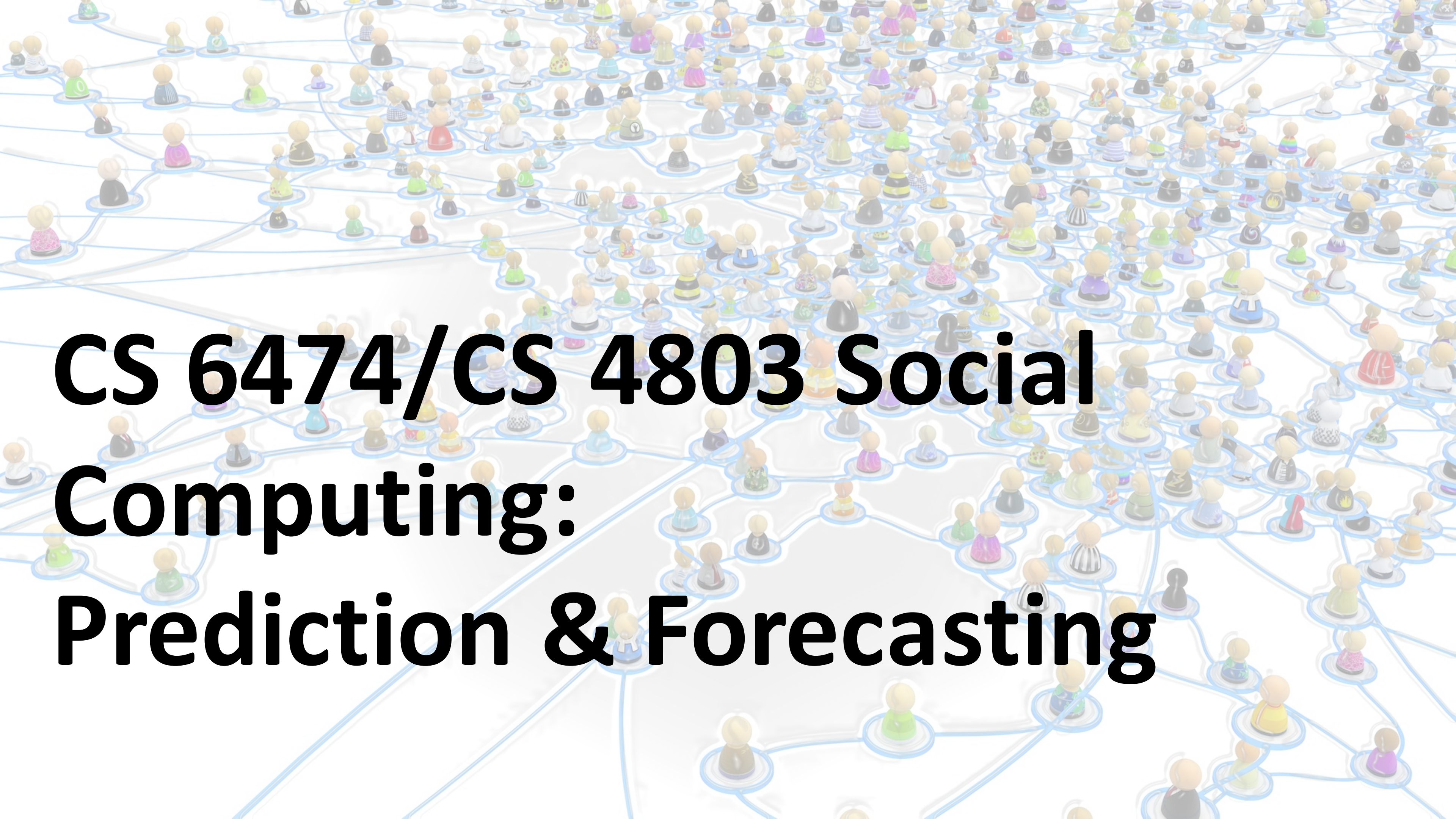


CS 6474/CS 4803 Social Computing: Prediction & Forecasting

Munmun De Choudhury

munmund@gatech.edu

Week 13 | November 14, 2018

Predicting the Future With Social Media

Sitaram Asur
Social Computing Lab
HP Labs
Palo Alto, California
Email: sitaram.asur@hp.com

Bernardo A. Huberman
Social Computing Lab
HP Labs
Palo Alto, California
Email: bernardo.huberman@hp.com

***Abstract*—In recent years, social media has become ubiquitous and important for social networking and content sharing. And yet, the content that is generated from these websites remains largely untapped. In this paper, we demonstrate how social media content can be used to predict real-world outcomes. In particular, we use the chatter from Twitter.com to forecast box-office revenues for movies. We show that a simple model built from the rate at which tweets are created about particular topics can outperform market-based predictors. We further demonstrate how sentiments extracted from Twitter can be further utilized to improve the forecasting power of social media.**

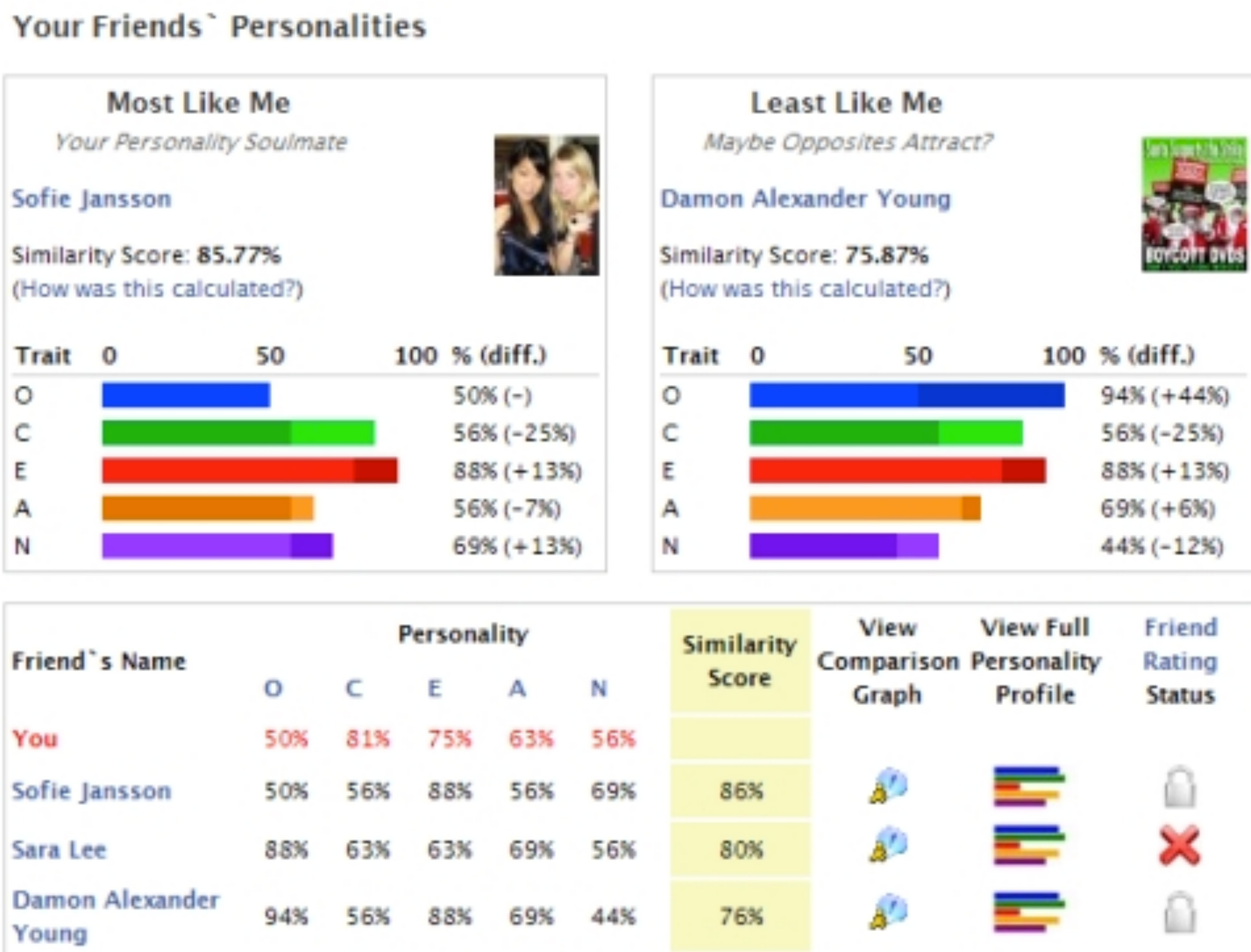
This paper reports on such a study. Specifically we consider the task of predicting box-office revenues for movies using the chatter from Twitter, one of the fastest growing social networks in the Internet. Twitter [1], a micro-blogging network, has experienced a burst of popularity in recent months leading to a huge user-base, consisting of several tens of millions of users who actively participate in the creation and propagation of content.

We have focused on movies in this study for two main reasons.

Private Traits and Attributes are Predictable from Digital Records of Human Behavior

Summary

- Facebook “likes” used to predict a range of highly sensitive personal attributes like ethnicity, religious and political views, intelligence, happiness, parental separation, age and gender.
- 58K users of Facebook who consented to authorize the mypersonality app
- Participants took many sociometric and psychometric tests
- Predictive accuracies were very high for sexual orientation, parental separation, political views, and the openness attribute of Big Five personality scale
  - The algorithms proved 88% accurate for determining male sexuality, 95% accurate in distinguishing African-American from Caucasian-American and 85% for differentiating Republican from Democrat.
  - Christians and Muslims were correctly classified in 82% of cases and relationship status and substance abuse was predicted with an accuracy between 65% and 73%.


Your Friends` Personalities
Most Like Me
Your Personality Soulmate
Sofie Jansson
Similarity Score: 85.77%
(How was this calculated?)
Trait 0 50 100 % (diff.)
O 50% (-)
C 56% (-25%)
E 88% (+13%)
A 56% (-7%)
N 69% (+13%)
Least Like Me
Maybe Opposites Attract?
Damon Alexander Young
Similarity Score: 75.87%
(How was this calculated?)
Trait 0 50 100 % (diff.)
O 94% (+44%)
C 56% (-25%)
E 88% (+13%)
A 69% (+6%)
N 44% (-12%)
Friend`s Name
Personality
O C E A N
Similarity Score
View Comparison Graph
View Full Personality Profile
Friend Rating Status
You 50% 81% 75% 63% 56%
Sofie Jansson 50% 56% 88% 56% 69% 86%
Sara Lee 88% 63% 63% 69% 56% 80%
Damon Alexander Young 94% 56% 88% 69% 44% 76%

- Start
- About myPersonality
- Register as a collaborator
- List of available variables
- Download databases
- Collaborators and Publications
- Research Ideas
- How to cite myPersonality
- Additional resources
- WCPR13

start

Table of Contents

Welcome to the myPersonality Project Website

++

If you're here because of the news coverage:

This wiki is aimed at researchers, although you're welcome to look around and see what we do.

We also encourage you to try http://www.YouAreWhatYouLike.com which predicts your personality based on your Facebook Likes.

++

News

- 2013-04-22 Added Smiley data in the download section
- 2013-02-12 There were 4 new papers based on our data in January of 2013 alone - in PNAS, PLOS ONE, WWW2013, and CWSM2013. Congratulations to authors!
- 2012-10-24 **LOADS OF NEW DATA** AND IMPROVED LAYOUT! Check out download databases section.
- Last.FM music DB collected by Liam McNamara. Click here for full details.
- news archive

Introduction

11 March 2013 Last updated at 15:00 ET

Share

Facebook 'likes' predict personality

Sexuality, political leanings and even intelligence can be gleaned from the things you choose to "like" on Facebook, a study suggests.

What do your Facebook likes say about you?

Researchers at Cambridge University used algorithms to predict religion, politics, race and sexual orientation.

The research, published in the journal PNAS, forms surprisingly accurate personal portraits, researchers said.

The findings should "ring alarm bells" for users, privacy campaigners said.

The study used 58,000 volunteers who alongside their Facebook "likes" and demographic information also provided psychometric testing results - designed to highlight personality traits.

The Facebook likes were fed into algorithms and matched with the

Related Stories

Facebook changes privacy settings

Privacy groups lobby Facebook

Top stories

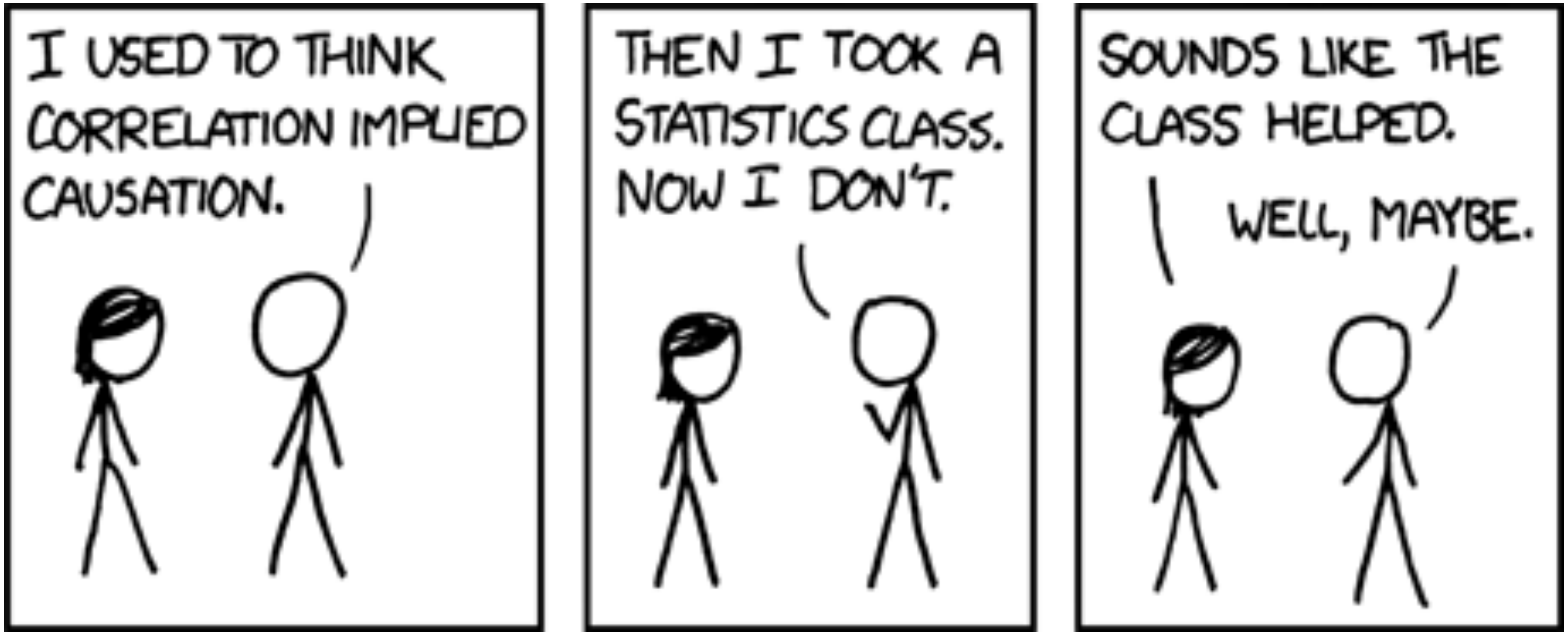
Soldier killed in Canada shootings

WHO crisis team holds Ebola talks

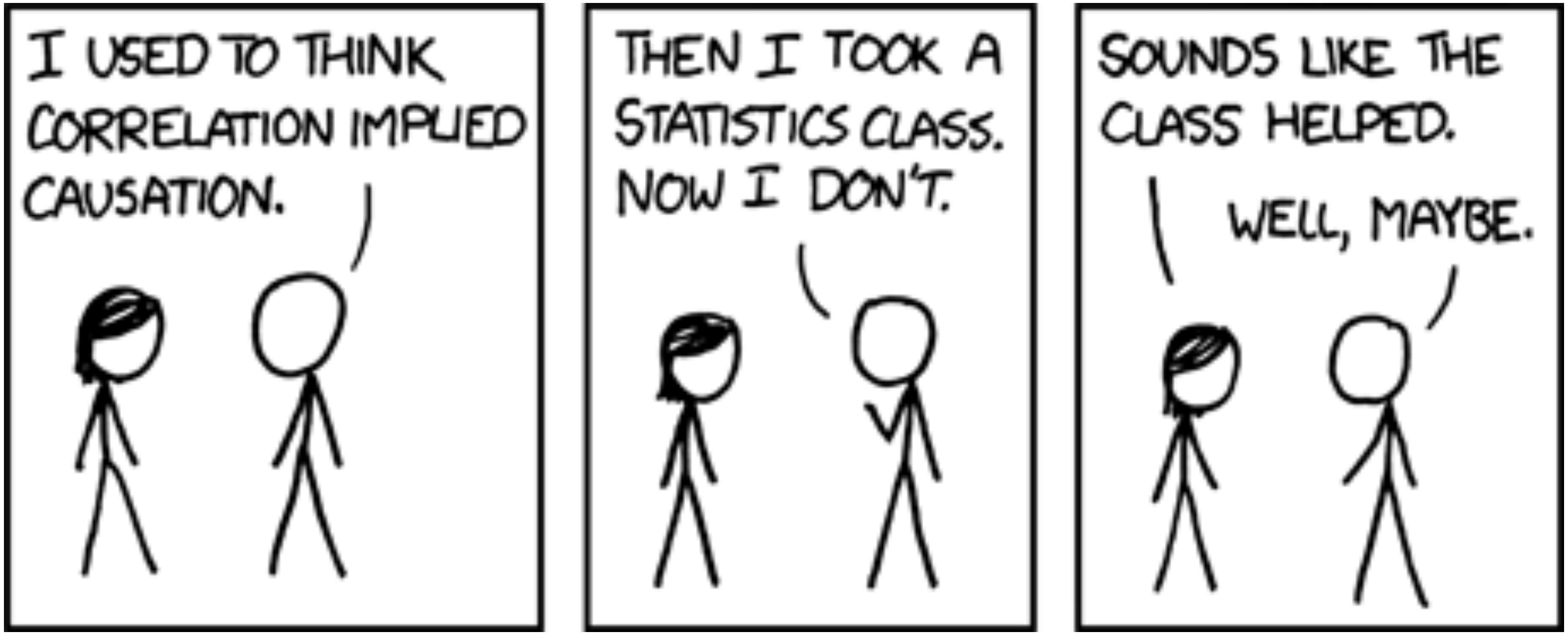
Sweden could use force in sub hunt

U-boat found with ship it sank

Total plane crash driver 'got lost'



Correlation and causation



Do you think the links between Facebook likes and private traits like personality can be causal or purely correlational? If causal, what is the direction of causality?

Class Exercise I

Instead of Facebook “likes”, if you were to predict individual attributes on Twitter, which cues would you use and why?

"This research should ring alarm bells for anyone who thinks that privacy settings are the solution to protecting information online. We need to fundamentally re-think how much data we are voluntarily sharing," said Nick Pickles, director of privacy campaign group Big Brother Watch.

"Yet again, it is clear the lack of transparency about how users' data is being used will lead to entirely justified fears about our data being exploited for commercial gain."

Class Exercise II

Ability to infer accurately individual traits can have implications in better personalization and search, what are its risks in privacy?

What are the other implications of such inferences?

Class Exercise III

Do you think it is ethical to release a Facebook app and use it to collect people's data? What are the challenges of such approaches? What type of biases does it introduce?

Deep neural networks are more accurate than humans at detecting sexual orientation from facial images.

Yilun Wang, Michal Kosinski

Created on: September 07, 2017 | Last edited: October 16, 2017

Page: 1 of 47 Automatic Zoom

DEEP NEURAL NETWORKS CAN DETECT SEXUAL ORIENTATION FROM FACES

THIS IS A PREPRINT OF THE PEER REVIEWED ARTICLE TO APPEAR IN JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY.

THE MOST RECENT VERSION IS AVAILABLE AT https://osf.io/zn79k/

AUTHOR NOTES ARE AVAILABLE AT: https://goo.gl/9b2aR2

Deep neural networks are more accurate than humans at detecting sexual orientation from facial images

Yilun Wang, Michal Kosinski

Graduate School of Business, Stanford University, Stanford, CA94305, USA

michalk@stanford.edu

The study has been approved by the IRB at Stanford University

Citation: Wang, Y., & Kosinski, M. (in press). Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. *Journal of Personality and*

wang_kosinski.pdf

Download preprint Version: 10

Abstract

We show that faces contain much more information about sexual orientation than can be perceived and interpreted by the human brain. We used deep neural networks to extract features from 35,326 facial images. These features were entered into a logistic regression aimed at classifying sexual orientation. Given a single facial image, a classifier ...

See more

Preprint DOI

10.17605/OSF.IO/HV28A

License

Disciplines

Social and Behavioral Sciences | Psychology

Tags

Artificial Intelligence | Big Data | Computational Social Science | Computer Vision | Facial recognition | Prenatal Hormone Theory | Privacy | Sexual orientation

Citations

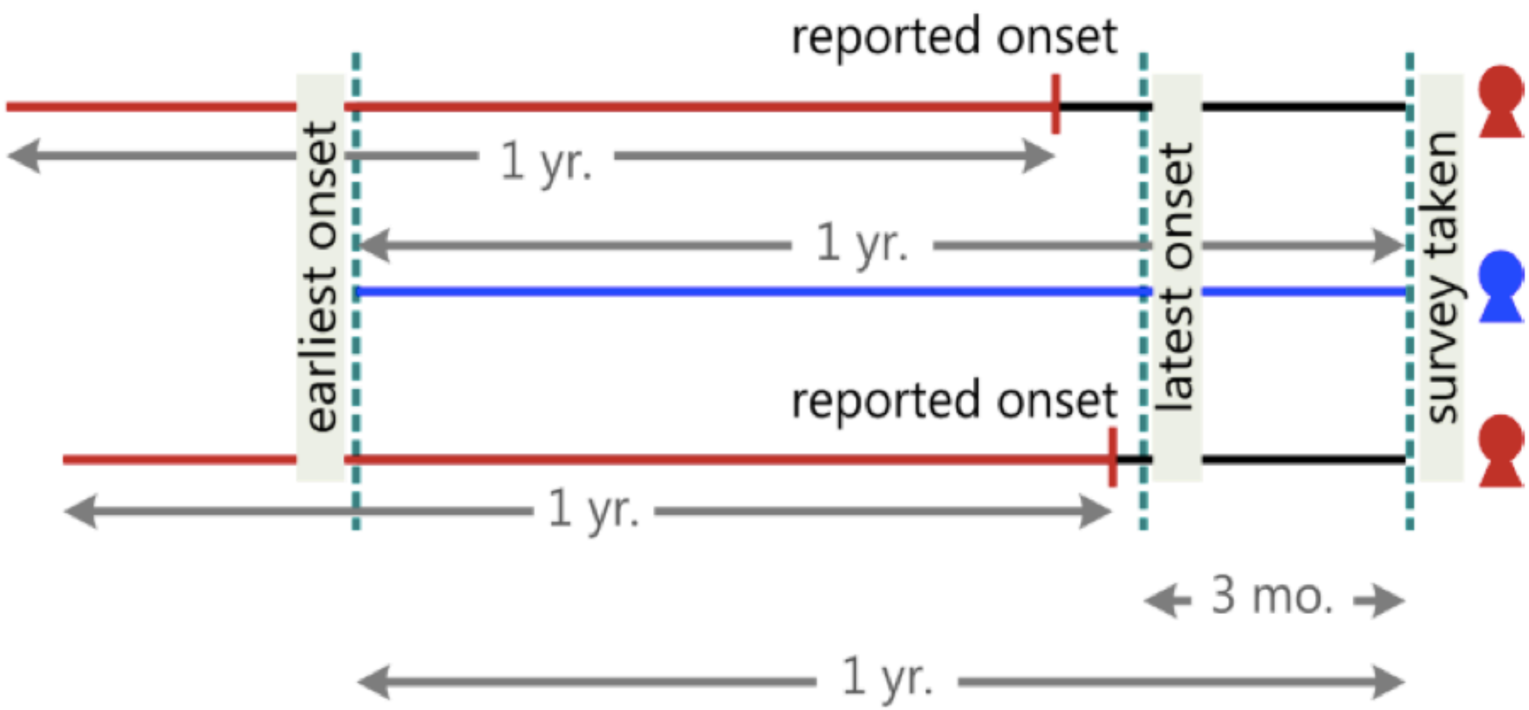
APA

Wang, Y., & Kosinski, M. (2017, October 16). Deep neural networks are more accurate than humans at detecting sexual orientation from facial images.. Retrieved from psyarxiv.com/hv28a

Predicting Depression via Social Media

Summary

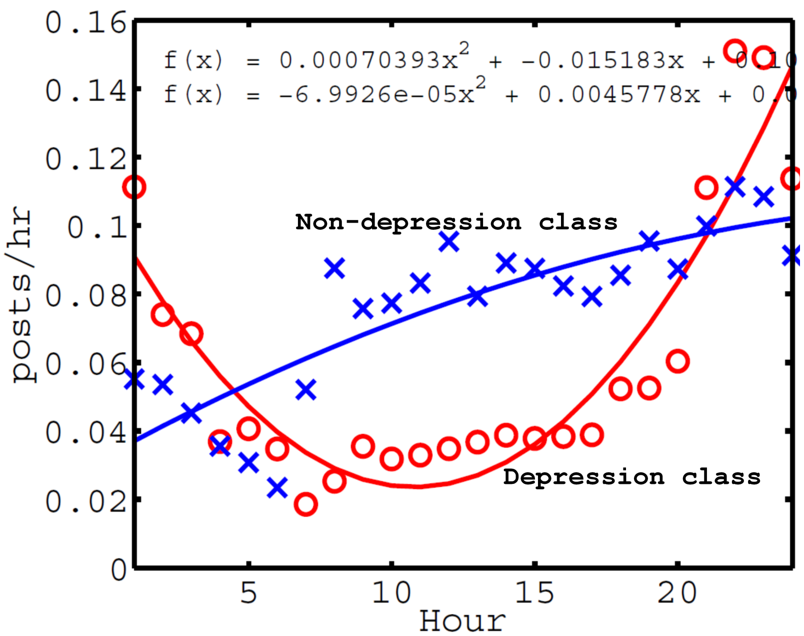
- Can social media activities and connectedness predict risk to major depressive disorder?
- Recruitment of a sample of Twitter users through a survey methodology over Amazon's Mechanical Turk
  - ~40% provided access to Twitter data



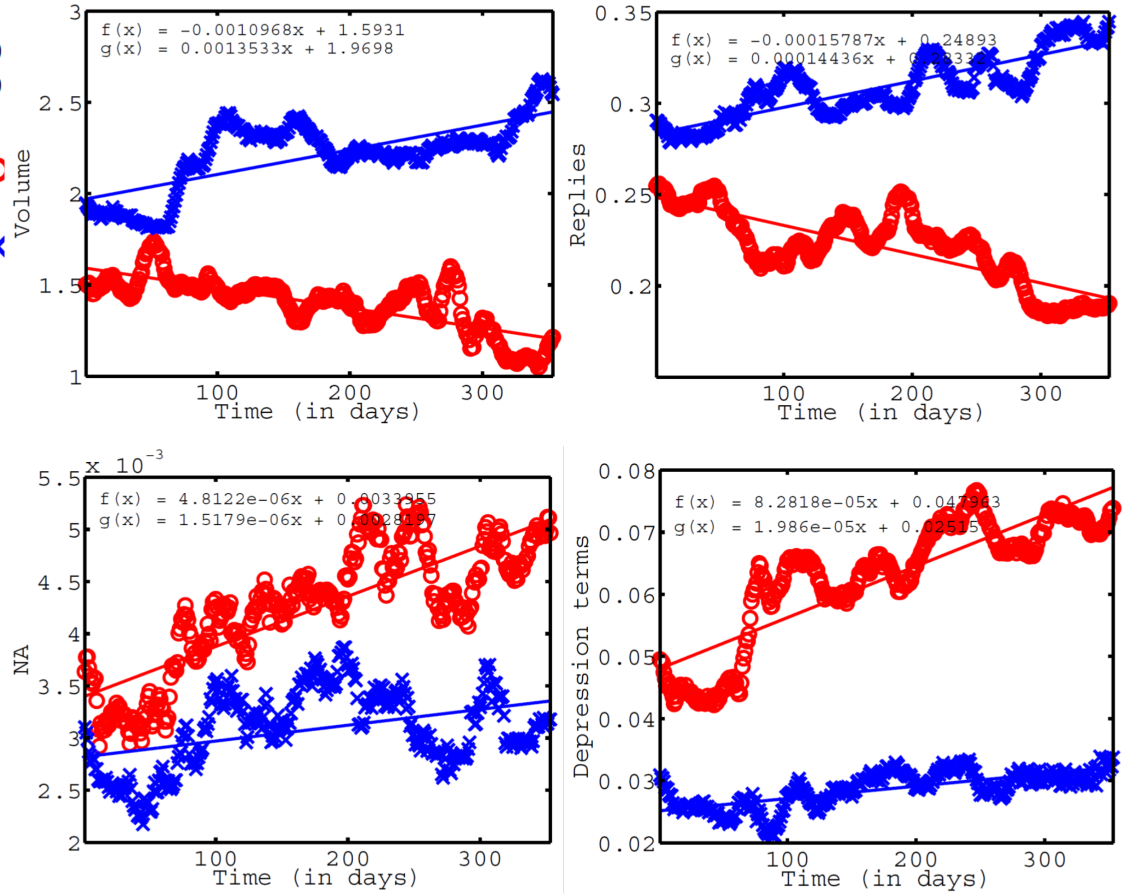
Summary

- Social engagement
- “Insomnia index” – mean *z*-score of an individual’s volume of Twitter activity per hour
- Ego-centric social graph – nodal properties (*inlinks, outlinks*); dyadic properties (*reciprocity, interpersonal exchange*); neighborhood properties (*density, clustering coefficient, two-hop neighborhood, embeddedness, number of ego components*)
- Language
  - Depression lexicon – top uni- and bigrams compiled from Yahoo! Answers category on mental health
  - Linguistic style

Summary



f(x) = 0.00070393x2 + -0.015183x + 0.10
f(x) = -6.9926e-05x2 + 0.0045778x + 0.0
Non-depression class
Depression class
posts/hr
Hour



f(x) = -0.0010968x + 1.5931
g(x) = 0.0013533x + 1.9698
Volume
Time (in days)
f(x) = -0.00015787x + 0.24893
g(x) = 0.00014436x + 0.28332
Replies
Time (in days)
x 10-3
f(x) = 4.8122e-06x + 0.0033955
g(x) = 1.5179e-06x + 0.0028197
NA
Time (in days)
f(x) = 8.2818e-05x + 0.047963
g(x) = 1.986e-05x + 0.02515
Depression terms
Time (in days)

Summary

Egonetwork measures	**Depres. class**	**Non-depres. class**
#followers/inlinks	26.9 (σ=78.3)	45.32 (σ=90.74)
#followees/outlinks	19.2 (σ=52.4)	40.06 (σ=63.25)
Reciprocity	0.77 (σ=0.09)	1.364 (σ=0.186)
Prestige ratio	0.98 (σ=0.13)	0.613 (σ=0.277)
Graph density	0.01 (σ=0.03)	0.019 (σ=0.051)
Clustering coefficient	0.02 (σ=0.05)	0.011 (σ=0.072)
2-hop neighborhood	104 (σ=82.42)	198.4 (σ=110.3)
Embeddedness	0.38 (σ=0.14)	0.226 (σ=0.192)
#ego components	15.3 (σ=3.25)	7.851 (σ=6.294)

Class Exercise IV

In the depression prediction paper, the ground truth was obtained from Amazon mechanical turk workers. Anything unique about this population that may have affected the findings? What would be alternative ways of recruiting people or gathering high quality ground truth?

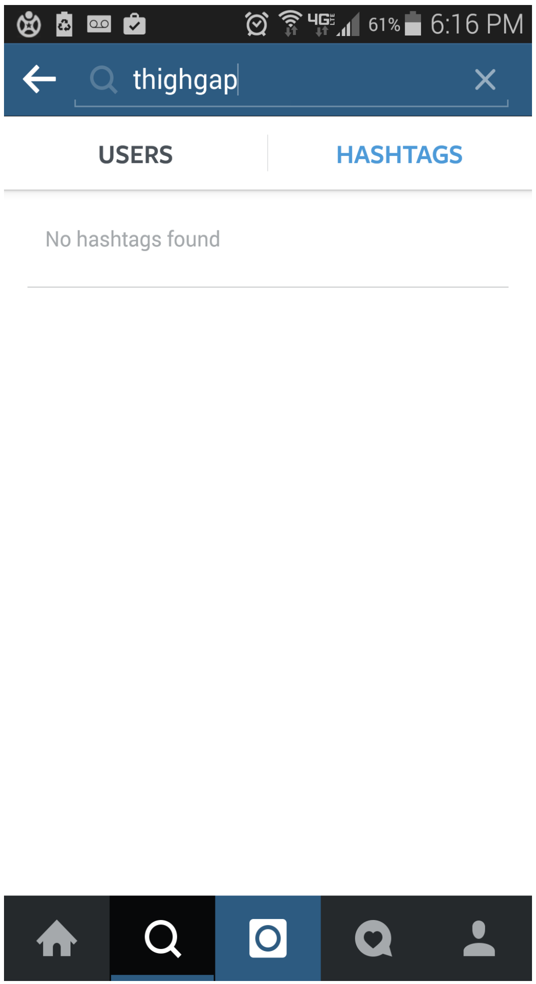
Class Exercise V

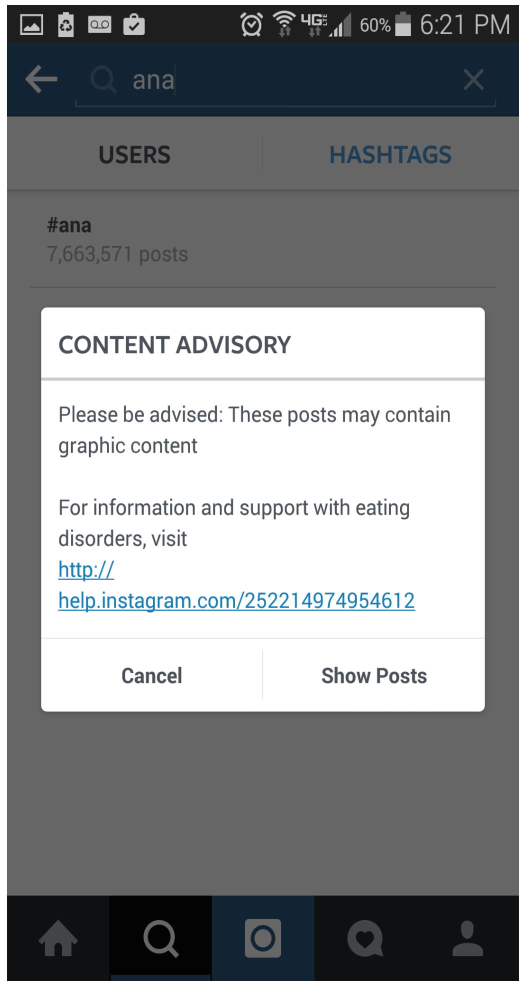
Depression is not an online condition, but one that spans both the online and the offline life. The paper does not take offline attributes into their models.

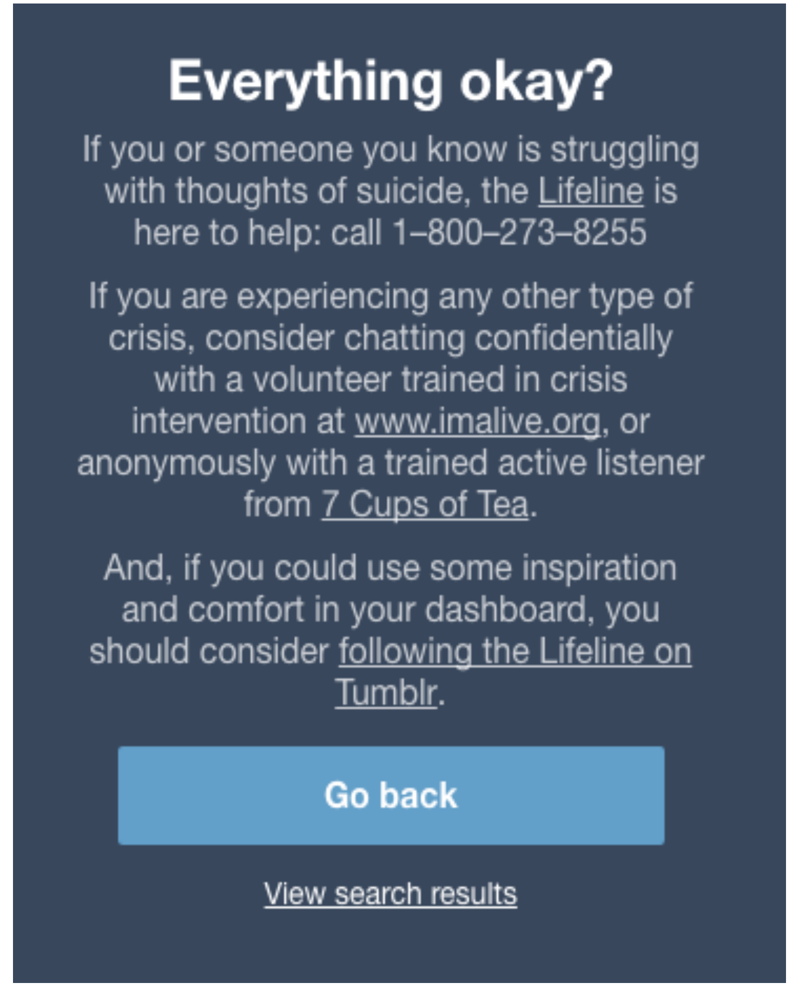
Is there a way to that into account? What would be the most significant offline attributes to consider?

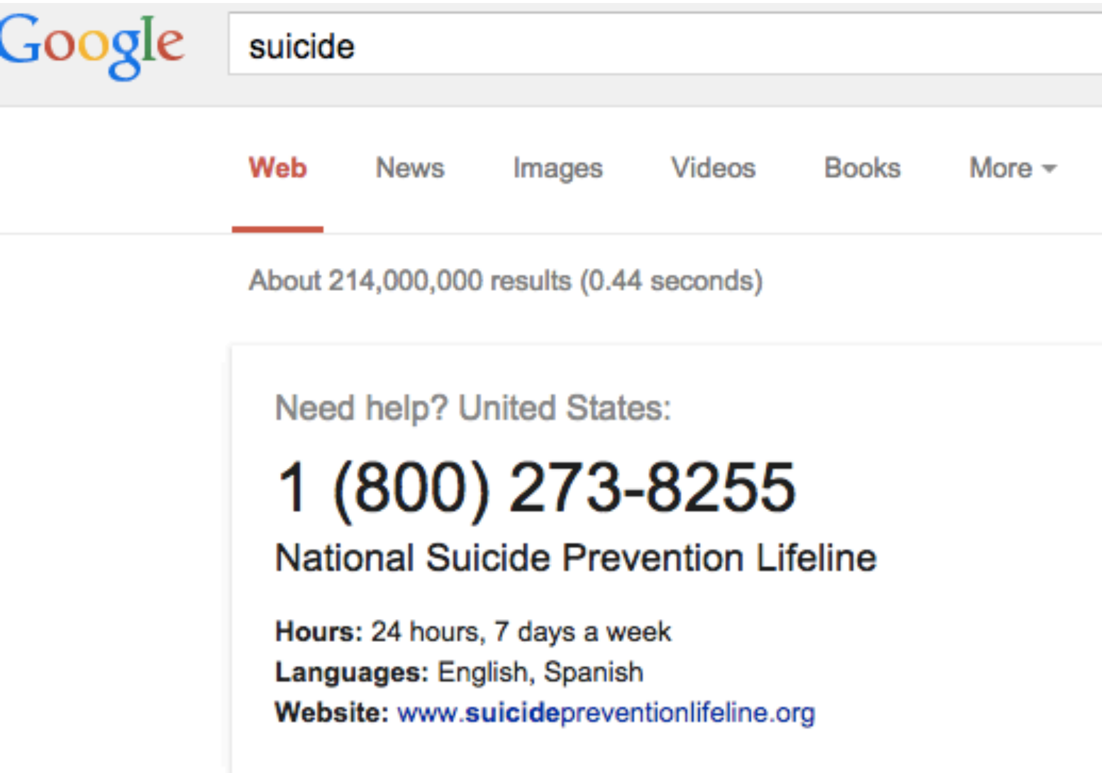
Discuss how can social media based depression (or other mental health condition) predictors could be helpful to people.

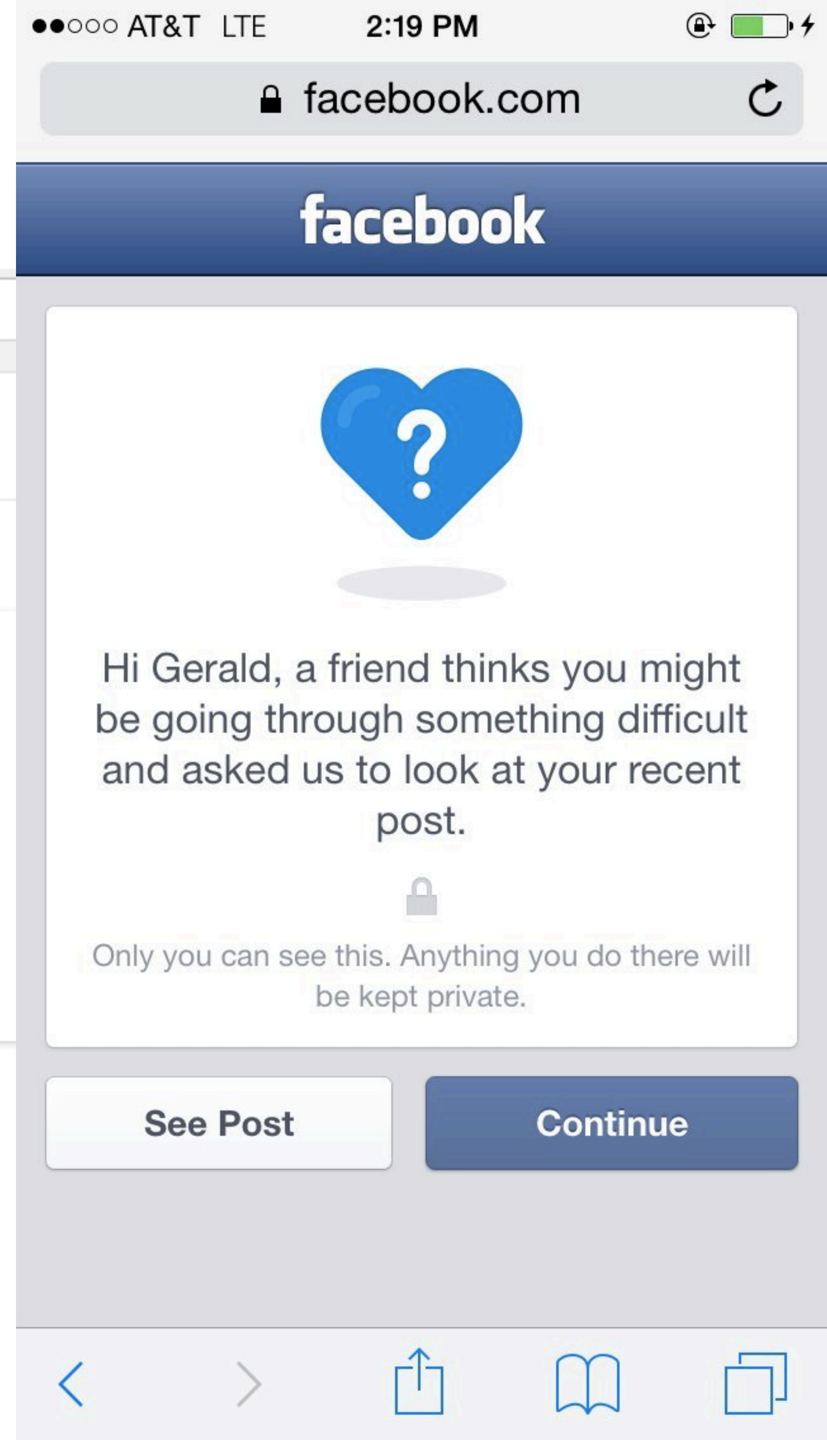
Improving “Blanket” Interventions








Google
suicide
Web
News
Images
Videos
Books
More
About 214,000,000 results (0.44 seconds)
Need help? United States:
1 (800) 273-8255
National Suicide Prevention Lifeline
Hours: 24 hours, 7 days a week
Languages: English, Spanish
Website: www.suicidepreventionlifeline.org


AT&T LTE
2:19 PM
facebook.com
facebook
Hi Gerald, a friend thinks you might be going through something difficult and asked us to look at your recent post.
Only you can see this. Anything you do there will be kept private.
See Post
Continue

Prediction and explanation in social systems

Summary

- Prediction is becoming central in the study of (online) social systems
- Papers presents three issues that need resolution to be able to derive value out of predictive approaches

- Standards of prediction:
  - Use multiple evaluation metrics
  - Models are evaluated by third parties (e.g., the Netflix prize)
  - Begin with exploratory research – move to confirmatory; register research design data etc.
- Limits of prediction
  - Theoretical limit to predictive accuracy
  - Consideration of confounding factors
  - Calibrate expectations; reduce false optimism
- Prediction versus interpretation
  - Prediction and interpretation do not have to be a trade-off
  - Hybrid approach where simple and complex methods are combined and the solution is question driven

What are the limits of prediction? Can they fail?

BIG DATA

The Parable of Google Flu: Traps in Big Data Analysis

Large errors in flu prediction were largely avoidable, which offers lessons for the use of big data.

David Lazer,[1,2*] Ryan Kennedy,[1,3,4] Gary King,[3] Alessandro Vespignani[3,5,6]

In February 2013, Google Flu Trends (GFT) made headlines but not for a reason that Google executives or the creators of the flu tracking system would have hoped. *Nature* reported that GFT was predicting more than double the proportion of doctor visits for influenza-like illness (ILI) than the Centers for Disease Control and Prevention (CDC), which bases its estimates on surveillance reports from laboratories across the United States (*1*, *2*). This happened despite the fact that GFT was built to predict CDC reports. Given that GFT is often held up as an exemplary use of big data (*3*, *4*), what lessons can we draw from this error?

The problems we identify are not limited to GFT. Research on whether search or social media can

the algorithm in 2009, and this model has run ever since, with a few changes announced in October 2013 (*10*, *15*).

Although not widely reported until 2013, the new GFT has been persistently overestimating flu prevalence for a much longer time. GFT also missed by a very large margin in the 2011–2012 flu season and has missed high for 100 out of 108 weeks starting with August 2011 (see the graph). These errors are not randomly distributed. For example, last week's errors predict this week's errors (temporal autocorrelation), and the direction and magnitude of error varies with the time of year (seasonality). These patterns mean that GFT overlooks considerable information that could be extracted by traditional

Meaningless comparisons lead to false optimism in medical machine learning

Orianna DeMasi[1], Konrad Kording[2, 3], and Benjamin Recht[1]

[1]Department of Electrical Engineering and Computer Sciences, University of California Berkeley, Berkeley, CA, USA
[2]Department of Bioengineering, University of Pennsylvania, Philadelphia, PA, USA
[3]Department of Neuroscience, University of Pennsylvania, Philadelphia, PA, USA
odemasi@berkeley.edu, kording@upenn.edu, brecht@berkeley.edu

July 21, 2017

Abstract

A new trend in medicine is the use of algorithms to analyze big datasets, e.g. using everything your phone measures about you for diagnostics or monitoring. However, these algorithms are commonly compared against weak baselines, which may contribute to excessive optimism. To assess how well an algorithm works, scientists typically ask how well its output correlates with medically assigned scores. Here we perform a meta-analysis to quantify how the literature evaluates their algorithms for monitoring mental wellbeing. We find that the bulk of the literature ($\sim$77%) uses meaningless comparisons that ignore patient baseline state. For example, having an algorithm that uses phone data to diagnose mood disorders would be useful. However, it is possible to over 80% of the variance of some mood measures in the population by simply guessing that each patient has their own average mood - the patient-specific baseline. Thus, an algorithm that just predicts that our mood is like it usually is can explain the majority of variance, but is, obviously, entirely useless. Comparing to the wrong (population) baseline has a massive effect on the perceived quality of algorithms and produces baseless optimism in the field. To solve this problem we propose "user lift" that reduces these systematic errors in the evaluation of personalized medical monitoring.

n terms that match the propen-
. but are structurally unrelated,
t predict the future, were quite
evelopers, in fact, report weed-
nal search terms unrelated to the
ly correlated to the CDC data,
regarding high school basket-
s should have been a warning
lata were overfitting the small
ses, a standard concern in data
s ad hoc method of throwing
search terms failed when GFT
nissed the nonseasonal 2009
-H1N1 pandemic (*2*, *14*). In
ial version of GFT was part flu
winter detector. GFT engineers

Considering the large number of approaches that provide inference on influenza activity (*16–19*), does this mean that the current version of GFT is not useful? No, greater value can be obtained by combining GFT with other near–real time health data (*2*, *20*). For example, by combining GFT and lagged CDC data, as well as dynamically recalibrating GFT, we can substantially improve on the performance of GFT or the CDC alone (see the chart). This is no substitute for ongoing evaluation and improvement, but, by incorporating this information, GFT could have largely healed itself and would have likely remained out of the headlines.

Source: D. Lazer, R. Kennedy, G. King, A. Vespignani. "The Parable of Google Flu: Traps in Big Data Analysis," Science, vol. 343, March 14, 2014

Class Exercise VI

Assess whether in each of the following cases interpretation or prediction (or both) is/are preferred.

i) How people react on a new product release (e.g., an iphone), as observed on social media
ii) Whether greater anonymity leads to greater hate speech on social media
iii) Whether more ads on YouTube videos leads to lesser YouTube use
iv) Whether exposure to similar ideological content leads to reinforcement of existing ideologies
v) Assessments of inconvenience of the I85 highway crash, based on people's social media activity