computational social media

lecture 6: watching

daniel gatica-perez

10.05.2019
announcements

reading #6 will be presented today

T. Gebru, J. Krause, Y. Wang, D. Chen, J. Deng, E. Lieberman Aiden, and L. Fei-Fei, Using deep learning and Google Street View to estimate the demographic makeup of neighborhoods across the United States, PNAS, 2017

Presenter: A. Sherstiuk
Discussant: R. Petitpierre
Scribe: J. Wang
this lecture

1. the rise of online video
2. conceptualizing YouTube
3. conversational social video
   uses of conversational video
   behavioral analysis & personality impressions
1.

the rise of online video
“Everything exists to end up in a book”
Stéphane Mallarmé, 1842-1998

“Everything exists to end up in a photograph”

“Everything exists to end up in YouTube”
Online Creators: Posting original photos and videos, 2012-2013

Among internet users, the % who post photos and videos and the % who post either

<table>
<thead>
<tr>
<th></th>
<th>2012</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post photos</td>
<td>45%</td>
<td>52%</td>
</tr>
<tr>
<td>Post videos</td>
<td>18%</td>
<td>25%</td>
</tr>
<tr>
<td>Creators</td>
<td>46%</td>
<td>54%</td>
</tr>
</tbody>
</table>

Online Curators: Sharing images and video from elsewhere on the web, 2012-2013

Among internet users, the % who have reposted images and videos they found elsewhere online and the % who repost both (curators)

<table>
<thead>
<tr>
<th></th>
<th>2012</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share others' images</td>
<td>35%</td>
<td>42%</td>
</tr>
<tr>
<td>Share others' videos</td>
<td>25%</td>
<td>36%</td>
</tr>
<tr>
<td>Curators</td>
<td>41%</td>
<td>47%</td>
</tr>
</tbody>
</table>

Source: Pew Research October Omnibus Survey, October 3-6, 2013. N=852 Internet users ages 18+. Interviews were conducted in English on landline and cell phones. The margin of error for results based on internet users is +/- 4.0 percentage points.

watch:
http://www.pewinternet.org/2013/10/10/video-the-rise-of-online-video/

http://www.pewinternet.org/2013/10/28/photo-and-video-sharing-grow-online/
http://www.pewinternet.org/2013/10/28/additional-analysis/
2.
conceptualizing YouTube
YouTube statistics
(may 2018, some numbers from earlier years)

founded in February 2005
most popular online video platform

1B unique users
1B hours of video watched each day
300h of video uploaded per minute (in 2015)
   (100h in 2014)

more 18-34 & 18-49 year-olds than any cable network in US
mobile: over 50% of global views

https://www.youtube.com/yt/about/press/
synchronous
few-to-many
passive
centralized

asynchronous
many-to-many
interactive
decentralized
Survey: YouTube Stars More Popular Than Mainstream Celebs Among U.S. Teens
View #1: YouTube is a weird place
YouTube and the silliness of crowds

rudeness of crowds

absurdity

J. Stossel, ABC 20/20:
"Do you like watching kids doing stupid and reckless things? Beauty queens falling down? Or a thousand prisoners dancing to the music of Thriller? It’s all in YouTube"

Burgess and Green:
"Rather than video about nothing, this could be situated in the much longer history of vernacular creativity – the wide range of everyday creative practices ... practiced outside the cultural value systems of high culture or commercial practice"

J. Burgess and J. Green,YouTube. Online Video and Participatory Culture, Polity, 2009
View #2: YouTube is a popular place
YouTube and popular videos

- analysis of video popularity distributions (power-law with truncated tails)
- large-scale analysis ($10^6$ videos)
- no content was analyzed

![Graph showing video popularity distribution of YouTube: power-law in waist, sharp decay in tail](image)

View #3: YouTube is a place for pirates
YouTube and copyright infringement

- near-duplicate (ND) detection: YouTube key business feature
- YouTube partners with TV and movie companies to remove unauthorized copies
- NDs are seen as redundant and deemed to be eliminated
- NDs are not always uploaded with bad intent

How Content ID works (launched 2007)

Copyright owners can use a system called Content ID to easily identify and manage their content on YouTube.

“July 2015: 8,000+ partners use Content ID, including major network broadcasters, movie studios and record labels. They have claimed over 400 million videos, helping them control their content and make money on videos containing copyrighted material.”

https://support.google.com/youtube/answer/2797370?hl=en
https://www.youtube.com/yt/about/press/
video
View #4: YouTube is a place for marketing
Be seen where everyone is watching

With TrueView ads, reach potential customers when they watch or search for videos on YouTube – and only pay when they show interest.

http://www.youtube.com/yt/advertise/index.html
View #5: YouTube is a place for expression
conversational social video: provide a rich communication experience
3. conversational social video
La piel va perdiendo hidratación y por tanto elasticidad a lo largo de los años. Además, las pieles maduras ya cuentan con arrugas y signos de la edad que deben tenerse en cuenta a la hora de elegir el maquillaje.

Para maquillar a mujeres maduras es bueno tener en cuenta algunas indicaciones para conseguir que resulte un maquillaje natural, que resalte la belleza de la mujer, tratando de iluminar una piel que va estando más apagada con la edad.
Member Testimonials

Written Testimonials

My story has only just begun, I am in my 40’s, unemployed, need hernia surgery and weighed 407 lbs. I could have given up on myself or I could have had obesity surgery. I refused to accept what life was giving me, I made a choice to live. I went to LA Fitness and spoke to one of the kindest men I have ever met. He seemed to really care (Oscar Mosqueda), he signed me up for a personal training which was the best decision of my life. I now weight less than 300 lbs with the extraordinary help of my trainer (Eric Diaz). Eric helps me with my diet and exercise and he also cares. I am at the gym five days a week. Best of all I found new supportive friends, it is my second home. Also, the other trainers have been incredibly supportive such as Ivan (the Adonis-fantastic body) who asks me how I am and how are my workouts. Then Patricia (the Greek Goddess-gorgeous) she is my cheer leader, whenever I am down I seek her out because she will make my day. I am still working at it!

Maribel Lebron

Video Testimonials

We are on this journey together.

Watch other members of LA Fitness share their stories of achievement. By sharing our success we can help each other. Help someone else. Make a video to tell your story.

https://www.lafitness.com/Pages/VidTestimonialPlay.aspx
two communication channels

Written Testimonials

My story has only just begun, I am in my 40’s, unemployed, need hernia surgery and weighed 407 lbs. I could have given up on myself or I could have had obesity surgery. I refused to accept what life was giving me, I made a choice to live. I went to LA Fitness and spoke to one of the kindest men I have ever met. He seemed to really care (Oscar Mosqueda), he signed me up for a personal training which was the best decision of my life. I now weight less than 300 lbs with the extraordinary help of my trainer (Eric Diaz). Eric helps me with my diet and exercise and he also cares. I am at the gym five days a week. Best of all I found new supportive friends, it is my second home. Also, the other trainers have been incredibly supportive such as Ivan (the Adonis- fantastic body) who asks me how I am and how are my workouts. Then Patricia (the Greek Goddess- gorgeous) she is my cheer leader, whenever I am down I seek her out because she will make my day. I am still working at it!

Maribel Lebron

sentiment

topic

language style

opinion

beyond
nonverbal communication

- gaze
- gestures
- prosody and speaking activity
- body posture

Alternative to spoken words. Indicate states, traits, relationships. Unconscious, hard to fake. Accurate judgments.
nonverbal & verbal communication in YouTube video: «in-the-wild» data
research framework

1. video crowdsourcing
   study social perception

2. behavioral analysis
   automatically characterize vloggers
1. video crowdsourcing
   study social perception

can we crowdsource reliable human impressions from vlogs?

a basic model for nonverbal behavior and interpersonal perception

**Figure 2.** The interpersonal circle. Adapted from “A Psychological Taxonomy of Trait-Descriptive Terms: The Interpersonal Domain” by J. S. Wiggins, *Journal of Personality and Social Psychology, 1979, Vol. 37, p. 400. Copyright © 1979 by the American Psychological Association. Adapted by permission.*

vlogger impressions

big-five traits
extraverted
agreeable
conscientious
stable
open

mood
happy
relaxed
bored
stressed
angry
vlog data

shot-based analysis

conversational data

442 YouTube vloggers
53% female / 47% male
first conversational minute
crowdsourcing impressions

watch one-minute slices
answer questionnaires
big-five & mood
5 annotators/vlog
113 workers (US & India)
crowdsourced vlogger demographics (majority vote)

groundtruth: 47% M; 53% F
# impression reliability: intra-class correlation (ICC)

ICC: agreement achieved with aggregated scores across annotators

<table>
<thead>
<tr>
<th>Trait</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extraversion</td>
<td>0.77</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>0.65</td>
</tr>
<tr>
<td>Openness</td>
<td>0.47</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>0.45</td>
</tr>
<tr>
<td>Emotional Stability</td>
<td>0.42</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mood</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Happy</td>
<td>.76</td>
</tr>
<tr>
<td>Excited</td>
<td>.74</td>
</tr>
<tr>
<td>Angry</td>
<td>.67</td>
</tr>
<tr>
<td>Disappointed</td>
<td>.61</td>
</tr>
<tr>
<td>Sad</td>
<td>.58</td>
</tr>
<tr>
<td>Relaxed</td>
<td>.54</td>
</tr>
<tr>
<td>Bored</td>
<td>.52</td>
</tr>
<tr>
<td>Stressed</td>
<td>.50</td>
</tr>
<tr>
<td>Surprised</td>
<td>.48</td>
</tr>
<tr>
<td>Nervous</td>
<td>.25</td>
</tr>
<tr>
<td>Over. mood</td>
<td>.75</td>
</tr>
</tbody>
</table>

- ICC < 0.40 — poor
- ICC in [0.40, 0.59] — fair
- ICC in [0.60, 0.74] — good
- ICC in [0.75, 1.00] — excellent

2. behavioral analysis
automatically characterize vloggers

are verbal & nonverbal cues linked with personality impressions in vlogging?

nonverbal cue extraction (1)

from activity segmentations

AUDIO
- Speaking time
- Num Turns

VISUAL
- Looking time
- Num Turns

MULTIMODAL
- Looking & Speaking
- Looking & not-speaking

AUDIO
Prosody
- Energy
- Pitch
- Voice rate

VISUAL
- Proximity to camera
- Framing

- Accumulated motion
nonverbal cues (2)

Computer Expression Recognition Toolbox (CERT)

- basic facial expressions
- smile

raw statistical aggregates from segmentations

- Active Time
- Num segments
verbal cues

Linguistic Inquiry Word Count (LIWC)

- 65 categories related to psychological constructs and personal concerns
- word count per category
### big-five impressions & audio/visual activity cues

#### Cue utilization (# of significant correlations)

<table>
<thead>
<tr>
<th>Trait</th>
<th>Extr</th>
<th>Cons</th>
<th>Open</th>
<th>Agr</th>
<th>Emot</th>
</tr>
</thead>
<tbody>
<tr>
<td># cues</td>
<td>24</td>
<td>16</td>
<td>12</td>
<td>10</td>
<td>3</td>
</tr>
</tbody>
</table>

Extraversion: highest cue utilization

Agreeableness: low cue utilization

Effects backed up by social psych literature

#### Selected Effects (p<0.05)

- **Extr**
  - Speaking time (talkative)
  - Visual activity (dynamic)
  - Dominance ratio
  - # Speaking turns (fluency)

- **Cons**
  - Looking time (persistent gaze)
  - Visual activity (quiet)

- **Agr**
  - Vertical framing (upper body)
big-five impressions & facial expression cues

Cue utilization (# of significant correlations)

<table>
<thead>
<tr>
<th>Trait</th>
<th>Extr</th>
<th>Open</th>
<th>Agr</th>
<th>Cons</th>
<th>Emot</th>
</tr>
</thead>
<tbody>
<tr>
<td># cues</td>
<td>75</td>
<td>37</td>
<td>27</td>
<td>17</td>
<td>9</td>
</tr>
</tbody>
</table>

Extraversion: highest cue utilization

Selected Effects (p<0.05)

<table>
<thead>
<tr>
<th>Trait</th>
<th>+</th>
<th>-</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extr</td>
<td>joy, smile</td>
<td>anger, disgust</td>
</tr>
<tr>
<td>Open</td>
<td>surprise</td>
<td>anger</td>
</tr>
<tr>
<td>Agr</td>
<td>joy, smile</td>
<td>anger</td>
</tr>
</tbody>
</table>

Facial expressions also favor impressions of Openness and Agreeableness
### big-five impressions & verbal cues

#### Cue utilization (# of significant correlations)

<table>
<thead>
<tr>
<th>Trait</th>
<th>Cons</th>
<th>Agr</th>
<th>Extr</th>
<th>Emot</th>
<th>Open</th>
</tr>
</thead>
<tbody>
<tr>
<td># cues</td>
<td>24</td>
<td>15</td>
<td>12</td>
<td>12</td>
<td>10</td>
</tr>
</tbody>
</table>

Higher cue utilization for Conscientiousness (and Agreeableness)

#### Selected Effects (p<0.05)

<table>
<thead>
<tr>
<th>Trait</th>
<th>+</th>
<th>-</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cons</td>
<td>work, achieve (long words)</td>
<td>negate, swear</td>
</tr>
<tr>
<td>Agr</td>
<td>posemo, i, friend</td>
<td>anger, negemo,</td>
</tr>
<tr>
<td>Emot</td>
<td>leisure, work</td>
<td>negemo, affect</td>
</tr>
<tr>
<td>Extr</td>
<td>you, social, sexual</td>
<td>tentative, exclusive</td>
</tr>
</tbody>
</table>

Many effects backed up by previous research in social psychology
personality inference

5 regression tasks on Big-Five scores machine learning: SVM and RF performance measured with $R^2$

$$R^2 = 100 \times \left( 1 - \frac{\sum (y_{obs} - y_{pred})^2}{\sum (y_{obs} - \bar{y}_{obs})^2} \right)$$

squared sum of prediction errors

squared sum of prediction errors using mean predictor
### Prediction Results (Random Forest)

<table>
<thead>
<tr>
<th>Trait</th>
<th>AV</th>
<th>FE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extr</td>
<td>.39</td>
<td>.24</td>
</tr>
<tr>
<td>Cons</td>
<td>.10</td>
<td>.06</td>
</tr>
<tr>
<td>Open</td>
<td>.10</td>
<td>.11</td>
</tr>
<tr>
<td>Agr</td>
<td>.06</td>
<td>.08</td>
</tr>
<tr>
<td>Emot</td>
<td>.06</td>
<td>.07</td>
</tr>
</tbody>
</table>

AV = audiovisual activity cues  
FE = facial expression cues

<table>
<thead>
<tr>
<th>Trait</th>
<th>Manual Transcripts</th>
<th>Automatic Transcripts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agr</td>
<td>.31</td>
<td>.10</td>
</tr>
<tr>
<td>Cons</td>
<td>.18</td>
<td>.08</td>
</tr>
<tr>
<td>Emot</td>
<td>.17</td>
<td>.05</td>
</tr>
<tr>
<td>Extr</td>
<td>.13</td>
<td>.02</td>
</tr>
<tr>
<td>Open</td>
<td>.04</td>
<td>.02</td>
</tr>
</tbody>
</table>

ASR has significantly improved since 2013
feature fusion for personality impression inference

Extraversion

\[ AV + FE \rightarrow R^2 = .48 \]

Openness:

\[ AV + FE \rightarrow R^2 = .17 \]

Agreeableness

\[ FE + VB \rightarrow R^2 = .39 \]
what to remember
online video is different than text & images continues to rise

YouTube is multifaceted

cconversational video
audio-visual behavioral analysis
inference of perceived traits & states
personal & cultural differences
many applications
questions?

gatica@idiap.ch
daniel.gatica-perez@epfl.ch