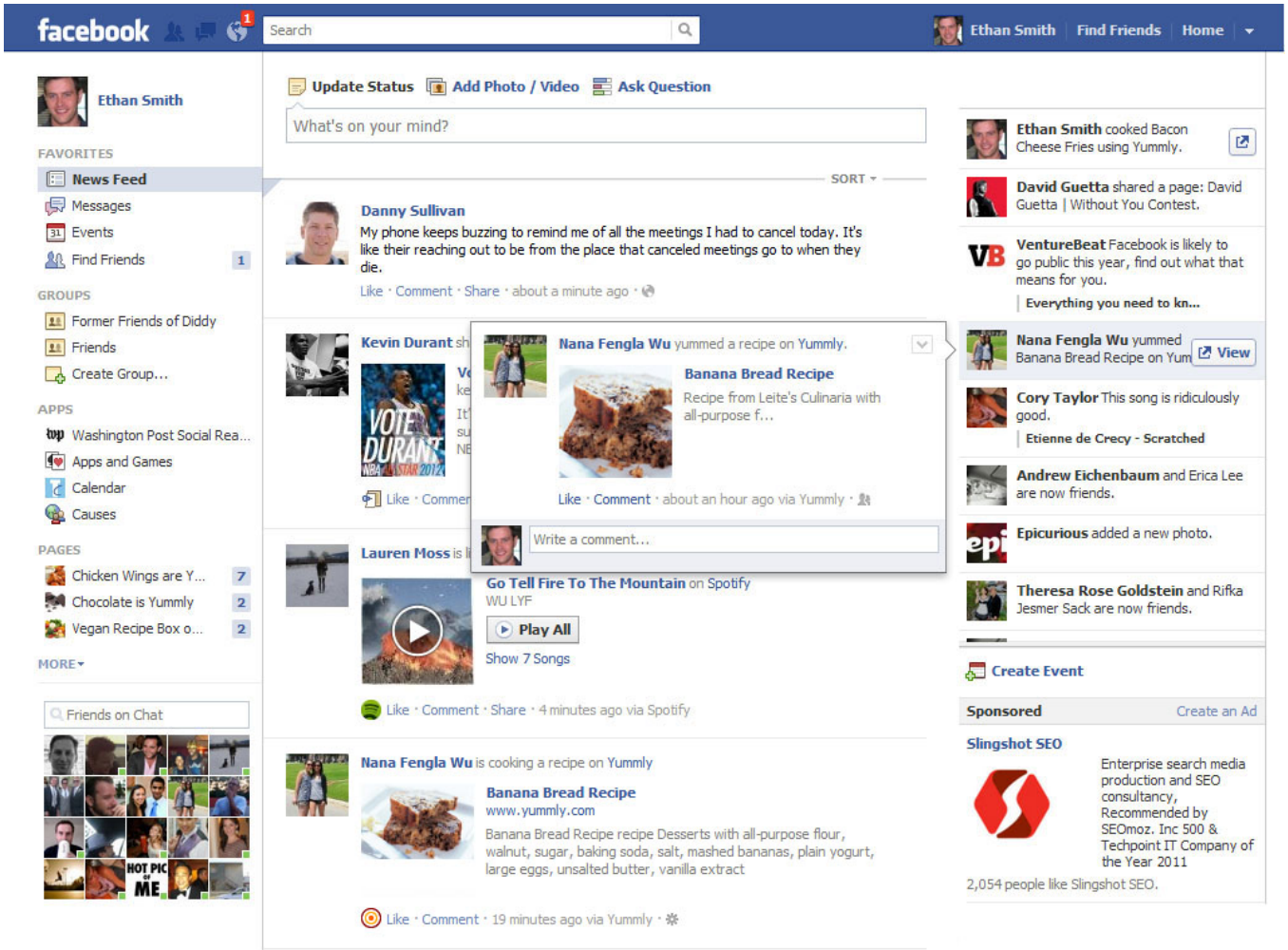
Facebook's numbers

- 1,3+ billion monthly active users
- 10 billion messages are sent everyday
- 1.3 trillion "like" activities since 2009
- 71% of online U.S. adults use Facebook
- 87% of young adults in US (18-29) report using Facebook actively
- 47% of Facebook users get news through this platform
- 99% of Members of U.S. Congress have a Facebook account

Facebook's main features

Barack Obama's Facebook Timeline

Facebook's main features



A random user's News Feed

Facebook's main features

Three main features

1. Timeline: profile with photos, lists of interests, contact information, personal background, etc.
2. News Feed: shows status updates by users and profile changes, which can be "liked", "shared" or "commented"
3. Private messages: allows users to communicate privately

Other features:

pages Public Facebook profiles for political figures, companies, celebrities...

like Positive feedback on a post, page, or link

share Re-publication of another user's content

The offline effects of Facebook

Three defining characteristics of Facebook

1. Most content is private
2. "Friends" are usually actual friends
3. Social metrics for every post

Consequences:

- Facilitates organization of collective action
  - → "Social Media and the Decision to Participate in Political Protest: Observations From Tahrir Square", by Tufekci and Wilson, *Journal of Communication* (2012)
- Channels social influence on political behavior
  - → "A 61-million-person experiment in social influence and political mobilization", by Bond et al, *Nature* (2012).
- Social cues that affect behavior
  - → "Endorsements Trump Partisan Source Affiliation when Selecting News Online", by Messing and Westwood, *Communication Research* (2014).

Learning from Facebook networks

"Private traits and attributes are predictable from digital records of human behavior", by Kosinski, Stillwell, and Graepel, *PNAS* (2013)

ABSTRACT: We show that easily accessible digital records of behavior, Facebook Likes, can be used to automatically and accurately predict a range of highly sensitive personal attributes including: sexual orientation, ethnicity, religious and political views, personality traits, intelligence, happiness, use of addictive substances, parental separation, age, and gender.

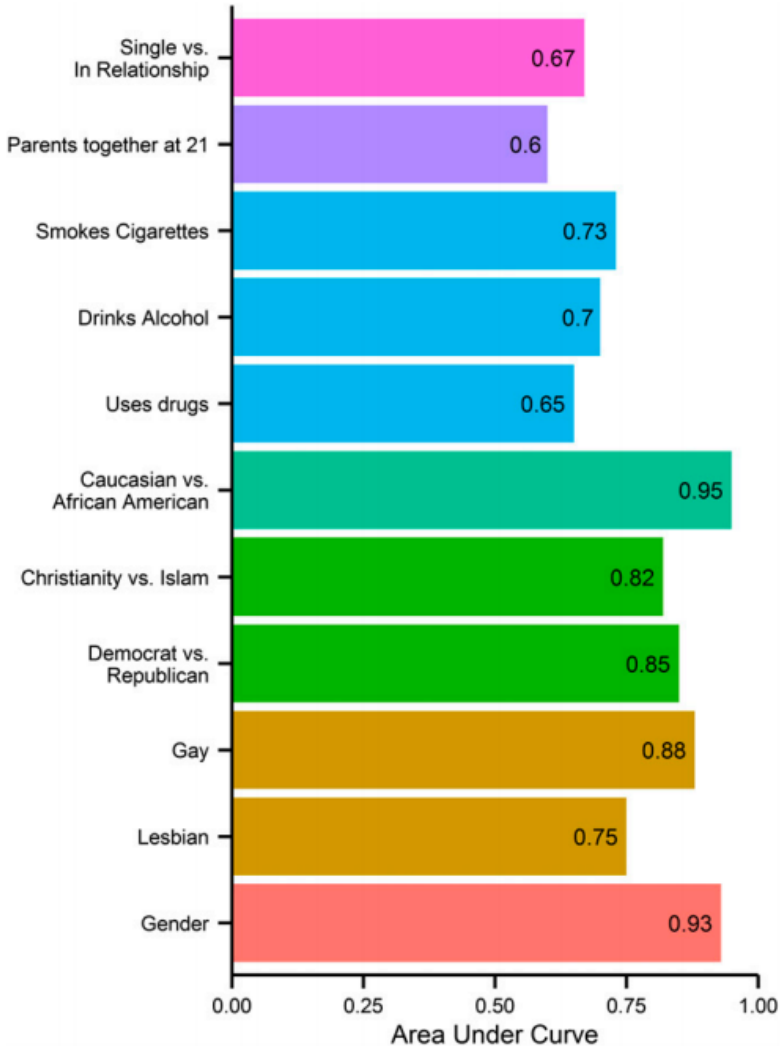


Fig. 2. Prediction accuracy of classification for dichotomous/dichotomized attributes expressed by the AUC.

Facebook API

Facebook API

API = *Application Programming Interface*

Facebook gives researchers access to two different types of data:

1. Data from Facebook pages (posts, likes, comments)
2. User's personal data (profile, checkins, likes...)

`Rfacebook` package gives access to both with the following functions:

1. `getPage` and `getPost`
2. `getUsers`, `getCheckins`, `getLikes`

Collecting Facebook Data

The R script `lab4_collecting_facebook_data` shows how to:

- Install R package to download Facebook data
- Use OAuth to authenticate
- Display your profile information
- Capture data from a Facebook page

Collecting Facebook Data

The R script `lab4_analyzing_facebook_data` shows how to:

- Load R package to download Facebook data
- Use OAuth to authenticate
- Analyze metrics of a Facebook page over time
- Prepare a wordcloud that summarizes comments on a page

In-class exercise

In-class exercise: collecting and analyzing Facebook data

Create your own R script (with comments) that:

1. Downloads the most recent 1000 posts on a Facebook page of a celebrity or politician.
2. Runs different commands to answer the following questions:
    1. Which of these 1000 posts received the most likes, comments, shares? Are these three different? If so, why?
    2. Create a plot that shows the evolution in the number of likes on posts over time. Is the popularity of this page growing?
    3. Choose a post and download all the likes on that post. (If they are too many, choose the first 500). Then, download the user information and look at the most common first names. What's the gender distribution?
    4. Download also the comments on that post (or the first 1000). Then, do a wordcloud of the most common words. What do you learn?

And send it to me via email (`pablo.barbera@nyu.edu`)