Class slides for Thursday, October 8: Social fads, part 2

Matthew J. Salganik

COS 597E/SOC 555 Limits to prediction
Fall 2020, Princeton University

Can Cascades be Predicted?

Justin Cheng
Stanford University
jcccf@cs.stanford.edu

Lada A. Adamic
Facebook
ladamic@fb.com

P. Alex Dow
Facebook
adow@fb.com

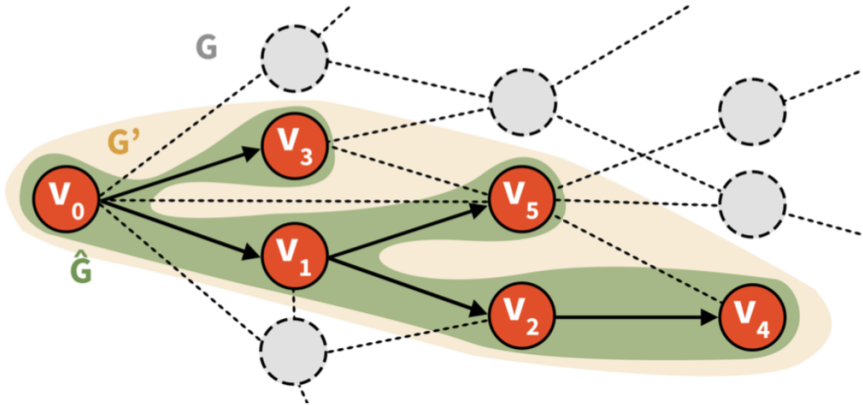
Jon Kleinberg
Cornell University
kleinber@cs.cornell.edu

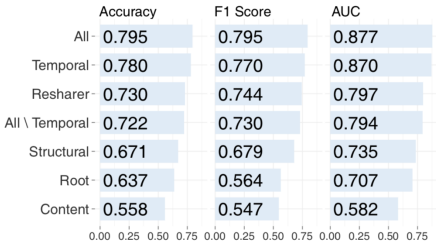
Jure Leskovec
Stanford University
jure@cs.stanford.edu

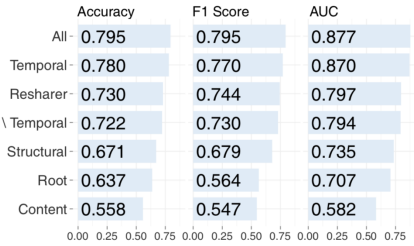
- *cascade growth prediction problem:* given a cascade that currently has size k, will it grow beyond the median size $f(k)$?

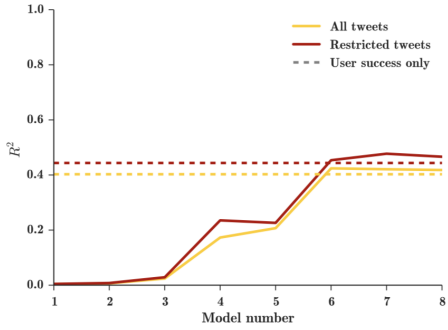
- *cascade growth prediction problem:* given a cascade that currently has size k, will it grow beyond the median size $f(k)$? In this setting, this is equivalent in this setting to: given that a cascade has size k will it grow beyond $2k$?

- *cascade growth prediction problem:* given a cascade that currently has size k, will it grow beyond the median size $f(k)$? In this setting, this is equivalent in this setting to: given that a cascade has size k will it grow beyond $2k$?
  - prediction problems that are balanced
  - creates one prediction problem for each k
  - it closely approximates real tasked needing to be solved for managing viral content

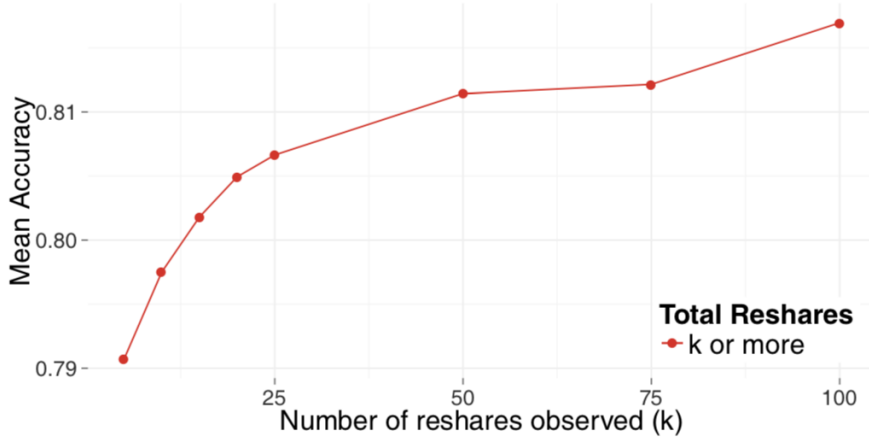

G
G'
Ĝ
V0
V1
V2
V3
V4
V5


Accuracy
F1 Score
AUC
All
Temporal
Resharer
All \ Temporal
Structural
Root
Content
0.795
0.795
0.877
0.780
0.770
0.870
0.730
0.744
0.797
0.722
0.730
0.794
0.671
0.679
0.735
0.637
0.564
0.707
0.558
0.547
0.582
0.00 0.25 0.50 0.75
0.00 0.25 0.50 0.75
0.00 0.25 0.50 0.75

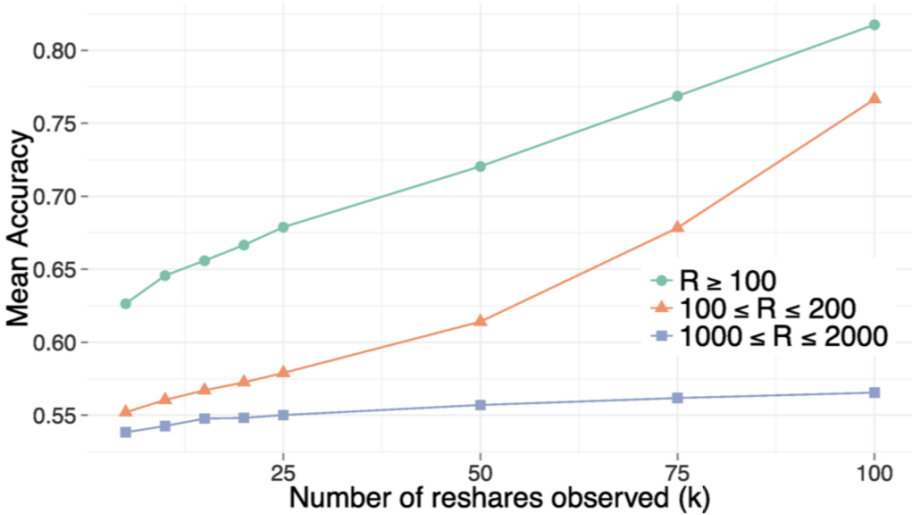




One (set) of pretty basic features does about as well as lots of features. Does this imply that we are close to a limit?



How could we apply this same idea to other domains?



How could we apply this same idea to other domains?

Can Cascades be Predicted?

Justin Cheng
Stanford University
jcccf@cs.stanford.edu

Lada A. Adamic
Facebook
ladamic@fb.com

P. Alex Dow
Facebook
adow@fb.com

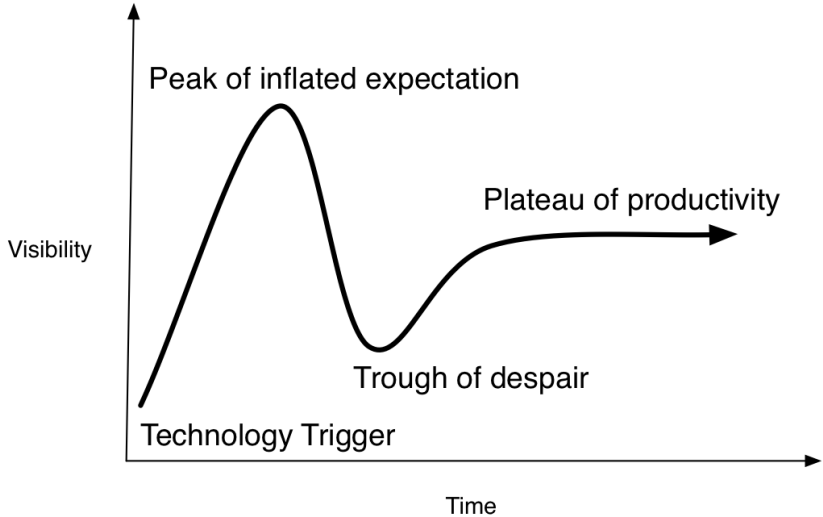
Jon Kleinberg
Cornell University
kleinber@cs.cornell.edu

Jure Leskovec
Stanford University
jure@cs.stanford.edu

Predicting consumer behavior with Web search

Sharad Goel[1], Jake M. Hofman[1], Sébastien Lahaie[1], David M. Pennock[1], and Duncan J. Watts[1]

Microeconomics and Social Systems, Yahoo! Research, 111 West 40th Street, New York, NY 10018



https://en.wikipedia.org/wiki/Hype_cycle

Detecting influenza epidemics using search engine query data

Jeremy Ginsberg[1], Matthew H. Mohebbi[1], Rajan S. Patel[1], Lynnette Brammer[2], Mark S. Smolinski[1] & Larry Brilliant[1]

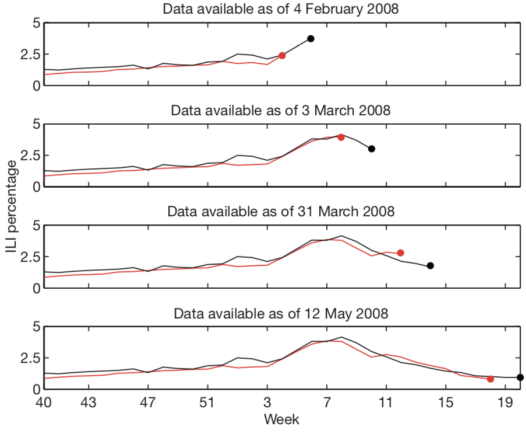


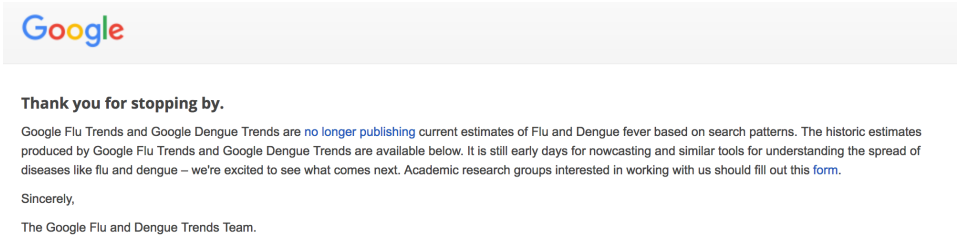
Figure 3 | ILI percentages estimated by our model (black) and provided by the CDC (red) in the mid-Atlantic region, showing data available at four points in the 2007-2008 influenza season. During week 5 we detected a sharply increasing ILI percentage in the mid-Atlantic region; similarly, on 3 March our model indicated that the peak ILI percentage had been reached during week 8, with sharp declines in weeks 9 and 10. Both results were later confirmed by CDC ILI data.

Predicting the Present with Google Trends

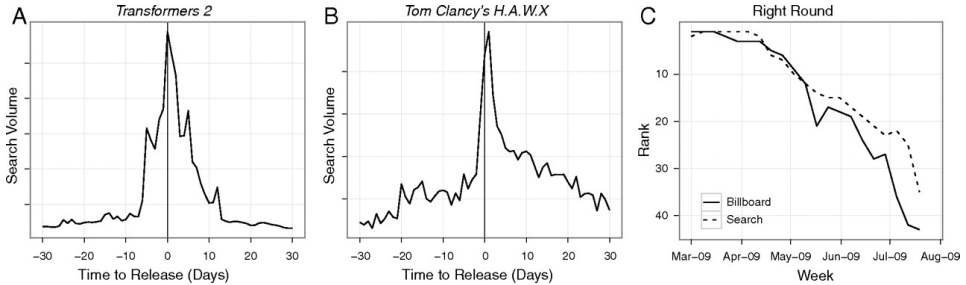
HYUNYOUNG CHOI and HAL VARIAN
Google, Inc., California, USA

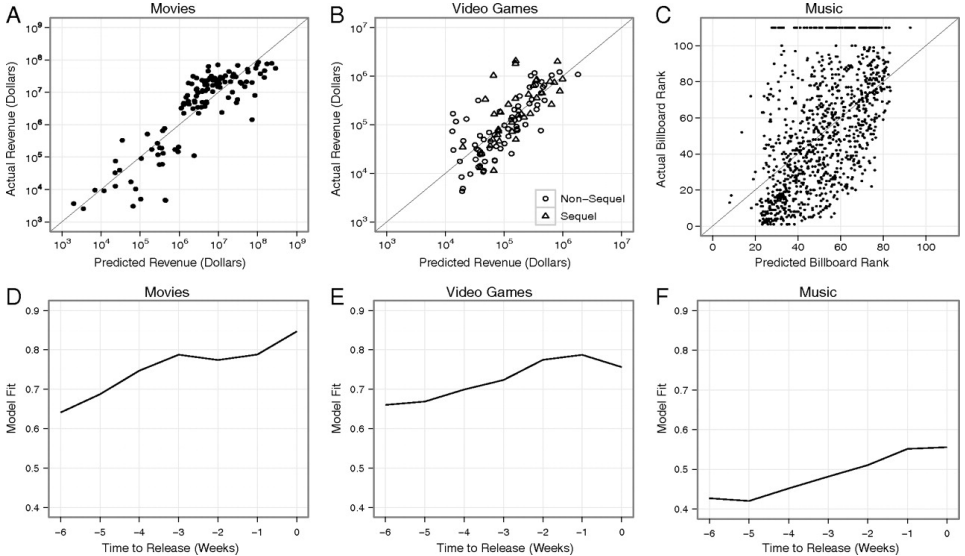
The Parable of Google Flu: Traps in Big Data Analysis

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

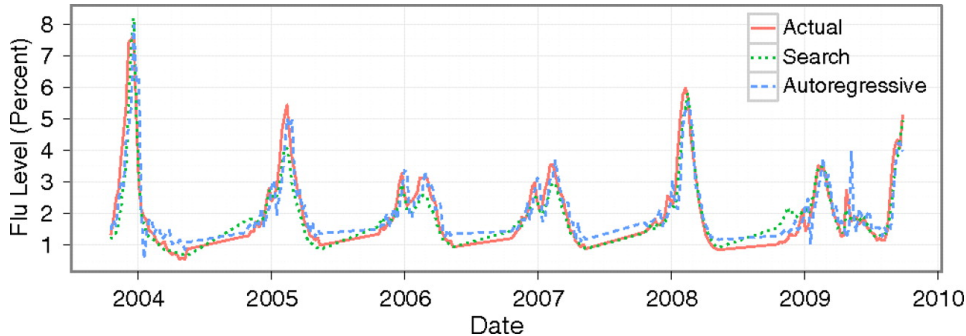


Nice history: https://www.theatlantic.com/technology/archive/2014/03/in-defense-of-google-flu-trends/359688/


A
Transformers 2
Search Volume
−30 −20 −10 0 10 20 30
Time to Release (Days)
B
Tom Clancy's H.A.W.X
Search Volume
−30 −20 −10 0 10 20 30
Time to Release (Days)
C
Right Round
Rank
10
20
30
40
Billboard
Search
Mar−09 Apr−09 May−09 Jun−09 Jul−09 Aug−09
Week

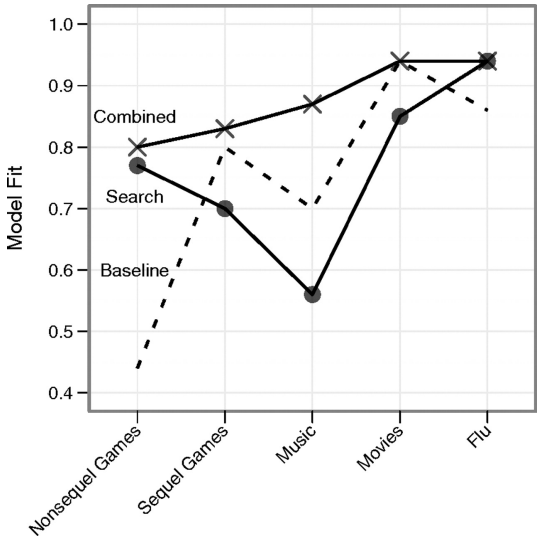


Note that measure of predictive performance is correlation. (same as Ginsburg)



correlation between prediction and truth

- simple auto-regressive model: 0.86
- Google Flu Trends: 0.94


1.0
0.9
0.8
0.7
0.6
0.5
0.4
Model Fit
Combined
Search
Baseline
Nonsequel Games
Sequel Games
Music
Movies
Flu

- For debunking: always compare to something simple

- For debunking: always compare to something simple
- Predicting the present is not the same as predicting the future

Predicting consumer behavior with Web search

Sharad Goel[1], Jake M. Hofman[1], Sébastien Lahaie[1], David M. Pennock[1], and Duncan J. Watts[1]

Microeconomics and Social Systems, Yahoo! Research, 111 West 40th Street, New York, NY 10018

Class slides for Thursday, October 8: Social fads, part 2

Matthew J. Salganik

COS 597E/SOC 555 Limits to prediction
Fall 2020, Princeton University