

CS 6474/CS 4803 Social Computing: Prediction & Forecasting II

Munmun De Choudhury

munmund@gatech.edu

Week 14 | November 20, 2017

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. Twitter¹, 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%.

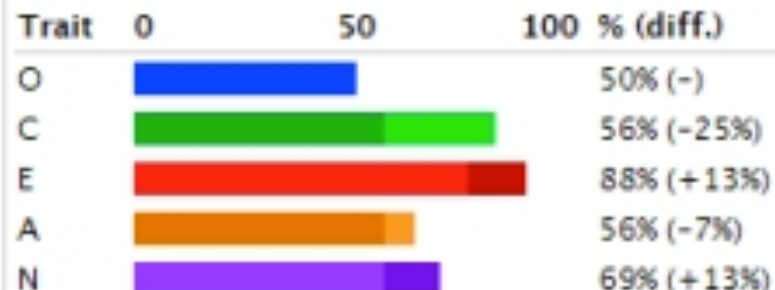
Your Friends` Personalities

Most Like Me

Your Personality Soulmate

Sofie Jansson

Similarity Score: 85.77%
(How was this calculated?)

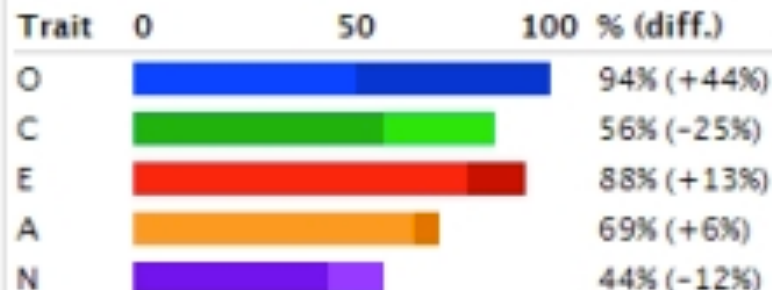


Least Like Me

Maybe Opposites Attract?

Damon Alexander Young

Similarity Score: 75.87%
(How was this calculated?)



Friend`s Name	Personality					Similarity Score	View Comparison Graph	View Full Personality Profile	Friend Rating Status
	O	C	E	A	N				
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%			

Trace: • [start](#)

- [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](#)

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](#)

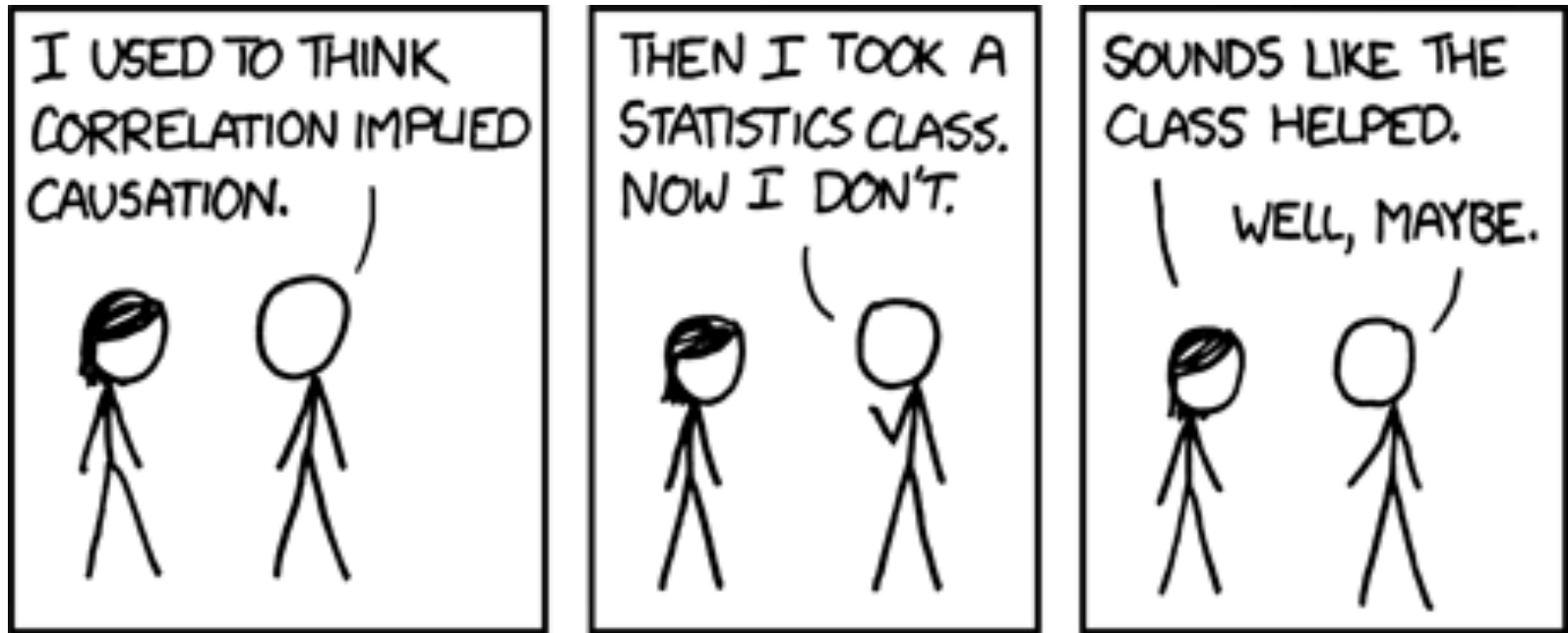
Introduction

Table of Contents

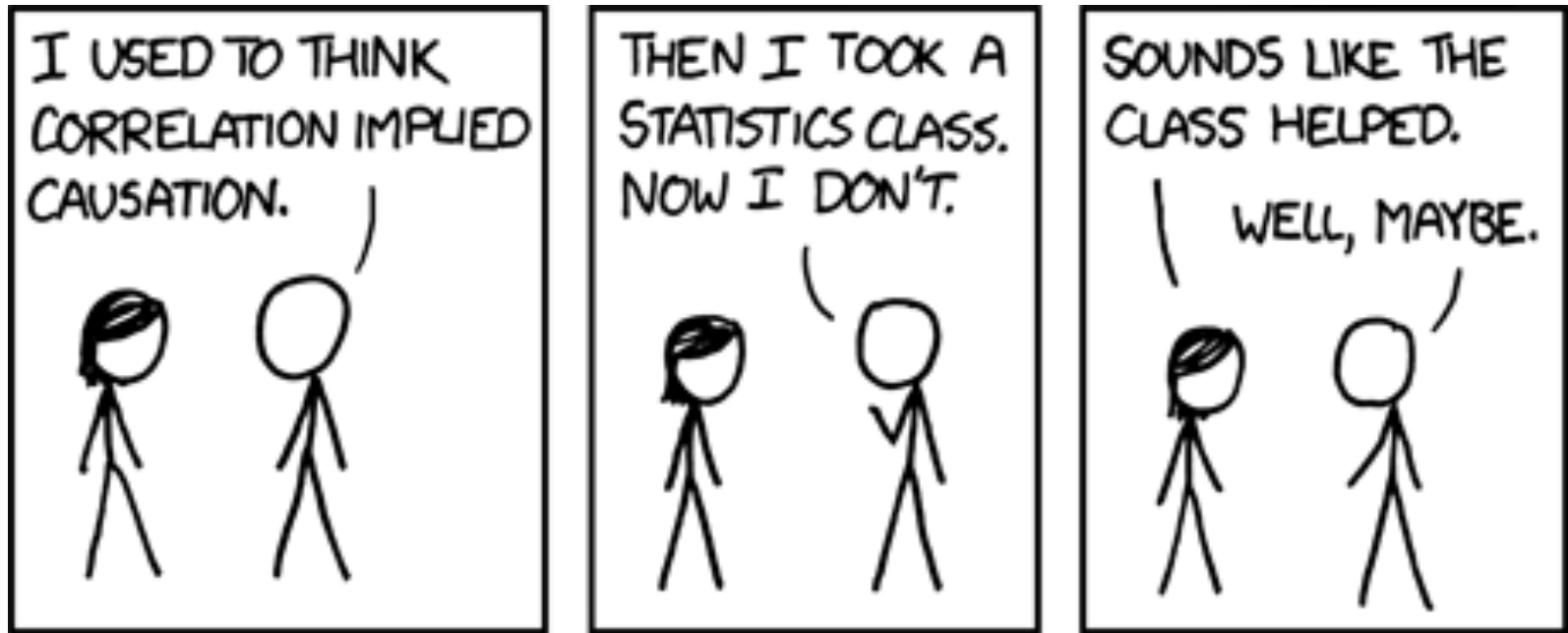
- ◊ [Welcome to the myPersonality Project Website](#)
 - ◊ [If you're here because of the news coverage:](#)
- ◊ [News](#)
- ◊ [Introduction](#)
 - ◊ [What are we talking about](#)

start





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?

Online tools may be providing “safe” places for introverts to express themselves; so they may appear less of an introvert than they actually are. Would this gap be accounted for by the analysis in the first paper?

One reason why Facebook “likes” are perhaps predictive of individual attributes is because it is a general platform. Could Instagram, Tumblr, or Reddit activities equally inform us of the same?

Class Exercise I

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

Class Exercise II

Do you think it is ethical to release a Facebook app and use it to collect people's data? How would you design such a study? What are the challenges of opt-in methods?

NEWS TECHNOLOGY

[Home](#) [US & Canada](#) [Latin America](#) [UK](#) [Africa](#) [Asia](#) [Australia](#) [Europe](#) [Mid-East](#) [Business](#) [Health](#) [Sci/Environment](#) [Tech](#) [Entertainment](#) [Video](#)

REGENERIST LUMINOUS
Add **NEW** Tone Perfecting Treatment to your regimen.

[SHOP NOW >](#)

YOUR BEST BEAUTIFUL™

ADVERTISEMENT

AdChoices



11 March 2013 Last updated at 15:00 ET



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.

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



What do your Facebook likes say about you?

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'](#)

ADVERTISEMENT



Patenta

"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."

Samsung Smart TV privacy policy

Recognition features to you. In addition, Samsung may collect and your device may capture voice commands and associated texts so that we can provide you with Voice Recognition features and evaluate and improve the features. Please be aware that if your spoken words include personal or other sensitive information, that information will be among the data captured and transmitted to a third party through your use of Voice Recognition.

If you do not enable Voice Recognition, you will not be able to use interactive voice recognition features, although you may be able to control your TV using certain predefined voice commands. While Samsung will not collect your spoken word, Samsung may still collect associated texts and other usage data so that

George Orwell, 1984

Behind Winston's back the voice from the telescreen was still babbling away about pig-iron and the overfulfilment of the Ninth Three-Year Plan. The telescreen received and transmitted simultaneously. Any sound that Winston made, above the level of a very low whisper, would be picked up by it, moreover, so long as he remained within the field of vision which the metal plaque commanded, he could be seen as well as heard. There was of course no way of knowing whether you were being watched at any given moment. How often, or on what system, the Thought Police plugged in on any individual wire was guesswork. It was even conceivable that they watched everybody all the time. But at any rate they could plug in your wire whenever they wanted to. You had to live--did live, from habit that became instinct--in the assumption that every sound you made was overheard, and, except in darkness, every movement scrutinized.

Class Exercise III

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?

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 47Automatic Zoom

DEEP NEURAL NETWORKS CAN DETECT SEXUAL ORIENTATION FROM FACES

1 THIS IS A PREPRINT OF THE PEER REVIEWED ARTICLE TO APPEAR IN JOURNAL OF

2 PERSONALITY AND SOCIAL PSYCHOLOGY.

3

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

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

6

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

8 images

9

10 Yilun Wang, Michal Kosinski

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

12 michalk@stanford.edu

13

14

15 *The study has been approved by the IRB at Stanford University*

16

17 Citation: Wang, Y., & Kosinski, M. (in press). Deep neural networks are more accurate than

18 humans at detecting sexual orientation from facial images. *Journal of Personality and*

wang_kosinski.pdfDownload preprintVersion: 10

Download preprintDownloads: 24829

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

No 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

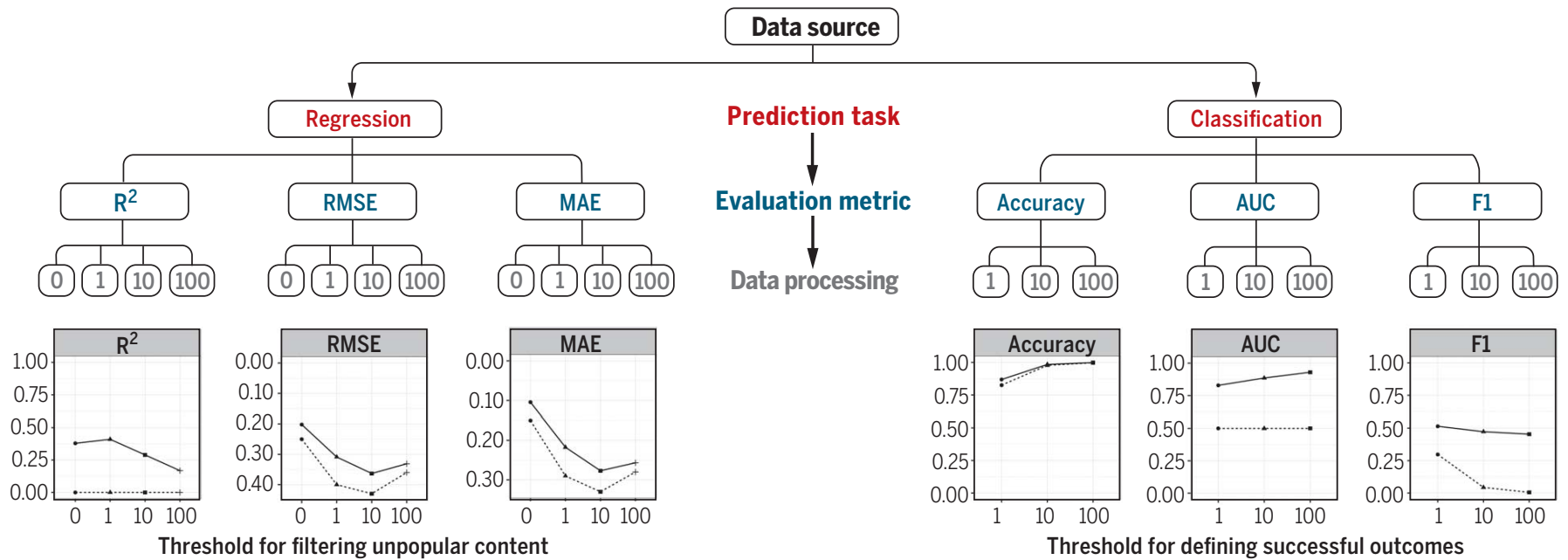
Wang, Y., & Kosinski, M. (2017, October 16). Deep neural networks are more accurate than humans at detecting sexual orientation from facial images.. Retrieved from psyarxiv.com/hv28a

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

Summary (Contd.)



What are the limits of prediction? Can they fail?

BIG DATA

The Parable of Google Flu: Traps in Big Data Analysis

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

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



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 DeMasi¹, Konrad Kording^{2, 3}, and Benjamin Recht¹

¹Department of Electrical Engineering and Computer Sciences, University of California Berkeley, Berkeley, CA, USA

²Department of Bioengineering, University of Pennsylvania, Philadelphia, PA, USA

³Department 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.

Class Exercise IV

Assess whether each of the following cases is an easier prediction or a harder one. (Example cases in the handout sheet)

...n terms that match the propen-
...but are structurally unrelated,
...t predict the future, were quite
...developers, 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
...ata were overfitting the small
...ses, a standard concern in data
...s ad hoc method of throwing
...search terms failed when GFT
...missed 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.

Class Exercise V

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